AGARD-CP-486

AGARD-CP-486

AGARD

ADVISORY GROUP FOR AEROSPACE RESEARCH & DEVELOPMENT
7 RUE ANCELLE 92200 NEUILLY SUR SEINE FRANCE

AGARD CONFERENCE PROCEEDINGS 486

Use or Reduction of Propagation and Noise Effects in Distributed Military Systems

(Utilisation ou Réduction des Effets de la Propagation et du Bruit dans les Systèmes Militaires Distribués)

NORTH ATLANTIC TREATY ORGANIZATION

AGARD-CP-486

AGARD

ADVISORY GROUP FOR AEROSPACE RESEARCH & DEVELOPMENT
7 RUE ANCELLE 92200 NEUILLY SUR SEINE FRANCE

AGARD CONFERENCE PROCEEDINGS 486

Use or Reduction of Propagation and Noise Effects in Distributed Military Systems

(Utilisation ou Réduction des Effets de la Propagation et du Bruit dans les Systèmes Militaires Distribués)

Papers presented at the Electromagnetic Wave Propagation Panel Symposium, held in Rethymno, Crete, Greece, 15th to 18th October 1990.

North Atlantic Treaty Organization
Organisation du Traité de l'Atlantique Nord

The Mission of AGARD

According to its Charter, the mission of AGARD is to bring together the leading personalities of the NATO nations in the fields of science and technology relating to aerospace for the following purposes:

- Recommending effective ways for the member nations to use their research and development capabilities for the common benefit of the NATO community;
- Providing scientific and technical advice and assistance to the Military Committee in the field of aerospace research and development (with particular regard to its military application);
- Continuously stimulating advances in the aerospace sciences relevant to strengthening the common defence posture;
- Improving the co-operation among member nations in aerospace research and development;
- Exchange of scientific and technical information;
- Providing assistance to member nations for the purpose of increasing their scientific and technical potential;
- Rendering scientific and technical assistance, as requested, to other NATO bodies and to member nations in connection with research and development problems in the aerospace field.

The highest authority within AGARD is the National Delegates Board consisting of officially appointed senior representatives from each member nation. The mission of AGARD is carried out through the Panels which are composed of experts appointed by the National Delegates, the Consultant and Exchange Programme and the Aerospace Applications Studies Programme. The results of AGARD work are reported to the member nations and the NATO Authorities through the AGARD series of publications of which this is one.

Participation in AGARD activities is by invitation only and is normally limited to citizens of the NATO nations.

Published June 1991

ISBN 92-835-0601-4

Printed by Specialised Printing Services Limited
40 Chigwell Lane, Loughton, Essex IG10 3TZ

Preface

Modern telecommunications are aimed at improving the safety and the data rate of messages, in particular by means of digital techniques, but also by increasing redundancy through the use of distributed systems.

The main feature of distributed systems is that transmitted and/or received signals are distributed into the Space-Time-Frequency space, at least along one dimension.

Signal processing capability has evolved quickly during the last two decades, following the development of D.S P. components, which allow fast and accurate treatment

New methods have made possible the handling of complex problems such as: receiving in large antenna arrays of which OTH radars are the best example, spread spectrum techniques, which greatly improve transmission viability, adaptive systems which rely on time distribution in order to mitigate multipath effects, and multistatic radars, able to detect stealth targets.

The objective of this Symposium has been to gather the best specialists in these fields, where propagation effects are present in the design of the systems

The present editing contains the texts of the lectures, along with following discussions or comments.

I would like to thank the Session Chairmen who had to direct the discussions, and the Technical Committee Members for the important work they have accomplished

I am very grateful to the Greek Authorities, the National Delegate, and the Local Coordinator and his staff, for their friendly welcome and for the excellent organization of the Symposium

I must also express my thanks to the EPP Administrator and his Secretary who provided, before, during, and after the Symposium, precious and appreciated support.

Finally, I cannot forget my colleagues, Dr. J.H. Richter, and the Ingénieur Général de l'Armement P. Fuerxer, Deputy Chairmen of this Symposium, with whom I was so pleased to work.

C Goutelard
Deputy Chairman
Co-Editor

Préface

Les télécommunications modernes visent à améliorer la rapidité de transmission et la protection des messages notamment par les techniques numériques mais aussi en augmentant la redondance par l'utilisation de systèmes distribués.

Les systèmes distribués doivent être caractérisés par le fait que l'émission et/ou la réception des signaux est répartie selon une dimension, au moins, de l'espace minimum qui permet de les représenter: Temps-Fréquence-Espace.

Les possibilités de traitement du signal ont évoluées rapidement depuis deux decennies grâce aux composants qui permettent d'effectuer des traitements numériques donc précis et rapides. De nouvelles méthodes ont donc permis de traiter des systèmes complexes, tels que la réception avec de grands réseaux d'antennes dont les radars transhorizon sont les exemples les plus clairs, les techniques d'étalement de spectre d'où découle une grande sécurité de transmission, les systèmes adaptatifs qui s'appuient sur la distribution temporelle pour lutter contre les trajets multiples ou enfin les radars multistatiques susceptibles d'assurer une riposte aux cibles furtives.

L'objectif de ce symposium a été de réunir les meilleurs spécialistes de ces domaines où l'influence de la propagation est présente dans toute conception de systèmes.

Cette édition reproduit les textes des conférences présentées ainsi que les discussions qui les ont accompagnées.

Je tiens à remercier les présidents de séances qui ont eu à conduire les discussions et les membres du comité technique pour l'important travail qu'ils ont accompli.

Que les autorités grecques, le délégué national, le coordonnateur local et son équipe acceptent l'expression de notre reconnaissance pour l'accueil si sympathique que nous avons reçu et l'excellente organisation de ce symposium.

Mes remerciements s'adressent aussi à l'administrateur de l'EPP et à son secrétaire qui ont apporté avant, pendant et après le symposium, une aide précieuse fort appréciée.

Comment pourrais-je oublier enfin dans ces remerciements mes collègues, le Docteur J.H. Richter et l'Ingénieur en Chef de l'Armament P. Fuerxer, co-Président de ce symposium avec qui j'ai eu tant de plaisir à travailler.

C. Goutelard
Co-Président
Co-Editeur

Theme

Modern military systems must be resistant to electronic countermeasures, have high reliability, and continuously assure high quality functions. Distributed systems are more and more considered as an interesting possibility, among them: Space Distribution, Time or Space-Time Distribution, and Frequency Distribution.

When dealing with these techniques, effects of the medium on the coherence of the electromagnetic waves, non-homogeneities, variability and dispersivity of the medium, must be taken into account. A better understanding of the physical phenomena will lead to better technical solutions and will help push back the frontiers of current technical limitations. These techniques apply to high reliability communications, radar systems and satellite systems for radio localisation and range finding. For military applications, modern systems largely rely on distributed systems, because of the additional benefits they bring, e.g. counter jamming and performance improvement. Sophisticated signal processing enables the development of these techniques, which provide military systems with the reliability, the hardening and the quality they require. In addition, these systems are able to survive in the battlefield environment.

The topics covered include:

1. Limitation of the medium,
2. Jamming and noise reduction effects on distributed systems,
3. Distributed systems.

Thème

Les systèmes militaires modernes doivent être résistants aux contremesures électroniques, faire preuve d'une grande fiabilité et assurer en permanence des fonctions de premier ordre. Les systèmes distribués sont de plus en plus considérés comme une option intéressante et en particulier : la distribution dans l'espace, la distribution dans le temps, la distribution dans l'espace/temps et l'attribution de fréquences.

En ce qui concerne ces techniques, il faut tenir compte des effets du milieu de propagation sur la cohérence des ondes électromagnétiques, ainsi que des non-homogénéités, et de la variabilité et de la dispersivité du milieu. Une meilleure compréhension du phénomène physique amènera des solutions techniques plus performantes, tout en faisant reculer les frontières imposées par les contraintes techniques actuelles.

Ces techniques s'appliquent aux réseaux de télécommunications à haute fiabilité qui sont les systèmes radar et satellite de radiolocalisation et de télémétrie. Pour les applications militaires, les systèmes modernes sont généralement des systèmes distribués en raison des avantages qui y sont associés : l'anti-brouillage et l'amélioration des performances. Ces techniques, qui garantissent aux systèmes militaires la fiabilité, le durcissement et la qualité dont ils ont besoin, doivent leur existence a des systèmes sophistiqués de traitement du signal. En outre, ces systèmes sont adaptés aux conditions de survie caractéristiques du champ de bataille.

Parmi les sujets examinés on distingue:

1. La limitation du milieu de propagation.
2. Les effets du brouillage et de la réduction du bruit sur les systèmes distribués.
3. Les systèmes distribués.

Electromagnetic Wave Propagation Panel

Chairman: Ir H. Vissinga
van Kempenstraat 30
2252 VH Voorschoten
The Netherlands

Deputy Chairman: Dr J.H. Richter
Head, Ocean and Atmospheric Sciences Division
Code 54
Naval Ocean Systems Center
San Diego, CA 92152-5000
United States

TECHNICAL PROGRAMME COMMITTEE
CO-CHAIRMEN

ICA P. Fuerxer
Chef du Groupe 2
Télécommunications et Détection
DRET
26, Bd Victor
75996 Paris Armées
France

Prof. C. Goutelard
Directeur du L.E.T.T.I.
Université Paris-Sud
9, Avenue de la Division I eclerc
94230 Cachan
France

Dr J.H. Richter
Head, Ocean and Atmospheric Sciences Division
Code 54
Naval Ocean Systems Center
San Diego, CA 92152-5000
United States

MEMBERS

Dr J.S. Belrose
P.O. Box 11490
Communications Research Ctr
Station H
Ottawa, ON K2H 8S2
Canada

Dr D Rother
Standard Elektrik Lorenz AG
GLP/ETF
Ostendstrasse 3
D-7530 Pforzheim
Germany

Dr P.S. Cannon
Flight Management
P 161 Building
Royal Aerospace Establishment
Farnborough, Hants GU14 6TD
United Kingdom

Prof. G. Tacconi
Universita' di Genova
Dip. Ingegneria Biofisica
Via all'Opera Pia 11/A
16145 Genova
Italy

Mr G.H. Hagn
Assistant Director, Information
and Telecom. Sciences Center
SRI International
1611 N. Kent Street
Arlington, VA 22209-2173
United States

Dr D. Yavuz
Head, Radio Branch
Communications Division
SHAPE Technical Center
P.O. Box 174
2501 CD The Hague
The Netherlands

PANEL EXECUTIVE

Lt Col. P.Brunelli

Mail from Europe:
AGARD-OTAN
Attn: EPP Executive
7, rue Ancelle
92200 Neuilly sur Seine
France

Mail from USA and Canada:
AGARD-NATO
Attn: EPP Executive
APO New York 09777

Telephone: 33(1)47385768 Telex: 610176 (France) Telefax. 33(1)47385799

Contents

* Printed in classified publication CP 486 (Supplement)

SESSION VI – DISTRIBUTED SYSTEMS
Session Chairman: Dr J.H. Richter

THE DYNAMICS OF F-LAYER IRREGULARITIES RELATIVE TO SPACE, TIME AND FREQUENCY DIVERSITY

by

Dr Jules Aarons
Center for Space Physics
Boston University
725 Commonwealth Avenue
Boston, MA 02215
United States

ABSTRACT

With the statistics of morphology of F-layer irregularities now in hand, it is possible to forecast in broad terms what to expect at equatorial, auroral and polar latitudes during various levels of solar flux. With the beginning of an understanding of the effect of the various phases of magnetic storms on generating irregularities as noted from the solar wind, ring current, convection, auroral index, and magnetic index parameters, it is possible to roughly forecast levels of F-layer irregularity intensity. With these in hand, the utility of space, time, and frequency diversity can be evaluated. Diversity could be used if forecasting in real time was possible. This study outlines the dynamics of irregularity generation and inhibition during various phases of ring current and magnetic activity, leaving the detailed use of diversity methods to the system operators. For the system operators, the duration and severity of the effects must be evaluated for the particular location.

INTRODUCTION

Irregularities at F-layer heights produce phase and amplitude fluctuations on signals traversing the ionosphere. They therefore have an effect on radio systems designed, developed or envisaged which utilize frequencies from 20 MHz to 6 GHz. If the system is in the field, the using organization is usually left to determine the need and utility of correction methods. The using organization has to develop the day to day perturbations that have an effect on the system. In the case of dealing with the problems that irregularities produce (fading, loss of signal, false targets) the designer of the equipment has to be able to supply methods of minimizing these effects. This has to mean a decrease in capability. It is there that the evaluation of frequency, time or space diversity has to be made.

For forecasting ionospheric propagation parameters we have two distinct needs i.e. long term statistical effects and dynamic effects. The planner is only concerned with the effects of magnetic storms such as were observed March 13 and 14, 1989 as a piece of his statistics, perhaps the 99 percentile bracket. Data and analyses are at hand to forecast the morphology of F-layer irregularities. Forecasting the morphology allows the evaluation of the utility of operating systems and the planning of back-ups.

However, the users of satellite communications at 250 MHz were dismayed on March 13 and 14, 1989 when auroral effects on transmission were noted as far from the auroral zone as Mexico and Puerto Rico. Forecasting the dynamics will allow the use of real time warning and possible use of diversity as a function of geophysical conditions. It allows the system to warn operators about impending problems. In this paper, we shall concentrate on two geophysical areas where it is possible to develop forecasting methods to determine the dynamics of irregularity generation; we shall look at the auroral and equatorial regions.

Using scintillation data from satellite transmissions, we shall note briefly the morphology of irregularities producing scintillation but direct the study to the development of the irregularity dynamics as a function of time and latitude.

GLOBAL PATTERNS OF F-LAYER IRREGULARITIES

1. HIGH LATITUDE MORPHOLOGY AND DYNAMICS

a. Morphology

We currently have a good idea as to the level of phase and amplitude scintillations and spread F in the regions of importance. Figure 1 is a cartoon of the nighttime F-layer irregularities affecting scintillation during years of high solar flux. The word cartoon is used to reflect the broad brush somewhat inaccurate nature of the picture. It is an updated version of similar global maps presented in the past (Aarons, 1982). It does establish the importance of the auroral, the polar, and the equatorial regions for this parameter. We know quite a bit about the statistics of F-layer irregularities in the auroral and equatorial regions, the areas to be discussed in this paper. Excellent references on the occurrence and intensity of both phase and amplitude fluctuations are given in Su. Basu et al, 1987 and Su. Basu et al, 1988.

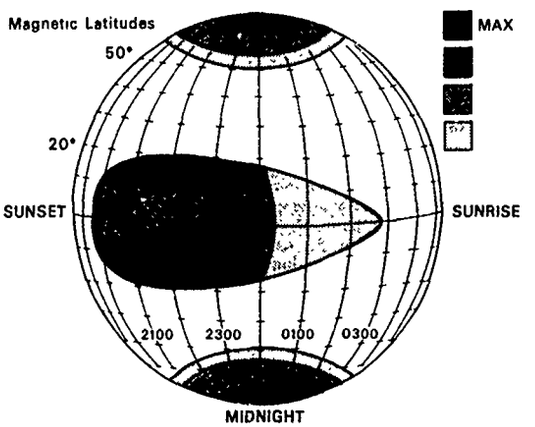


Fig. 1

Figure 1. The nighttime F-layer irregularity world in years of high solar flux This is an updated version of the scintillation activity as shown in Aarons (1982)

Much research has been done on the effect of magnetic activity on F-layer irregularities at auroral latitudes. At high latitudes, a review of the data indicates a rather complex picture with a seasonal, latitudinal and longitudinal dependence. In most of the studies (e.g. Rino and Matthews, 1980), the data are analyzed by sorting into magnetically disturbed and quiet periods of time depending on whether Kp or Ap is greater than or less than some particular value.

b. Dynamics

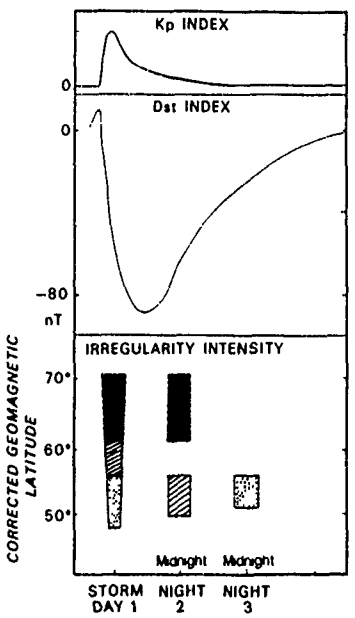
However, the effects of the magnetic storms on F region irregularities undoubtedly depend on the phases of the storms. For a particular observing point this means a correlation of the effects of storm development as a function of local time. This has been shown to be amply true in the case of total electron content obtained from satellite observations (Mendillo et al, 1972). We shall focus on the pattern of individual storm development.

From the modelling point of view, the irregularity data need a geophysical parameter around which to organize occurrence and intensity. For the auroral region the forcing function has been magnetic activity as shown by studies of low altitude satellites such as WIDEBAND and TRANSIT as well as earlier work with radio stars and synchronous satellites. Increased solar flux plays a role both pushing the F-layer irregularity region equatorwards in latitude and increasing the quiet day intensity.

In the study of the dynamics of the development of irregularities at auroral latitudes during a magnetic storm, F-layer irregularities are found to be generated during the injection phase of the magnetic storm, the strength and duration of which vary with the intensity and duration of the storm components. The effect during this intial phase of a magnetic storm is proposed to be the prompt electric field penetration to low latitudes due to the rate of motion of the equatorward edge of the auroral electric fields, before the shielding layer builds (Fejer et al, 1989b. and Foster and Aarons, 1988) When this penetration occurs, the conditions which produce the F-layer irregularities descend to what had been sub-auroral latitudes. When the electric field descends to lower latitudes there is a distinct time lapse along a particular longitude for the irregularities to form. The source weakens and descends to lower latitudes probably as a function of the electric field and the ambient ionospheric conditions extant at that time. There is no hole observed i.e. a region such as the trough where there are very weak irregularities, although our data does not definitively rule this out. The descent could take place during the day or during the night. During the night, the irregularity region under magnetically quiet circumstances has already descended to lower latitudes. Therefore, the latitude affected by the equatorward direction of the source can be lower than during the day.

b. Development of the Classic Geomagnetic Storm

Sugiura and Chapman (1960) have shown the Dst levels from classical magnetic storms; we have averaged and produced a curve of Dst in the diagram of Figure 2. We have added an idealized Kp, substituting this for the DS used by Suguira and Chapman. Although this is not an auroral electrojet index as suggested by Rostoker (1972) the use of Kp is justified on the basis that we are working with auroral, sub-auroral and equatorial latitudes as well as a history of correlation of irregularities with Kp. These curves indicate the immediate magnetic variations and the injection of ions into the ring current (Dst) either the result of substorm injections or Birkeland Region 2 currents. These occur in weak , moderate and great magnetic storms. However, it should be noted individual storms rarely appear as the average storm of Figure 2. From this classical picture there are two elements of interest. In the initial period of a magnetic storm, Kp and AE increase. Ions are injected into the ring current but there is most often a time lapse between the initial start of the storm and the time period of maximum energy in the ring current as measured by Dst.



MODEL OF IRREGULARITY INTENSITY DURING AND AFTER MAGNETIC STORM COMMENCEMENT

Fig. 2

Figure 2. High latitude development of F-layer irregularities according to the model of Rodgers and Aarons. During the initial injection phase of the magnetic storm, the electric field penetrates to lower latitudes. However, with the end of the magnetic storm, the ring current decays, producing the conditions for irregularity development at sub-auroral latitudes.

A descriptive model of the dynamic effects of magnetic storms on auroral and sub-auroral irregularities was developed by A.S. Rodger and myself. In Figure 2 we show the apparent motion of the irregularity effects of the magnetic storm, in reality the electric field penetration as a function of latitude. It first affects the auroral region with very close correlation between the form of magnetic activity and the timing of scintillation at auroral latitudes. As time proceeds, the effects descend to lower latitudes encompassing regions termed in quiet magnetic periods, sub-auroral latitudes. This is essentially shown in the 'storm day' which we term Day 1. During this injection phase, the ring current stores the ionospheric and solar ions. After the storm, the ring current decays and this decay is the source for the generation of sub-auroral irregularities - and Stable Auroral Red Arcs.

In this model of the time development of irregularities in the F region, the initial irregularities observed are noted in the auroral oval. Within a matter of hours, the irregularity development moves equatorwards. While this can take place during daytime hours, the presence of an E layer shorts out the irregularities in a matter of minutes to an hour. During night hours, the penetration of the electric field to lower latitudes takes place along with the development of irregularities.

During the recovery phase of the magnetic storm (Day 2 and at times Day 3), the effect of the decay of ring current ions on the sub-auroral region is such that irregularities of low intensities are formed without auroral or magnetic activity. The appearence of intense F-layer irregularities at sub-auroral latitudes over several days in the recovery phase of magnetic storms is in a complex manner correlated with the level and recovery pattern of Dst, the measure of ring current energy density (Rodger and Aarons 1988, Aarons and Rodger, 1989).

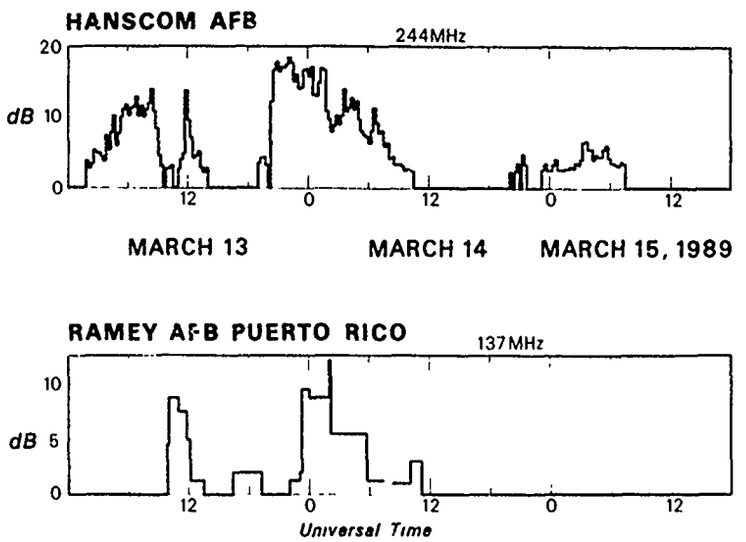
We shall illustrate (1) the time lapse of the descent during both night and day (2) the weakening and (3) the failure to find a latitude where the irregularity intensity (an irregularity trough) decreases during the this early period of magnetic storm development.

THE DESCENT AND WEAKENING OF THE IRREGULARITY GENERATING SOURCE

MARCH 13-15,1989

The great magnetic storms of March 13-15, 1989 indicated the depth of penetration and weakening of the electric field in an equatorward motion. In Figure 3a, scintillation data from Hanscom Field, MA (52.5° CGL) and from Ramey AFB, Puerto Rico with an ionospheric propagation intersection of 17° Geographic and 30.5° CGL are plotted. Longitudes of the two intersections differ by 2°.

As shown in Figure 3a, the Hanscom Field observations at 250 MHz show very high levels on March 13 and March 14 and lower levels on March 15. The Ramey AFB data at 137 MHz show high levels for the first two days but fail to show scintillations on March 15.



DESCENT AND WEAKENING

Fig. 3a

Figure 3.a. Scintillation activity for data taken from the Boston area and from Puerto Rico during the great magnetic storms of March 13-15, 1989.

For the night of March 13-14, the time when intense irregularities are first noted from the Boston area is approximately 2040 UT; the time when intense irregularities develop on the Puerto Rico path is 2315 UT. On the night of March 13-14, scintillation levels at 244 MHz reached 19 dB peak to peak for the path from Hamscom AFB. For the lower latitude, path levels were much lower at a lower frequency, reaching 9 dB except for a short burst.

The timing of the descent, i.e. essentially 2 hours 35 minutes to move 22° of latitude and the weakening as shown by the low levels on the Puerto Rico path on March 13-14 are indicative of the way the dynamic development takes place during the initial phases of the magnetic storm.

At Millstone Hill, MA, all sky observations of 6300 A aurora and SAR arcs are made routinely by the Boston University Mobile Ionospheric Observatory. For the night of March 13-14, when observations were allowable through clouds, saturated aurora were observed from 30°. to 55° Geographic North. In Figure 3b, we have plotted the meridional component of the all sky optical data for the next night, March 15. The meridional scans show continuous high levels from 01 to 0510 UT with a high level isolated patch noted from 0535 to 0550. On the Hanscom Field scintillation data in Figure 3a, high levels can also be noted in these time periods. The Ramey AFB data for March 15 do not show any scintillation activity. The electric field on this day did not penetrate to this latitude. For this night, irregularities and aurora reached (and passed) 53° CGL but failed to reach the 30.5° CGL intersection of Ramey AFB.

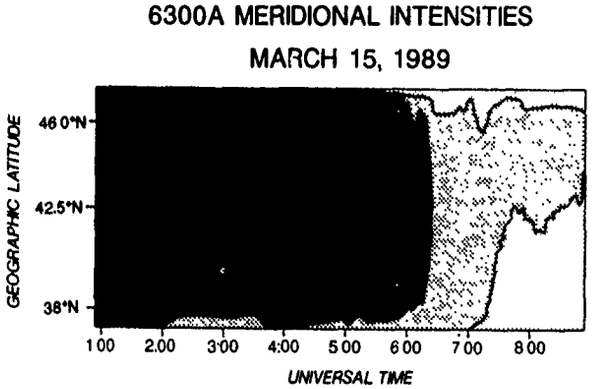


Fig. 3b

Figure 3.b. Meridional scan of 6300 A auroral emissions taken by the Boston University Mobile Ionospheric Observatory at Millstone Hill, MA. on March 15, 1989

SEPTEMBER 19-20, 1984

With observations from Millstone Hill Incoherent Scatter Radar, the electric field was determined from convection patterns for September 19-20, 1984 (Figure 4 from Foster and Aarons, 1988). In this illustration, a weak storm occurring during morning hours showed penetration equatorward and weakening of the electric field as the field moves to lower latititudes. Using the same satellite, GOES 2, transmitting at 137 MHz, peak to peak scintillation levels reached 16 dB on the path through 61° CGL but only 2.5 dB through 53°. In addition, a time difference develops between the onset of irregularities at 60° and at 53° CGL, a difference of at least 45 minutes for the 7° geographic latitude difference.

Thus we have an equatorward motion of the source that is weakened in its effect as it moves away from the auroral region.

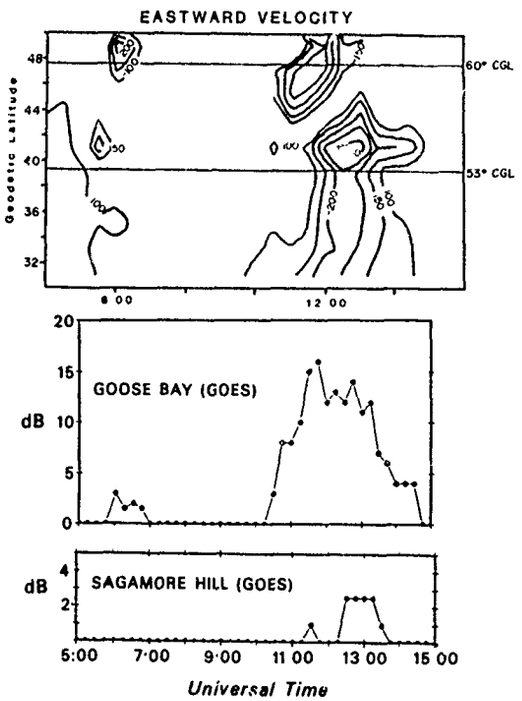


Fig. 4

Figure 4. Observations of convection and scintillation for the Dst excursion and electric field changes of September 19-20, 1984. The electric field penetration takes place later at the lower latitude of 53° CGL than at 61° CGL. In addition, the effects are considerably weakened.

2. THE EQUATORIAL REGION: MORPHOLOGY AND DYNAMICS

a. Morphology

Solar flux has a first order effect on equatorial irregularities. With high solar flux the plume regions show greater intensity and greater height. With high solar flux, the anomaly regions at the edges of the equatorial zone show high electron densities at sunset hours before irregularities appear and these high electron densities continue after sunset. When irregularities appear in the anomaly region intense microwave scintillations are observed during the pre-midnight time period. During years of low solar flux, the anomaly region fails to show such levels. Studies with synchronous satellites and with optical observations have allowed the morphology and nature of the irregularities in this region to be clearly catalogued (Su. Basu et al, 1988). Another 'forcing function' to be noted is the importance of the month of the year at various longitudes with a lowering of irregularity occurrence and intensity during December and January in the Pacific sector and a lowering of these levels in June in the $0-75^{\circ}$ West sector.

b. Dynamics

The data on the correlation of the magnetic activity and equatorial irregularities has been supplemented recently by research on individual storms. Among these, Dabas, et al, 1988, concluded that suppression of irregularities occurs if the recovery phase started during local daytime hours. On the night that follows, scintillation and spread F were suppressed completely or partially. Their conclusion about 'partially' may come about because it is difficult to identify irregularities coming from the west and developing under different circumstances. They used 52 events at Lunping (Magnetic or Dip Latitude of intersection 14.9°). The difficulty with this analysis is that Lunping is in the equatorial anomaly region and the absence of irregularities can be due to two possibilities (1) only a thin layer of irregularities was formed at the equator and therefore the height of the plume did not reach the altitudes necessary to produce irregularities on the field line through which Lunping makes the observations or (2) the absence of irregularities.

In the case of equatorial F-layer irregularities, several studies have shown correlation with the phase of the storm in the equatorial region (DasGupta et al, 1985). We shall focus on the pattern of the effects of the phases of the magnetic storms coupled with the local time when the ring current parameter, Dst, is maximum.

The concept that is being developed is as follows: if the maximum level of the ring current energy, as shown by the value of Dst, is reached in the early afternoon, then the F-layer height will be affected. The reduction of the normal height during months when irregularity activity is noted almost every night, essentially inhibits irregularity activity during the night. If the maximum Dst level occurs in the midnight to post-midnight time period then the layer height will fall and irregularities will be produced. It should be noted that these concepts follow from early work in spread F and scintillation data. The difference is the use of the Dst maximum as the keynote in the effect and in careful use of the local time.

c. DECEMBER 20, 1980 MAGNETIC STORM

An illustration of the effects of the ring current can be shown by using our own and published observations of irregularities in the equatorial region for the magnetic storm of December 19-20, 1980. This is shown in Figure 5. The two top panels show Kp for the period and Dst. The data set illustrates the inhibition of irregularities when the maximum Dst is in the noon to afternoon time period, the increase in irregularity activity when the maximum Dst excursion is in the midnight to post-midnight time period and the absence of an effect on irregularities when the Dst maximum excursion occurs after sunset.

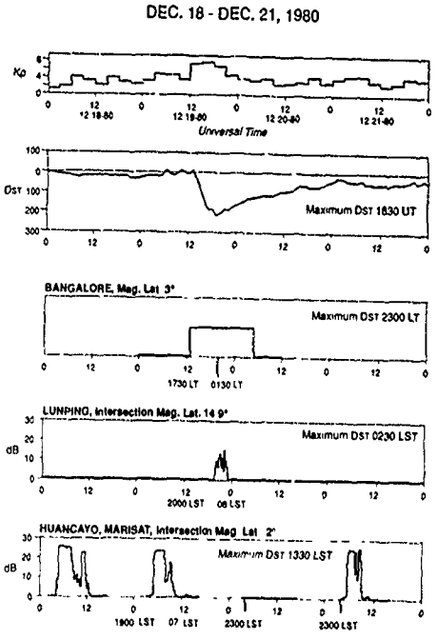


Fig. 5

Figure 5 Irregularity data for December 18-21, 1980 for various longitudes. At Bangalore, India at 3° dip latitude, the maximum Dst excursion was too late to have an effect on F-layer height; scintillations were observed. At Lunping, near the equatorial anomaly, the maximum excursion was in the post-midnight time period; scintillations were observed. At Huancayo where the maximum excursion was in the early afternoon, irregularities were inhibited. Bangalore data taken from Dabas et al (1988), and Lunping data taken from Huang and Chen(1980).

Kp (Panel 1) shows activity started at 03-06 UT on December 19, 1980 when Kp reached 5. Dst went positive reaching a value of +10 nT at 05 UT and then going to +12 nT at 1200 UT. However Dst reached its maximum value of -240 nT from 1800-1900 UT. Thus there is separation of the two indices Kp and Dst. For this example, the injection and recovery stages are clearly outlined with little magnetic or ring current activity before or after the storm.

The effect on the F-layer irregularities is shown by noting the Local Standard Time of maximum ring current energy. Data from Huancayo, Peru (Panel 5) indicate that during the magnetically quiet and low Dst night of December 18-19 scintillations were present. On December 19, the maximum Dst for this longitude occurred at 1300 LST and stayed greater than -200 nT until 1700 LST. The excursion, therefore, was enough to inhibit the development of scintillation that night. This serves as an illustration of the inhibition of irregularities when the maximum Dst occurs in the pre-sunset time period.

From data published in Dabas et al, 1988, for Bangalore, India at 3° from the magnetic equator, scintillations started at the post-sunset time of 1800 Local Time about the same time that Dst began to increase. Ring current maximum was at 2300 LT. There was no effect of the excursion of Dst since the high values of Dst occurred after the start of scintillation activity. This is an illustration of the absence of effects when Dst maximum is reached after sunset. It should be noted that the Kp changes occurring in local morning and before and during the sunset time period had little effect.

In Panel 3 in the post midnight time period, the Lunping, Taiwan station with the ionospheric intersection to the satellite ETS-2 of 14.9° recorded scintillations (Panel 4) in their logs (Huang et al, 1981). It was the only time in December that scintillations were recorded. (December in this longitude region normally shows a low occurrence of irregularities). On this night, irregularities were recorded beginning at 0330 LST and continuing through 0645 LST; This is an illustration of the post-midnight generation of irregularities.

To summarize: the Huancayo data shows inhibition, the Bangalore observations with maximum Dst occurring post-sunset failed to show an effect on irregularity generation and the Lunping scintillation activity occurring in the post-midnight time period showed irregularity activity when maximum Dst occurs in the midnight to post-midnight time period.

While we attempt to demonstrate one hypothesis for the inhibition of irregularities on equatorial paths, we do not state that this is the only reason for the absence of irregularities during periods of high occurrence. Localized effects produced by neutral winds or by subtle pattern changes in the electric field for a particular region produce conditions under which irregularities are not generated.

2. THE MAGNETIC STORMS OF MARCH 4-9, 1981

The complexities of using ring current and Kp levels to forecast observations are shown in the data for the period from March 4-9. A great magnetic storm took place on March 5, 1981 with Kp reaching 7+ in the time period 1500-1800 UT. Dst reached -215 nT at 1700 UT.

Huancayo Data

Before the storm took place on the nights of March 3-4 and March 4-5 scintillation activity showed its normal equinox pattern as illustrated in Figure 6. The large excursion of -215 nT took place at 1200 LST on March 5, decreasing to -173 nT at 1600 LST on March 5. These pre-sunset levels meant that the effect on the electric field at the equator was to inhibit the normal rise in height of the F-2 layer. Scintillations were not observed on the Huancayo MARISAT propagation path on the night of March 5-6.

Decay was slow reaching -99 nT at 0700 LST on March 6. Another storm took place on March 7 with Kp maxima at 1000-1600 LST and a further excursion in Dst of -90 nT at 1200 LST on March 7th. This level produced the same effect as the earlier inhibition of irregularities; no scintillations were noted on the night of March 7-8th.

Manila Data

The data from Manila for March 5-6 with a difference in time of 13 hours relative to Huancayo indicates the effect of increased Dst during the midnight-post midnight time period. The night of March 4 with relatively quiet conditions showed scintillation activity beginning at 2000 LST. The next night on March 5 with the magnetic storm in progress showed inhibition of post- sunset scintillations but very high levels of scintillations in the time period after midnight.

Dst for the storm on March 7th (with an excursion of -90 nT) peaked at 1700 UT (01 LST) for Manila. However there was no post-midnight increase in scintillation activity. It should be noted however that the Dst levels for this period were high throughout March 7th and March 8th ranging from a low of -40 nT to many values in the 60s and 80s. Kp values for the 1700 LST to 04 LST for the UT day of March 7 ranged from 4 to 5.

On March 8th scintillations were inhibited with a slow decay and an injection around sunset on March 8th for the Manila longitude.

Summary:

The Huancayo data showed inhibition with pre-sunset maximum excursions in two cases. The Manila data showed an increase in level in the post-midnight time period on March 5th. With large excursions through the period, the March 7 excursion in Dst failed to show increased post-sunset activity. The complexity of behavior during periods of injection and recovery made this one time period for this location difficult to categorize.

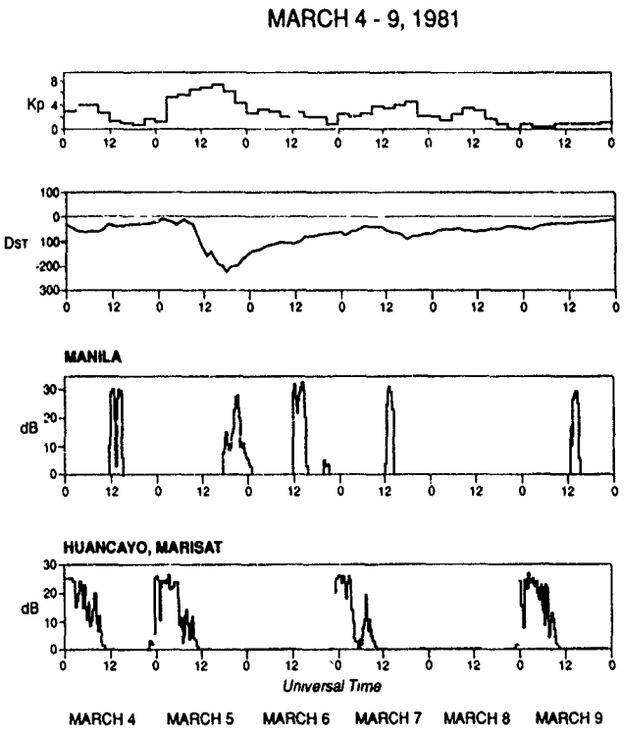


Fig. 6

Figure 6. March 1981 For this magnetic storm, Huancayo showed inhibition of irregularities on the night of March 5-6 with maximum DST excursion at 1500 LST while Manila with maximum Dst excursion in the post-midnight time period showed increased activity.

DISCUSSION

1. AURORAL ELECTRIC FIELD PENETRATION

The effect during the injection phase of a magnetic storm insofar as the sub-auroral latitudes are concerned is proposed to be the prompt electric field penetration to sub-auroral latitudes due to the motion of the equatorward edge of the auroral zone electric fields. On September 19-20, 1984 (Figure 4) south of Millstone Hill, a strong southward component of the F region neutral wind was observed and convection reversal was noted. This followed the model of Spiro et al (1988) who found storm enhanced neutral winds at latitudes equatorward of the shielding layer before the shielding layer was set up (Fejer et al, 1989, Foster and Aarons, 1988). Thus one expects the increase of irregularities at these latitudes during the initial phases of a magnetic storm with the irregularities appearing within hours of their appearance at auroral latitudes. Blanc et al (1983) after studying several events concluded that large southward values of Bz and large values of the rate of energy injection into the ring current were associated with the extension of magnetospheric convection electric fields to mid-latitudes.

As for the actual generation of conditions for the creation of auroral irregularities there have been several mechanisms which have been validated. For auroral latitudes, phenomena correlated with irregularities include precipitation (Basu et al, 1983), shear effects (Basu et al, 1984), convection of plasma containing irregularities (Weber et al 1985; Basu et al, 1988) as well as other factors in instability processes (Hudson and Kelley, 1976). For both auroral and sub-auroral latitudes recent studies, in progress, indicate that irregularity intensity increases when high ion convection is observed on incoherent scatter radars.

2. SUB-AURORAL LATITUDES

With the storm recovery in those cases when magnetic activity dwindles to low levels, sub-auroral irregularities can be shown to be relatively high and are correlated with Dst (Aarons and Rodger, 1988). These data suggest that the sub-auroral irregularities observed in the recovery phase of magnetic storms are the result of energy stored in the ring current which is then slowly released.

The mechanism suggested for the generation of SAR arcs which are closely correlated with the occurrence of sub-auroral irregularities involve ring current ions. Charge exchange, heat transfer to electrons or Coulomb collisions from low energy (< 20 kev) heavy ions (oxygen, nitrogen, helium) have been discussed by Kosyra et al, 1987.

The effects of the ring current are difficult to assess. In part this is due to the asymmetry of the ring current. In addition, the composition changes from storm to storm. During both great and lesser magnetic storms, the proton flux increases. However, protons do not increase as much as O+ and N+ (Lui et al, 1987). Thus the interaction of ring current ions with the ionosphere in the charge exchange process is rather difficult to assess. The asymmetry, the ion composition, and the energy of the ions make the process of interaction difficult to evaluate for any particular storm. While the ring current appears to be the source of much of the sub-auroral F layer irregularities, the Dst measure may not reflect in detail the development of those parameters of importance to irregularity generation. The Dst measure in the February 6-9, 1986 storm was well correlated with the inner ring current energy density from storm maximum well into recovery but the correlation was not good during the long developing main phase of this storm (Hamilton et al, 1988).

3. EQUATORIAL LATITUDES

The predominant flow in the ever present ring current is westward, which is responsible for the depression in horizontal intensity. Equatorial eastward electric fields produce an upward vertical drift during the day. Increased excursion (an increasing westward current in the ring current) decreases the eastward electric field before sunset which has produced the rising of the F-layer.

After midnight, the sudden and large increase of ring current produces an immediate rise of the F-layer in the midnight to post-midnight time period. The layer height then falls and irregularities are generated.

Rastogi and Aarons (1980) have observed that a reversal of the vertical drift to the upward direction during the night is followed by strong VHF scintillations near the magnetic equator. This had been noted earlier on spread F data. In a paper with many cases of storms individually analyzed, DasGupta et al (1985) reported that post-midnight VHF scintillations are related to the maximum negative excursion of the horizontal intensity of the magnetic field occurring in the 0000-0600 local time interval whereas no such relation is observed with the pre-midnight excursions.

CONCLUSIONS

For the high latitudes the developing electric field is first observed at auroral latitudes then descends to lower latitudes and weakens. The irregularities developed during this period have been observed in this set of measurements as far south as 30.5 CGL during some great magnetic storms.

In the recovery part of the storm, the ring current ions decay. The effect of the ions on the ionosphere is to produce irregularities at plasmapause latitudes. During the recovery phase, when irregularities at the auroral latitude become weak both in strength and duration, irregularities in the sub-auroral latitudes are found to be of the intensity noted during the injection phase of the storm. These irregularities while relatively intense for these latitudes are still much less intense than the injection phase auroral levels.

In the equatorial region, the maximum level of the ring current, the time just before recovery sets in, appears to be the most relevant time for the inhibition or generation of irregularities. If the maximum is reached before sunset, the ring current has the effect of decreasing the altitude of the F-layer which in turn inhibits the generation of irregularities. If the maximum level of ring current occurs after sunset but before midnight, there is no inhibition of irregularities. If the maximum Dst level is reached in the midnight to post-midnight time period, then the F-layer rises, quickly falls and irregularities are generated.

Data and analyses are at hand to forecast the morphology and dynamics of global F-layer irregularities. For the intense irregularities at auroral and sub-auroral latitudes, the forcing parameters are the form, level, and development of the magnetic storm. For the equatorial latitudes the ring currents peak excursion has differing effects as a function of local time.

During any year we have thousands of users of the transmissions for the satellites in the 250 MHz range. The effect of amplitude fluctuations can be minimized with diversity methods used during magnetic storms; it is up to the user community to demand attention to this problem.

ACKNOWLEDGEMENTS

The studies of J. Aarons are supported by the Office of Naval Research. The author thanks S. Basu and J.A. Klobuchar for use of observations and data from the Geophysics Laboratory. The data for the 6300 A observations were supplied by M. Mendillo.

REFERENCES

Aarons, J.(1982) Global morphology of ionospheric scintillations, Proc. IEEE 70, 360-378

Aarons, J., and A.S. Rodger (1989) Sub-auroral F-layer irregularities and the ring current, Ann. Geophysicae 7, (2), 169-176.

Basu, Su., S. Basu, E. MacKenzie and H.E. Whitney (1985) Morphology of phase and intensity scintillations in the auroral oval and polar cap, Radio Science 20, 347-356.

Basu, Su., E. MacKenzie, and S. Basu (1988) Ionospheric constraints on VHF/UHF communication links during solar maximum and minimum periods, Radio Science 23, 363-378.

Blanc, M. D. Acayde, and J.D. Kelly (1983) Magnetospheric convection effects at mid-latitudes 2. A coordinated Chatanika/Saint-Santin study of the April 10-14, 1978 magnetic storm J. Geophys Res. 88, 224-234.

Dabas, R.S., Lakshmi, D.R., and Reddy, B.M. (1988) Effect of geomagnetic disturbances on the VHF nighttime scintillation activity at equatorial and low latitudes, Radio Science 24, 563-573.

DasGupta, A., A. Maitra, and S.K. Das, Postmidnight equatorial scintillation activity in relation to geomagnetic disturbances, (1985) J. Atmos. Terr. Phys., 47, 911-916.

Fejer, B.G., E.R. de Paula, I.S. Batista, E. Bonelli, and R.F. Woodman (1989) Equatorial F region vertical plasma drifts during solar maxima, J. Geophys. Res. 94, A9, 12,049-12,054.

Fejer, B.G., R.W. Spiro, R.A. Wolf, and J.C. Foster (1989) Latitudinal variation of perturbation electric fields during magnetically disturbed periods: 1986 SUNDIAL observations and model results. Submitted to Ann. Geophysic.

Foster, J.C., and J. Aarons (1988) Enhanced anti-sunward convection and F-region scintillations at mid-latitudes during storm onset J. Geophys. Res. 93, 11,537-11,542.

Hamilton, D.C. , G. Gloeckler, and F.M. Ipavich, W. Studemann, B. Wilken and G. Kremser (1988) Ring current development during the great geomagnetic storm of February 1986 J. Geophys Res. 93, 14,343-14,355.

Huang, Y.,and S. Chen (1980) Report of Lunping Observatory January-December 1980.

Hudson, M.K. and M.C. Kelley (1976) The temperature gradient drift

instability at the equatorward edge of the ionospheric plasma trough, J. Geophys. Res. 81, 3913.

Kersley, L., Huank, and H. Chandra (1988) Scintillation at high latitudes during winter magnetic storms J. Atmos. Terr. Phys., 50, 565-572.

Kozyra, J.U., E.G. Shelley, R.H. Comfort, L.H. Brace, T.E. Cravens and A.F. Nagy (1987) The role of ring current O+ in the formation of stable auroral red arcs, J. Geophys. Res., 92, 7487

Lui, A.T.Y., R.W. McEntire, and S.M. Krimigis (1987) Evolution of the ring current during two geomagnetic storms J. Geophys. Res. 92, 7459-7470.

Mendillo, M., M.D. Papagiannis, and J.A. Klobuchar (1972) Average behavior of the mid-latitude F-region parameters NT, N max, and t during geomagnetic storms, J. Geophys. Res. 77, 4891.

Rastogi, R.G. and R.F. Woodman (1978) Spread F in equatorial ionograms associated with reversal of horizontal F-region electric field Ann. Geophys. 34,31-36.

Rastogi, R.G. and J. Aarons (1980) Nighttime ionospheric radio scintillations and vertical drifts at the magnetic equator, J. Atmos. Terr. Phys. 42, 583-591.

Rino, C.L. and S.J. Matthews (1980) On the morphology of auroral zone radio wave scintillation, J. Geophys. Res. 85, 4139.

Rodger, A.S. and J. Aarons (1988) Studies of ionospheric F-region irregularities from geomagnetic mid-latitude conjugate regions J. Atmos. and Terr. Phys. 50, 63-72

Rostoker, G., (1972) Geomagnetic Indices, Reviews of Geophysic. 10, 935-950.

Spiro, R.W., R.A. Wolf, and B.G. Fejer (1984) Penetration of high-latitude electric field effects to low latitudes during SUNDIAL, Ann. Geophys., 6, 39-50.

Suguira, M. and S. Chapman (1960) The average morphology of geomagnetic storms with sudden commencement Abh.Akad.Wiss. Goettingen Math. Phys Kl., 3. 10, 935-950.

Weber, E.J., R.T. Tsunoda, J. Buchau, R.E. Sheehan, D.J. Strickland, W. Whiting and J.G. Moore (1985) Coordinated measurements of auroral zone plasma enhancements, J. Geophyus. Res. 90, 6497.

DISCUSSION

C. GOUTELARD, FR

Pouvez-vous commenter l'extension des perturbations aurorales vers les zones tempérées et leur corrélation avec Kp comme cela a été établi dans les zones subaurorales.

Quelle est votre opinion sur les modèles de perturbations équatoriales au voisinage du terminateur (cf. Hanuise).

Pouvez-vous suggérer des modèles plus précis ?

AUTHOR'S REPLY

In my new work I have found that the parameter Dst fits the equatorial data better than Kp. The ring current which lasts longer than the initiating magnetic activity plays a role in the night electric field even after Kp has decreased. I believe we are searching for a model of F layer electron density height and height changes which could be linked to the generation and inhibiting of equatorial patches and plumes.

Optimum space, time and frequency sampling of the ionosphere with advanced digital ionosondes

Adolf K. Paul and David B. Sailors

Ocean and Atmospheric Sciences Division

Naval Ocean Systems Center, San Diego, CA 92152-5000

Introduction

Real time knowledge of the state of the ionosphere is a prerequisite for optimum usage of the ionosphere as an HF propagation medium. All properties of the ionosphere, such as absorption and refraction change with location, time and frequency. Consequently it is desirable to update the description of the ionosphere frequently, at many locations and in small frequency increments while, at the same time, keeping interference of the probing sounder to a minimum. The optimum compromise of course is dictated by structure scales of the medium.

Ionosondes were always an important tool for the observation of the ionosphere. The traditional analog instruments were only capable of measuring the travel time or virtual height as a function of frequency with limited accuracy. Sometimes, with special effort, estimates of the ionospheric absorption at a few frequencies could be obtained.

Advanced digital ionosondes are now capable of recording additional information, such as the radio phase and the amplitude of each echo. If the echoes are recorded at several receiving antennas, the change of phase with antenna location permits a highly accurate estimate of the angle of arrival or the apparent location of the reflection point. Repeating the transmission of a given frequency a short time later yields the Doppler frequency as a function of the radio frequency (and of the height, after profile computation). The change of the phase over a small frequency step gives improved accuracy of the virtual heights.

The recording time for an ionogram including the type of information mentioned above can be as short as 20 seconds. Under computer control the antenna, frequency switching can take place in any random sequence.

At a given ionosonde location and for a given measurement accuracy of the system the precision of the ionospheric information obtainable depends then on the distances between receiving antennas, the time interval between repetitions of sounding at the same frequency, the number of frequencies per ionogram and the time interval between consecutive ionograms. Any operational procedure has to take precaution to avoid undersampling of nonmonotonic phenomena.

Scale sizes of ionospheric variability and structure

Ionospheric variations usually take place in more than one dimension. At sunrise the rapid increase of the intensity of the ionizing radiation in time causes a strong horizontal gradient of the electron density perpendicular to the daylight boundary as the earth rotates around its axis. This is a typical example where the temporal variation is easily observable at a given location and the spatial scale is deduced from the temporal scale and a known velocity, in this case the rotational velocity of the earth. As an example, figure 1 shows the maximum electron density of the F-region derived from the penetration frequency foF2 as a function of time. The electron density is approximately constant for the first five hours of the day, then increases rapidly in a linear fashion for almost four hours during the sunrise period before levelling off for the noon hours. At the latitude of Brighton (40 degrees N) the 4 hour temporal scale size corresponds to a horizontal east-west scale size of 5000 km.

Also visible in figure 1 is a small sinusoidal variation between 1000 and 1300 hours with a period of about 2.5 hours. This is most likely due to an acoustic gravity wave, which may propagate with a speed of 500 to 800 km/h, resulting to a corresponding wavelength of 1200 to 2000 km. In this case the direction of propagation is not known since the ionosonde can only determine the orientation of the wave front, which may or may not coincide with the propagation direction.

Most of the acoustic gravity waves observed in the midlatitude ionosphere have much shorter periods. A more typical behavior is demonstrated in figure 2 which shows strong variations of the height of the F-region maximum, hmF2, with periods of approximately 20 minutes. With similar propagation velocities as mentioned above the wavelengths are now in the range between 170 and 270 km.

A different type of short term variation is visible in figure 3. Beginning at 0300 hours we see a fast increase of the height of the F-layer maximum by approximately 100 km over a period of one hour and then a rapid decrease by 200 km over a 2 hour period. Between 0300 and 0500 hours (before sunrise), when the height of maximum goes through half a cyle of its variation, there is very little change of the maximum electron density. The continuation of the decrease, however, is then coupled with the sunrise increase of the electron density in the F-region. This up and down movement of the F-layer with a vertical scale size of approximately 100 km appears frequently, but not always, immediately before sunrise. Based on the time of the appearence of this effect one could translate its duration, using again the rotational velocity of the earth, into a horizontal east-west spatial scale and a corresponding gradient. Angle of arrival measurements taken at 90% of foF2, however show that this assumption is wrong. Zenith angles of 8 degrees and more are observed with apparent reflection points south of the observation site between 0300 and 0430 hours. This indicates the presence of a strong north-south gradient and a rather complex structure for the phenomenon.

The upper part of figure 4 gives an example of the variation of the apparent reflection point observed at a fixed frequency of 9.1108 MHz in the extraordinary mode for a time interval of 19.6 seconds. The trace begins at the coordinates (4.0,-10.0), moves then toward south east,returns almost on the same track,changes direction toward south-west, completes a loop, moves toward west and finally toward south. The initial stage, i.e. the back and forth movement of the apparent reflection point, represents an oscillation with a period of about 7 seconds. Contrary to the previous examples given above, there are reasons to believe that the variation of the apparent reflection position is not a change of the orientation of the reflection area, but the result of variations of small scale irregularities affecting the phase paths of the echoes to the four receivers in slightly different ways. This

explanation is supported by the appearence of weak spread-F on the ionogram shown in the bottom part of figure 4, which was taken immediately after the fixed frequency recording. If this interpretation is correct, the motion of the apparent reflection point will depend (among other factiors) on the horizontal scale of the irregularities relative to the distances between the receiving antennas. Varying those distances in future experiments could provide a test on this hypothesis.

Contrary to horizontal scale sizes which are derived with some known or assumed velocities from the temporal change of an ionospheric quantity, vertical scale sizes can be observed directly with a single ionosonde. The best known parameter is the half-thickness of the electron density profile of a layer which is related to the scale height of the neutral atmospheric density. An example for the daytime behavior of the F-layer half-thickness is given in figure 5. As expected, the quantity shows variations within limits of 10 to 15 % which are clearly due to the ever present acoustic gravity waves. The magnitude of such variations mainly reflects the deformation of the layer in the vicinity of its maximum and should not be interpreted as a measure for contraction and expansion, since in most cases much smaller relative changes of the maximum electron density are observed.

A different vertical scale size can be defined based on the height dependance often found in Doppler data as in figure 6. The Doppler frequency as a function of the radio frequency near the maximum of the F-layer for the extraordinary mode was derived from the change of the radio phase with time for three consecutive ionograms recorded in 3 minute intervals. At 1506 hours the Doppler frequency is negative and nearly constant for the frequencies shown except for the very highest frequencies near penetration. In the next ionogram the Doppler frequency varies rapidly from approximately -1.0 Hz to +0.6 Hz. Again 3 minutes later the Doppler frequency reaches a positive value of about 1.1 Hz and is practically constant over the entire range shown. The frequency range for which data are presented in figure 6 corresponds to a height range of about 32 km according to the profile parameters computed. This means that in the F-region the Doppler frequency can change from -1.0 Hz to +1.1 Hz over a height range of about 32 km at a given time or, during a time interval of 6 minutes, at a given height.

Effects of structure and variability on HF operations

The ionosphere provides a natural mechanism for radio waves to propagate over the entire earth. There is virtually no lower frequency limit and extreme low frequencies (ELF, 30 Hz - 30 kHz) are actually used in special communication systems. Most of the long distance communication, however, takes place at higher frequencies in the HF/VHF range.Its upper frequency limit is determined by the maximum electron density in the reflecting area and the ray direction relative to the direction of the electron density gradient (angle of incidence). The maximum electron density at a given location is a function of the time of day, season and solar activity. There also is a strong dependence on the latitude modified by the local direction of the earth's magnetic field. The angle of incidence is mainly a function of the distance between the transmitter and receiver, the number of hops required to propagate over this distance and the reflection height. The bandwidth for ground to ground communication provided by the ionosphere is therefore variable. Its upper frequency limit can be as high as 50 MHz in the sunlit part of the earth, but often is much lower, e.g. at high latitudes and in the nightside of the earth. On the other hand the enhanced ionization present at lower heights of the ionosphere in the day side causes a reduction of the signal strenght due to absorption for the lower frequencies in this band. This means that no continuous operation in the HF/VHF band intended for long distance propagation can use one frequency all the time. For example, shortwave broadcast station use a wide spread set of frequencies depending on the propagation conditions.

Medium and small scale structures and short term temporal variations, superimposed on the larger scales generated by the differences between day and night and the latitude dependence of the electron density, affect the propagation conditions in a variety of ways. If the propagation distance is of the order of the horzontal scale or longer, multipath propagation is possible. This means that two or more ray paths exist between transmitter and receiver. Differences of travel times and phase delays between the different paths then lead to signal distortion. In this situation, but also if the distance is short compared to the horizontal scale,the path usually becomes asymmetric and deviates from propagation in a vertical plane if the reflection area is tilted. This can lead to large errors in HF-direction finding. An asymmetric path or, equivalently, a tilt of the reflecting area, also means a reduction of the maximum usable frequency over the given distance and a given reflection height relative to symmetric propagation, since the asymmetry reduces the effective angle of incidence. When the F-layer is strongly distorted, i.e. when the orientation of electron density gradient changes significantly over the height interval passed by a ray, the propagation distance (in broadcasting) and the maximum usable frequency (point to point communication) can change largely either way compared with symmetric conditions.

Doppler shifts caused by temporal changes of the ionosphere can interfer with over the horizon radar observations, especially if they change rapidly with height (distance), as in figure 5. Generally, under highly variable conditions as in figure 5, signals of any wide band HF-systems will be strongly distorted. Since very often the temporal changes are quasi periodic, errors in HF-direction finding can be reduced by averaging over time, if the duration of the period in question can be estimated. Similarly, observations of quasi periodic variations at one location could predict tilts (for error correction) at some other location, if the direction of the wave propagation can be determined.

Sampling of ionospheric parameters

Figure 7 shows an amplitude spectrum of the height variations given in figure 2, computed with AFA (Anharmonic Frequency Analysis, Paul, 1972). The data were collected over a 7 hour period in 3 minute intervals. Besides the high amplitude at the lowest end of the frequency scale, reflecting the diurnal variation of the height, there is a definite peak at 0.05 cycles/minute corresponding to a period of 20 minutes. The periods at the half amplitude points of this peak are 22.2 and 16.7 minutes respectively. The amplitude of this narrow band oscillation is approximately 7 km and there is no doubt that such a large variation has a significant effect on all propagation aspects of the kind mentioned above. Averaging angle of arrival measurements for the purpose of direction finding over a time period of 20 minutes will definitely improve the accuracy under such conditions. Small secondary peaks at 0.02, 0.03 and also at 0.075 cycles per minute may or may not be significant. The amplitude of the spectrum decreases and stays low for frequencies above 0.1

cycles per minute.

This example, which is typical for all high temporal resolution sequences available, demands a minimum sampling rate of at least 12 ionograms per hour,corresponding to 5 minute time intervals between recordings, assuming that no significant spectral contribution are to be found at periods shorter than 10 minutes. Slower sampling rates cause aliasing and misinterpretation of the periodicities and other properties of acoustic gravity waves.

Sampling in the frequency domain is always coupled with sampling in the height domain. For a given frequency interval the resulting height interval can be small e.g. in the lower portion of the F-layer, or relatively large near the maximum of the layer. If the ionosphere were to be sampled in equal height steps, the reflection height as a function of frequency would have to be known first. Besides during times of very rapid changes, this could be accomplished approximately assuming continuity of the true height profile in time and extrapolating from the previous ionogram. Since modern ionosondes are computer controlled and have a very high frequency resolution, such procedures could be programed into the system, where minimum interference and/or maximum temporal resolution are essential. The resulting frequency pattern would be nonuniform and would change continuously with time. If the operation of an ionosonde would be optimized in such a way a height resolution of 10 km would require only 20 -30 well chosen frequencies and an ionogram containing the basic information could be recorded in a few seconds.

On the other hand a signal with a given radio frequency may be reflected from several different heights. This is definitely true for the ordinary and extraordinary components. Several reflections from different locations in the ionosphere can also be observed when the electron density has a relatively short scale horizontal structure. This situation is usually caused by very strong gravity waves called travelling ionospheric disturbances recognizable by complex multiple echo traces. True height computations are not very reliable under such circumstances, since large deviations from vertical propagation occur. For detailed studies of these phenomena sampling rates higher than 12 ionograms per hour are needed.

In order to obtain reliable estimates of horizontal scale sizes, the installation of additional receiving sites is necessary. Distances should be of the order of a fraction of a typical horizontal wavelength. A triangle with an ionosonde and two additional receivers with a base of about 50 -100 km should be short enough to avoid spatial undersampling and at the same time long enough to give sufficient accuracy for the determination of magnitude and direction of the propagation velocity of acoustic gravity waves.

References:

Paul, A. K., Anharmonic Frequency Analysis, Math. Comp. 26, 118, 437, 1972

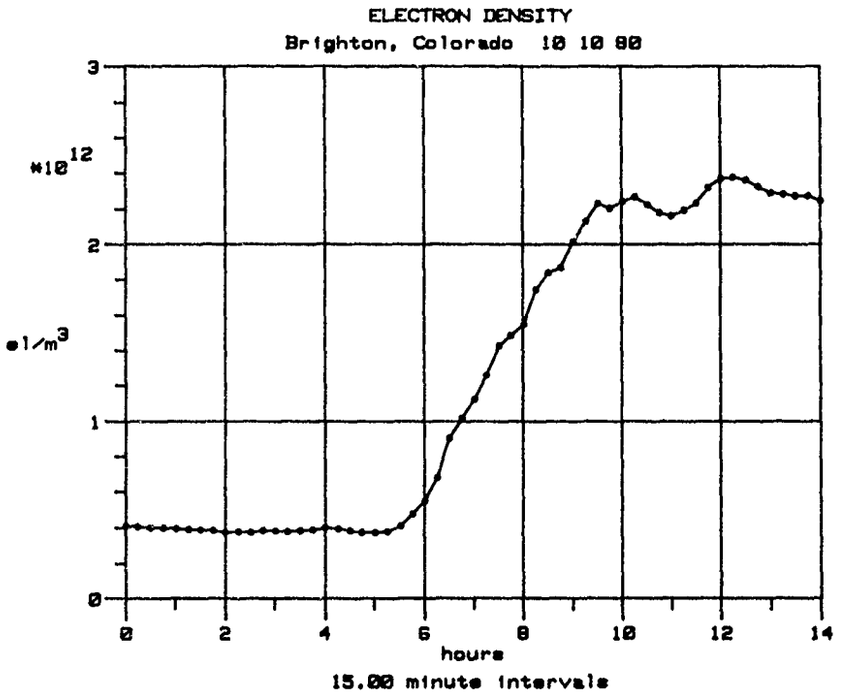


Figure 1. Night-day change of the maximum electron density of the F-region

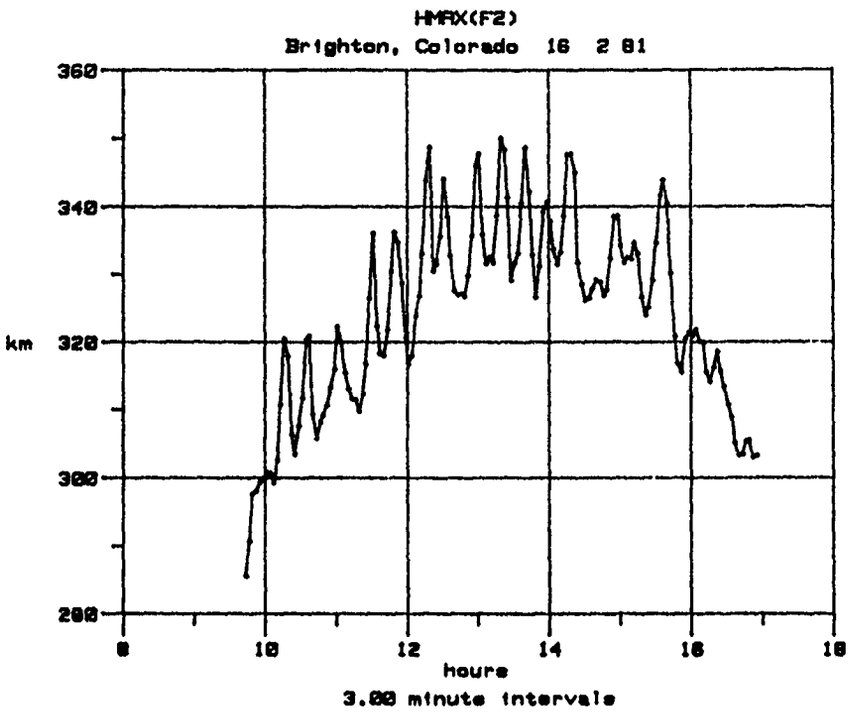


Figure 2. Oscillations of the F-region height of maximum

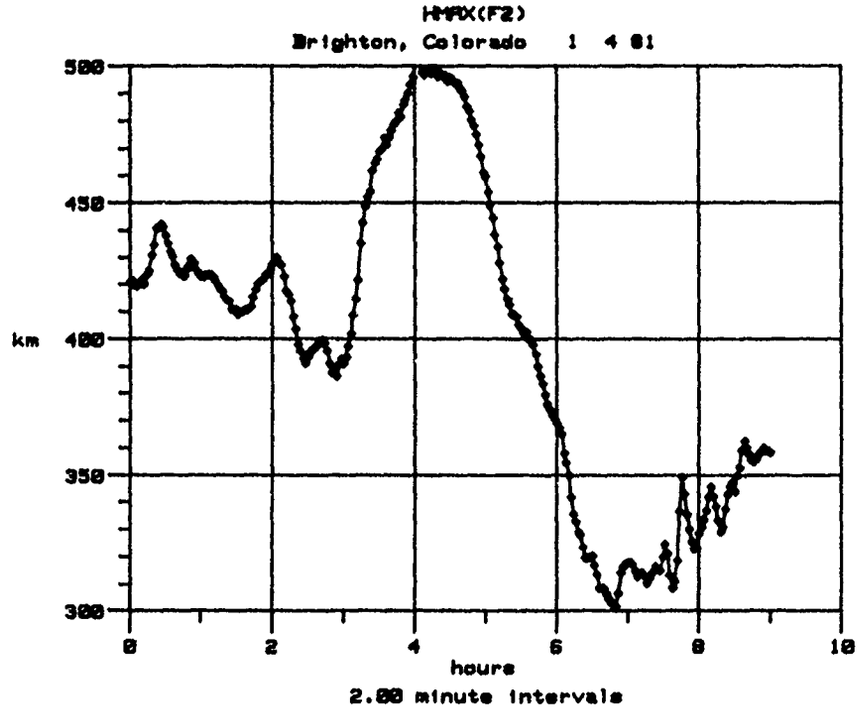


Figure 3. F-region height of maximum variation during the morning hours

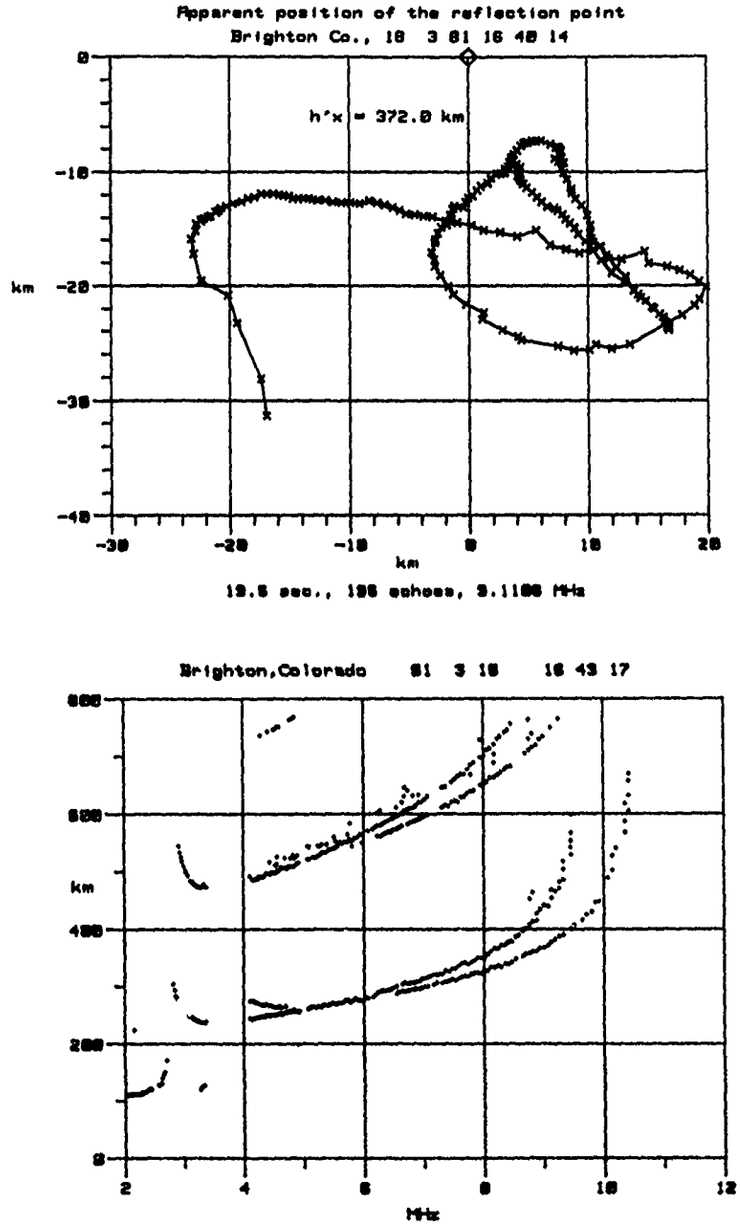


Figure 4. Change of the apparent position of the reflection point (top) probably caused by weak irregularities as indicated by light spread-F in the ionogram (bottom)

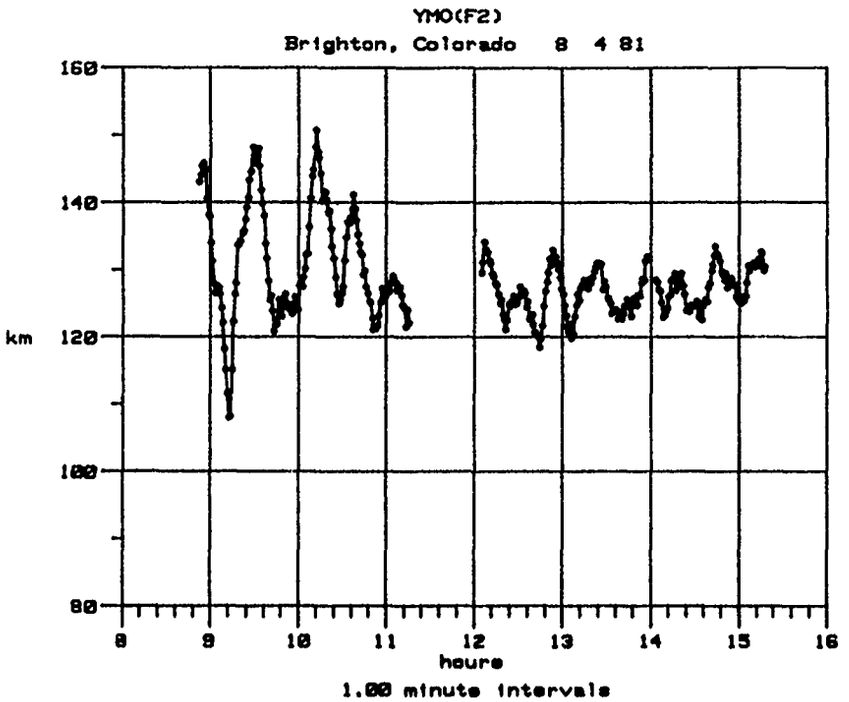


Figure 5. Oscillations of the F-region thickness parameter mainly due to changing deformations of the layer by acoustic gravity waves

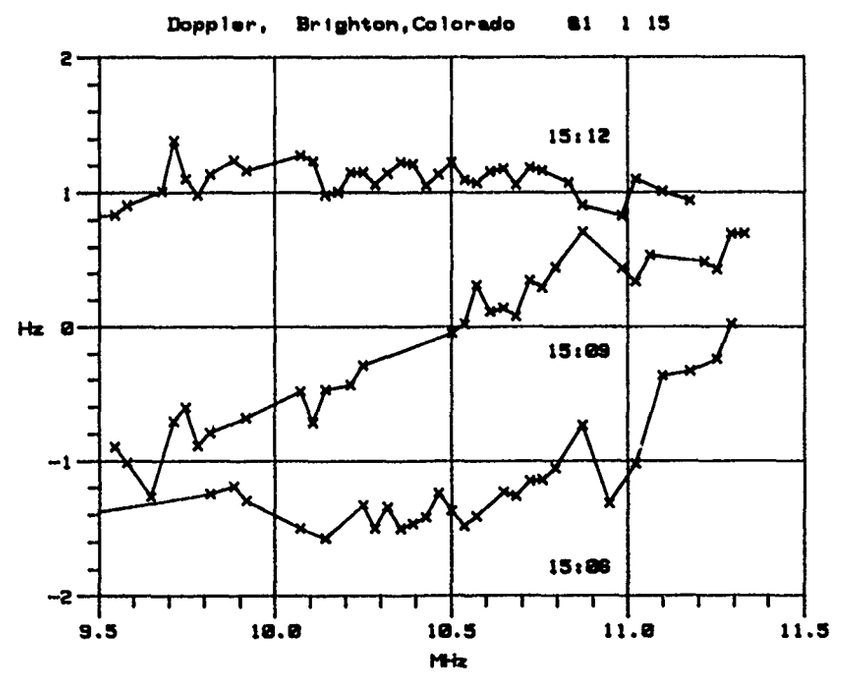


Figure 6. Doppler frequency as a function of radio frequency (height) and its change with time

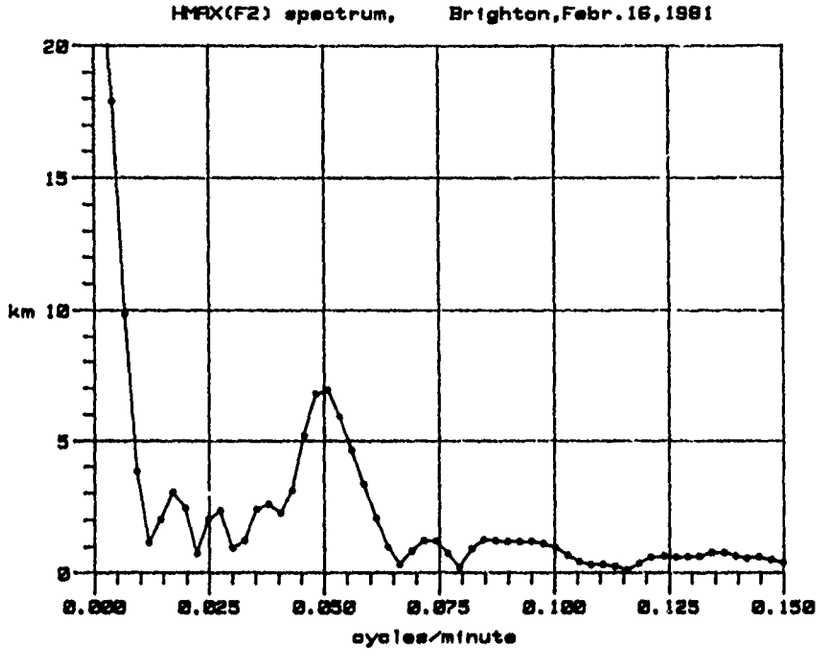


Figure 7. Amplitude spectrum of the F-region height variations of figure 2

DISCUSSION

R. JENKINS, CA

In the first part of figure 4, the apparent reflection point position is plotted as a function of time. With regard to the array used to take these measurements, what was the array aperture, and, if there were more than 2 antennas in each direction, were the phase changes measured across the array consistent with plane waves ?

AUTHOR'S REPLY

The phase values measured at the four receiving antennas are tested for consistency with the plane wave assumption. If the test fails, the data are usually rejected for angle of arrival computations. Exceptions are cases of overlapping ordinary and extraordinary components, if their separation is desirable (Paul, 1983).

C. GOUTELARD, FR

Si vous utilisez des antennes distantes de 100 m, vous avez une bonne mesure aux fréquences basses (1/2MHz) mais aux fréquences élevées, vous avez une ambiguité dans la mesure des angles d'arrivée à cause du fait que votre spectre devient lacunaire. Comment levez-vous l'ambiguité ?

AUTHOR'S REPLY

If echoes are reflected from strongly tilted areas, e.g. sporadic-E, the angle of arrival computation may be ambiguous, if the radio wavelength is much shorter than the distance between receiving antennas. In such cases, a unique determination of the angle of arrival is still possible if continuity of the phase observations with frequency can be assumed and observations start at low enough frequencies (Paul, 1990).

Paul, Adolf K., Numerical separation of overlapping ordinary and extraordinary echoes in digital ionograms, Radio Sci. 18, 416-420, 1983.

Paul, Adolf K., On the variability of sporadic-E, Radio Sci. 25, 49-60, 1990

THE UTILITY OF HIGH FREQUENCY GROUND WAVE IN A DISTRIBUTED COMMUNICATION SYSTEM

by
James R. Champion
The Johns Hopkins University Applied Physics Laboratory
Laurel, Maryland 20723
USA

SUMMARY

The Johns Hopkins University Applied Physics Laboratory (JHU/APL) has investigated high frequency ground wave (HFGW) communication in mountainous terrain. Using a broadband, discone antenna and 100 watts of power, communication links in the 20 to 30 MHz band of at least 50 and, in some cases, as much as 115 km have been established over mountainous paths. Links were non-line-of-sight, and immune to ionospheric fading. Digital and wide-band frequency hopping communications have been demonstrated using HFGW. The discone antenna is about 3.5 m high when assembled and can be made extremely portable. The special characteristics of the 20 to 30 MHz band coupled with the properties of the discone antenna would offer advantages to distributed communication systems in mountainous terrain. This paper presents HFGW results and discusses how this medium may be exploited in a distributed system which uses both frequency and spatial distribution.

PREFACE

Under the sponsorship of the Defense Nuclear Agency, The Johns Hopkins University Applied Physics Laboratory (JHU/APL) has been investigating high frequency ground wave (HFGW) as a nuclear survivable means of communication for land-mobile missile systems in Europe. Using a frequency band of 20 to 30 MHz, and a portable, broadband, discone antenna, HFGW propagation characteristics have been demonstrated which could be exploited in distributed communication systems (Reference 1). Test results show that by means of portable, wide band discone antennas the upper part of the HF band could be used alone or as an extension of the very high frequency (VHF) band to achieve non-line-of-sight, non-fading, spread spectrum communication. As discussed below, both spatial and frequency distribution strategies could be used.

DISCUSSION

1. ADVANTAGES OF THE 20 TO 30 MHz BAND

In order to maximize the range, it is desirable to use a resonant antenna for both transmission and reception. On the other hand, optimum portability is wanted. With these constraints in mind, 20 MHz was chosen as the lower limit of the frequency band to be tested. A resonant antenna could then be constructed which had a height of less than 4 meters. The upper frequency was taken to be 30 MHz because this is the limit of most military HF transceivers and because it was anticipated that the ability of the signal to provide non-line-of-sight coverage by diffraction in mountainous terrain would significantly diminish above 30 MHz.

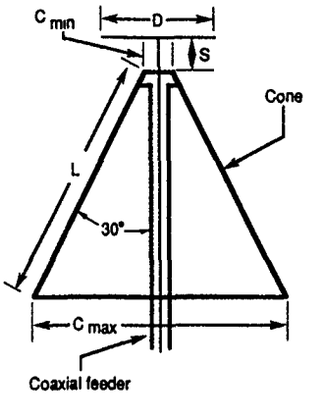
The 20 to 30 MHz band offers additional advantages. As will be shown, ranges of interest are achievable in mountainous terrain with reasonable transmission power levels. Beyond the ground wave range, however, there is a zone extending for hundreds of miles where no signal, ground wave or sky wave, is present. Accordingly, there is a large region surrounding the area of communication coverage where it is not possible to intercept or DF. Beginning at about 500 miles, it is possible to intercept a skipped signal a few percent of the time. Although the skip probability increases somewhat beyond this point, the large skip angle challenges DF accuracy.

A further advantage of the 20 to 30 MHz band is the relatively low noise level. Within the HFGW range, competitive noise sources are not likely to be transmitted from a resonant antenna, and consequently, their effect is limited. Interfering distant stations are limited by the skip characteristics discussed above. At night this band is particularly quiet as local noise sources are quiet and the probability of distant signals skipping in is almost nil. A study by Laycock et al of percent congestion values vs frequency range illustrates the point (Reference 2). Based on data taken at a rural site in central England, their data show zero percent congestion for most frequency intervals in the 20 to 30 MHz band. By contrast, the lower region of the HF band is almost 100 percent occupied, particularly at night.

2. PROPERTIES OF THE DISCONE ANTENNA

Military users of HFGW require the capability of changing frequencies quickly to avoid jamming and reduce the probability of interception. Further, portability and speed of erection and disassembly are highly important. The discone antenna was found to have the desired characteristics. The discone antenna is comprised of a disc and a cone. It is a wide band antenna with a nominal 50 ohm impedance which produces vertically polarized radiation (References 3 and 4). For HF and very high frequency (VHF) applications the disc and cone are usually not solid, but are comprised of spokes. The height of the antenna is set by the length of the cone spokes, which are chosen to be equal to one quarter the wavelength of the lowest desired frequency (cut-off frequency) and the cone angle, which is typically 30 degrees. (For best performance the cone spokes are chosen so that the cut-off frequency is taken to be about 1 MHz lower than the lowest frequency to be used.) Figure 1. shows the dimensions of the discone antenna used in the work described here. Eight spokes were used on the disc and cone for the first discone built at APL. Subsequently a discone was tested which only used six spokes on the disc and six for the cone. No difference in performance was measured between the two configurations.

Beginning just above the cut-off frequency the discone antenna has a good voltage standing wave ratio (VSWR) for many decades. Figure 2 shows VSWR vs frequency measured for a discone antenna having a cut-off frequency of 19 MHz. As a consequence, a 50 ohm HF transceiver will transmit and receive efficiently on this antenna throughout the 20 to 30 MHz band without the need for an antenna matching unit. Further, the discone lends itself to using spread spectrum or frequency hopping techniques throughout this 10 MHz band.



$L = (\lambda/4) \times 0.96 = C_{max}$

$D = 0.7\ C_{max}$

$C_{min} = < L/22$

$S = 0.3\ C_{min}$

Relations for selecting discone parameters

Cut-off freq. MHz	C max (m)	C min (m)	D (m)	S (m)	L (m)
19	3.78	.172	2.65	.0516	3.78

Dimensions for discone antenna with cut-off frequency of 19 MHz

Fig. 1 Basic structure of a discone antenna.

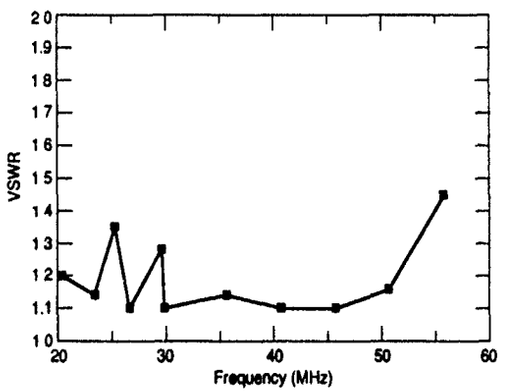


Fig. 2 VSWR vs frequency for a discone antenna with a cut-off frequency of 19 MHz.

The discone antenna can be made extremely portable. Erected, the discone stands about 4 meters high. Experience has shown that this size is reasonable for use in wooded, tactical positions. The antenna can be fabricated so that when disassembled it makes up a package weighing less than three kilograms which can be carried by backpack. Assembly and disassembly each take about 10 minutes.

3. MEASUREMENTS OF HFGW PROPAGATION

Using discone antennas and frequencies in the 20 to 30 MHz band, HF ground wave propagation has been studied in a number of different environments ranging from the mountains of West Germany to the Mojave desert in the United States. The results are described in the following:

Range Measurements Made in the Mountains of West Germany

Using a discone as the base station antenna, single sideband (SSB) voice and continuous wave (CW) test signals were transmitted at 22.733 and 26.725 MHz. The output power was selected to be 100 watts. Both military and civilian radio equipment was used to transmit, with identical results. The military transceiver was the AN/GRC 193A. The civilian transceiver was a Yaesu model 757GX. A second discone was moved to positions of increasing distance from the base station. At each test site voice contact was established and a CW signal transmitted. Measurements were made of signal-plus-noise and noise using another Yaesu 757GX transceiver which had been calibrated by means of injecting known power levels and measuring the audio power out. Table 1. shows the signal-to-noise ratio (SNR) and the received power levels measured over four mountainous paths. Note that only path 1 represents a maximum range. Even though signal levels indicated that greater range was possible along the other paths, the test team could not proceed farther for administrative reasons.

All paths listed in Table 1 were mountainous and non-line-of-sight. To illustrate, Figures 3 and 4 show the terrain profiles for paths 1 and 4, respectively. Elevation in meters is plotted against range in kilometers. The transmitter was located at the zero kilometer point and the signal measurement was made at the end of the path. It should be noted that, in addition to the measurements made at the end of the path, measurements were made at intermediate points, including positions in valleys, and the SNR for some of these measurements are shown in the figures.

HFGW Propagation across South Sandia Peak

A test was performed to determine whether HFGW communication could be conducted across South Sandia Peak near Albuquerque, New Mexico. The results were of interest not only because of the severe terrain but also because, in contrast to West Germany, the ground was extremely dry since the rainfall in the Albuquerque area is less than 10 inches per year. Theoretically,

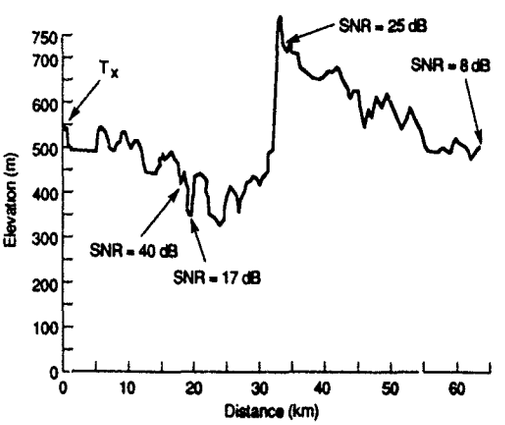


Fig. 3 Terrain profile for path 1.

Table 1. Summary of HFGW Range Survey (All measurements made at 26.725 MHz using a transmission power of 100 watts.)

Path	Tx Elevation (m)	Rx Elevation (m)	Range (km)	SNR (dB)	Received Signal Power (dBm)
1	572	498	64	8	-114
2	572	97	89	16	-102
3	450	525	52	28	-92
4	561	350	115	16	-102

Note: The ranges for path 1 represents a maximum value. The signal received on other paths indicated greater ranges could be achieved, but administrative reasons prevented further investigation.

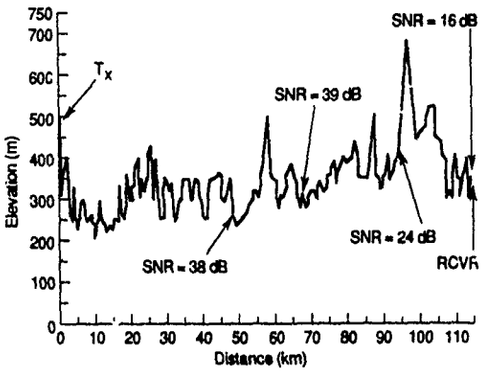


Fig. 4 Terrain profile for path 4.

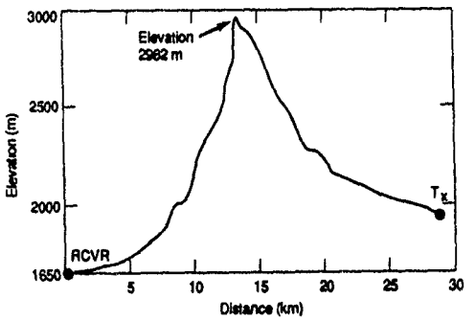


Fig. 5 Terrain profile across south Sandia peak.

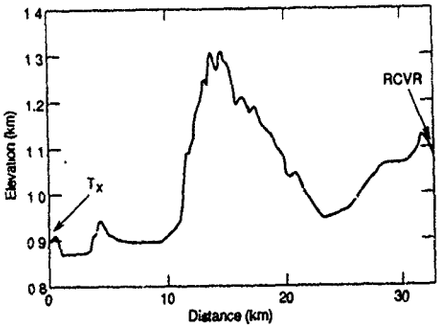


Fig. 6 Terrain profile for HFGW link in Mojave Desert.

ground wave propagation is sensitive to the moisture content of the ground. As can be seen in the terrain profile of Figure 5, the peak rises over a kilometer from the base station, which was located on Kirtland Air Force Base. From that site the top of South Sandia Peak subtends an angle of almost six degrees. Using two discone antennas, 100 watts output power and frequencies in the 20 to 30 MHz band, a 29 km SSB voice link was established at will over a period of several days. Signal level measurements were made over a period of several hours spanning sunset and are presented in the Section 4, which discusses signal stability.

HFGW Propagation in the Mojave Desert

In a further test of HFGW performance in a desert environment, a propagation path was investigated in the Mojave desert, near Barstow, California. The terrain profile for the 33 km test path is shown in Figure 6. This region gets less than three inches of rain yearly. Using two discone antennas and 10 watts of output power from a AN/PRC 104, good SSB voice communication was established on all but one of five channels distributed throughout the 20 to 30 MHz band. SNR levels were measured using CW signals and the results are tabulated in Table 2. As can be seen, the 26.720 MHz frequency was quite noisy and this was caused by an interfering signal which was present due to a transient skip condition.

4. SIGNAL STABILITY OF HFGW

While the sky wave signal level varies widely with time as the ionosphere changes, the ground wave signal level is expected to be more stable, thereby increasing its utility for communication systems. Accordingly, if the received power level remains steady during the time the ionosphere changes, this supports the claim that the signal is truly ground wave.

Three twenty-four hour tests were performed in the mountains of Germany during which signal levels were measured over a period of twenty-four hours. The dates and paths are listed in Table 3. CW signals were transmitted at 100 watts using a discone antenna. Another discone at the end of the path was used to receive the signals and measurements were made using a calibrated Yaesu 757GX transceiver. Signals were transmitted and received levels measured on a regular basis. Figure 7 shows the results for 26.725 MHz taken along path 3 (Test 3) in February 1986. Measurements were made every fifteen minutes. The signal levels stayed constant within plus or minus 2 dBm, which was near the resolution of the method used to make the measurement. These results are representative of the data obtained during all three tests. No evidence of sky wave fading was seen.

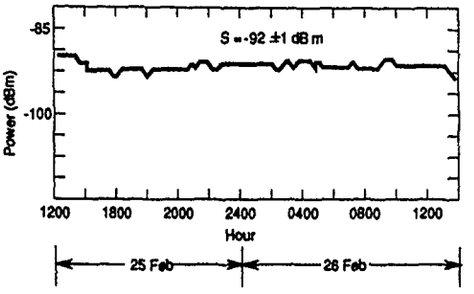


Fig. 7 Results of signal stability test for a 26.725 MHz groundwave signal.

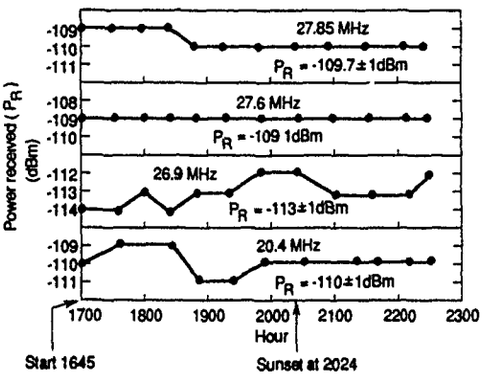


Fig. 8 Power received vs time of day for 29 km HFGW link across Sandia peak.

Table 2. HFGW Signal Levels Measured over 33 km path in the Mojave Desert (Discone to Discone Antenna Configuration)

Frequency (MHz)	Signal-Plus-Noise (dBm)	Noise (dBm)	SNR (db)
20.400	-115	-137	22
23.450	-110	-138	28
25.350	-111	-134	23
26.720	-113	-119	6
29.725	-116	-134	18

Table 3. Results of HFGW Signal Stability Measurements Performed during 24-Hour Periods over Non-line of Sight Paths

Test	Range (km)	Frequency (MHz)	Received Signal Power (dBm)
1	49	22.733	-105±2
		26.725	-102±2
2	49	22.733	-98±2
		26.725	-96±2
3	52	27.733	-98±2
		26.725	-92±2

In July 1989, during the test of propagation over South Sandia Peak, it was desired to demonstrate that the signal was indeed ground wave with no significant sky wave component. As a result, a test was performed in which the signal levels of four frequencies were measured as a function of time through the period of sunset. The results are shown in Figure 8. As can be seen, signal levels were constant within plus or minus 1 dBm and there is no evidence that any change occurred as a consequence of sunset, which took place at 2024 hours.

5. DEMONSTRATION OF DIGITAL CAPABILITY

Several tests have been conducted which demonstrated digital communication over HFGW links. One test used terminal node controllers tied to HF transceivers and personal computers to transmit frequency-shift-keyed packets over mountainous terrain in West Germany. The signal was comprised of two tones, one at 2.069 kHz and another at 1.460 kHz (600 Hz shift). The carrier frequency was 26.850 MHz transmitted at 100 watts. The test packet message was 1800 characters in length, divided into 80 character packets.

After each packet a handshake in the form of a checksum was required from the receiving station before the message continued. If an incorrect checksum was received, the packet was automatically retransmitted. As a result, only complete, perfect messages could be passed. The data rate was 300 bites per second, i.e. 300 baud. Signal-plus-noise to noise ratios (SNR's) were also measured over every path. To accomplish this a CW signal was transmitted using 100 watts at the test frequency. Further, SSB voice contact was also attempted for each path. The ranges of each path, the SNR's, and whether or not packet and voice messages were successfully passed are indicated in Table 4. Communication over path 3 was unsuccessful because of a high noise background. Only partial success was achieved over paths 8 and 11 for the same reason. If alternative frequencies had been available it is possible that communication would have been achieved over all or most of these three paths. The tests over paths 13 and 14 were affected by a thunderstorm. Because of the length of the message, the length of the packets and the frequency of lightning strikes, it was impossible to pass a perfect message in a reasonable amount of time. This might have been avoided by dividing the message into shorter packets or eliminating the handshake requirement. Virtually all paths were non-line-of-sight. Because the test schedule was coordinated with a military exercise, some paths were investigated during the day and others at night, as the opportunity presented itself.

Testing has also been conducted using mobile military radio teletype (RTT) equipment which uses a data rate of 50 baud and an output power of 200 watts. The RTT equipment produced a coded signal consisting of a 2 kHz audio tone which is shifted plus-or-minus 425 Hz. The two resulting audio tones are processed and transmitted the same as SSB signals. Accordingly, the outgoing RF signal is comprised of two signals, one at 2.425 kHz above the carrier frequency and one 1.575 kHz below. Using discone antennas and frequencies throughout the 20 to 30 MHz band, RTT units provided communications between battalion headquarters and six batteries both day and night over a four day period. The exercise took place in mountainous terrain with propagation distances ranging from

Table 4. Summary of HFGW Packet Radio Results

Path	Range (km)	SNR (db)	Packet Successful?	Voice Communications?
1	60	20	Yes	Yes
2	51	24	Yes	Yes
3	98	ND	No	No
4	40	13	Yes	Yes
5	67	17	Yes	Yes
6	22	3	Yes	Yes
7	89	8	Yes	Yes
8	69	ND	No	Marginal
9	48	ND	Yes	Yes
10	100	24	Yes	Yes
11	117	8	No	Yes
12	62	ND	Yes	Yes
13	69	11	Partial	Yes
14	94	8	Partial	Yes
15	38	16	Yes	Yes
16	59	8	Yes	Yes
17	59	17	Yes	Yes

ND = No SNR data available

as little as 14 to as much as 50 km. Messages were passed at will without difficulty and no problems attributed to HFGW were experienced.

6. WIDE BAND FREQUENCY HOPPING

Both the wide band nature of the discone and the propagation characteristics of ground wave in the 20 to 30 MHz band suggest that frequency hopping or other frequency distribution techniques could be exploited to achieve the lowest probability of interception and to minimize the possibility of jamming.

Figure 2, showing the VSWR of the discone versus frequency, demonstrates the ability of the discone to transmit and receive efficiently throughout the 20 to 30 MHz band. Accordingly, it should be possible, for example, to frequency hop over many MHz without sacrificing output signal strength and without requiring a matching circuit to protect the transmitter from reflected power.

Data obtained during range testing in a variety of environments have shown that the ground wave signal level does not change significantly throughout the 20 to 30 MHz band over a mountainous propagation path. Presumably this can be attributed to the properties of diffraction. The pertinent characteristic is illustrated in Figure 8. Here we see that, at the remote station across Sandia Peak, the signal level received near 20 MHz was equal to that received near 28 MHz, within 1 dBm.

HFGW wave wide band frequency hopping was tested in the United States over a 16 km propagation path through a densely populated section of New Jersey. Two discones were used, along with two frequency hopping transceivers provided by the U.S. Army Communication Electronics Command (CECOM). The output power was 20 watts. The frequency hopping transceivers were programmed with six sets of frequencies. These are shown in Table 5. Both the normal hopping rate and the faster, enhanced rate were tested for SSB voice quality.

Good voice communication was established on all hopping sets using the discone antennas. The transceivers had an option by means of which one can choose to have a matching circuit

Table 5. Hop Sets for Discone to Discone Frequency Hopping

Set	Frequency Range (MHz)	Number of Frequencies
1	20.0 - 23.9999	>100
2	24.0 - 29.9999	>100
3	20.0 - 23.9999	2
4	24.0 - 29.9999	2
5	20.5 - 23.5000	4
6	24.0 - 29.9999	6

between the transmitting circuit and an external antenna or, alternatively, connect the transmitter directly to the antenna. This latter function is to be used with a 50 ohm antenna. No difference was detected in the quality of the wide band frequency hopped voice signals, when the matching circuit was in series with the discone and when it was removed.

By contrast, a whip antenna used for the test was not able to provide an intelligible voice signal in the hopping mode for any of the sets listed in Table 5. Voice contact established using the whip antenna on a single frequency would immediately be lost when frequency hopping was begun, even though the matching circuit was used. This demonstrates the need for a wide band antenna when hopping sets which cover a band of several MHz are used for HFGW communications.

CONCLUSION

Empirical evidence shows that HFGW in the 20 to 30 MHz band could be used to provide nuclear survivable, non-line-of-sight communication links having ranges of at least 50 km and as much as 115 km in mountainous terrain. This could be accomplished with portable, broadband discone antennas and using transmission power levels which are already available. It is possible that signal processing techniques could extend the range. If spatially distributed systems are employed to relay messages, communication coverage could be provided over large regions with few, if any, inaccessible points, even in mountainous terrain. Wide band frequency hopping could be exploited to minimize the probability of interception and provide antijamming capability without reducing the communication coverage.

References

1. Champion, J. R., Spangler, R. W., Colson, J. D., and Brown, J. C., "The High Frequency Ground Wave Performance of a Discone Antenna," Proc. 5th International Conference on Antennas, JINA 88, 538 to 539.

2. Kando'an, A. G., "Three New Antenna Types and Their Applications," Waves and Electrons, Vol. 1, No.2, 70W-75W, February 1946.

3. Nail, J. J. "Designing Discone Antennas," Electronics, Vol. 26, 167 to 169, August 1953.

4. Laycock, P. J., Morrell, M., Gott, G. F., and Ray, A. R., "A Model for HF Spectral Occupancy," Proc. 4th Int.Conf. on HF Radio Systems and Techniques, IEE Conf. Pub. 284, April 1988, 165 to 171.

DISCUSSION

Paul S. CANNON

By moving from low frequency (< 6MHz) groundwave to high frequency (20MHz,f < 30MHz) groundwave you mitigate convertional skywave paths via ionospheric layers. By using high frequencies, with relatively little co-channel interference, there is a possibility that propagation via meteor trails can occur giving connectivity over a few hundred kilometers. Did you make any long distance measurements and did you receive any meteor burst signal ?

AUTHOR'S REPLY

We did perform measurements to determine the interceptibility of 20 to 30MHz signals at points 500 and 1200 miles distant from the transmitting discone antenna. However, the data do not allow one to distinguish between signals received after skywave hop and signals which might have propagated by reflection from meteor trails.

RESULTS OF MEASUREMENTS PERFORMED ON HF BACKSCATTER FOR EVALUATING THE INFLUENCE ON SHORT RANGE HF DATA LINKS

by

P.J. van Vliet and P.A. van der Vis
TNO Physics and Electronics Laboratory
P.O. Box 96864
2509 JG The Hague
The Netherlands

SUMMARY

On short range HF links, where the receiving station is relatively close to the transmitting station, reception of long delayed echoes (up to 20 ms) may occur due to backscatter. These backscatter signals will interfere with signals received through the 'legitimate' path, which is either groundwave or nearly vertically incident skywave. Even on low data rate links this time dispersion of signals may cause intersymbol interference. In order to assess the influence of HF backscatter on short range HF data links, measurements were performed. Path delays were measured using a direct-sequence spread-spectrum technique. A carrier, BPSK modulated with a pseudo-noise (PN) sequence, was transmitted and the received signal was correlated with the same PN sequence. Between June 1989 and June 1990, measurements were taken over a 66 km path in The Netherlands and, simultaneously, over a 958 km path between The Netherlands and Norway. Different frequencies in the HF band were used. The measurement results show that especially on the short range link, long delayed echoes appear. The backscatter effect depends on propagation conditions, thus on frequency, time of day, season and solar activity.

1. INTRODUCTION

A major problem experienced in high frequency (HF) data communications is that ionospheric propagation at HF suffers from dispersion in time. Components of a transmitted signal arrive at a receiving station along different propagation paths. Mixing of these components leads to multipath interference. The presence of multipath causes fading, frequency dependent channel characteristics and intersymbol interference.

After having observed many unsuccessful data transmissions during modem trials on short range links (< 100 km) in 1981, some limited measurements showed that long delayed echoes (up to 20 ms) of remarkable strength were often interfering with signals received through the 'legitimate' path, which was either groundwave or nearly vertically incident skywave (NVIS). Even with relative strong received signals, the quality of the data was often bad, due to intersymbol interference caused by the long delayed echoes. It was then postulated that most of these echoes were due to 'ground backscatter' and that the high probability of occurrence and the remarkable strength of these signals were related to the following two conditions:

1) The short range of the link; because of this almost all distant parts of the earth's surface that are - via skywave - illuminated by the transmitting station are 'seen' by the receiving one.

2) The use of omni-directional antennas; when it is not possible to use directional antennas, scattering from all directions will contribute to this sort of interference.

In order to assess the influence of backscatter on short range HF data links, measurements were performed, which were supported by the Royal Netherlands Army (RNLA). In this paper a brief description of the HF backscatter phenomenon, a description of the measurement set-up and some results of the measurements are given.

2. HF BACKSCATTER

Radio waves propagated via one or more ionospheric reflections do not always travel from transmitting to receiving station along the great-circle path. A radio wave propagated via the ionosphere is partially scattered by the irregularities of the ground or sea, and even to some extent by those of the ionosphere itself [1]. When scattered signals propagate back to the vicinity of the transmitting station along or near the direction of incidence this is called backscatter. Propagation of scattered signals in other directions is called side-scatter.

HF backscatter was first observed in 1926. In the late 1920's and the 1930's the scattering of signals into the skip zone was encountered frequently [2]. Early researchers supposed that all scattering took place in the ionosphere [3], but in the late 1940's and early 1950's it was concluded that the most common source of backscatter is scattering by irregularities on the surface of the earth [4, 5].

Sources of backscatter can be ground backscatter, caused by irregularities on the surface of the earth such as mountains and ocean waves, or direct backscatter, caused by irregularities in the ionosphere such as F-region irregularities and auroral zone ionospheric features. Backscatter can also be caused by aircraft or meteors. On short range HF links, where the receiving station is relatively close to the transmitting one, backscattered echoes may be received from different directions and with time delays in excess of those from signals propagating along the great-circle path between both stations. The strength of these signals depends on propagation conditions and properties of the scattering region. The spectral characteristics of the backscattered signals depend upon the propagation path through the ionosphere and on whether the signal is scattered by land, sea or the ionosphere itself. In general, propagation via the ionosphere will introduce a frequency offset (Doppler shift) and a frequency spread.

Applications in which use is made of HF backscatter signals are [6]:

- HF communication along non great-circle paths (through backscatter or rather side-scatter). This makes communication possible at frequencies higher than those normally usable;

- Monitoring the coverage of HF transmissions by determining the region from which signals are scattered back. This can also be used for obtaining an estimation of

propagation conditions for real-time choice of operating frequency;

- Monitoring the structure and dynamics of the ionosphere;
- Remote sensing: monitoring sea state and corresponding surface wind and ocean current at distances up to some thousand kilometres from the observing station;
- Target detection for surveillance (over-the-horizon (OTH) radar), e.g. aircraft tracking.

Besides these applications in which a positive use is made of backscatter signals; strong scatter signals, however, may cause degradation of communication circuits. For instance, on short range HF links, where the receiving station is relatively close to the transmitting station, reception of long delayed echoes may appear due to backscatter. These backscatter signals will interfere with signals received through the 'legitimate' path. Even on low data rate links this time dispersion of signals may cause intersymbol interference.

3. MEASUREMENT SET-UP

The aim of the measurements is to investigate HF backscatter behaviour in order to assess the interference caused by these signals on short range HF data links. The most important information required, concerns statistics on parameters like path delays and signal strengths of backscatter signals and how they are related to time of day, season, operating frequency and solar activity.

In order to gather such data, path delays and signal strengths were measured over a 66 km path between The Hague (52.07N 4.23E) and Dongen (51.37N 4.56E) in The Netherlands. To assess whether short range links suffer more from scattered signals than long range links, measurements were also taken over a longer path of 958 km between The Hague and Kjeller (60.06N 10.13E) in Norway (Fig. 1). With this configuration, backscatter on a short path and side-scatter on a substantially longer path can be monitored simultaneously.

Fig. 1 Locations of measurement stations

In this measurement set-up, the station in The Hague is the transmitting station, while the stations in Dongen and Kjeller are the receiving stations. Along the two paths between the transmitting station and the receiving stations we measured path delay - in the order of milliseconds -, received signal strength and frequency of the received signal.

3.1 Principle of path delay measurement

To measure path delay, a carrier which is binary-phase-shift-keying (BPSK) modulated with a pseudo-noise (PN) sequence, is transmitted. At the receiving station the received signal is demodulated with the same PN sequence. The effective chip rate of the receiver PN sequence is somewhat lower than the chip rate of the transmitter PN sequence, so in fact the receiver PN sequence is shifted along the received bit sequence, which is BPSK modulated on a carrier. When the received signal and the receiver PN sequence correlate, the BPSK modulation will be removed and the original carrier will be restored. When there is no correlation, the output of the demodulator will be 'noise-like'. At the beginning of a measurement, both transmitter and receiver PN sequences are reset at the same time. The path delay can be derived from the time interval between the beginning of the measurement and the appearance of the demodulated carrier.

This principle is shown in figure 2. At time t_0 there is no correlation between the received signal and the receiver PN sequence. At time t_1 the receiver PN sequence is shifted one bit, but there is still no correlation. At time t_2 the receiver PN sequence has shifted one bit again and now correlates with the received signal. The restored carrier appears at the point of output of the BPSK demodulator. This carrier is detectable and the path delay can be determined.

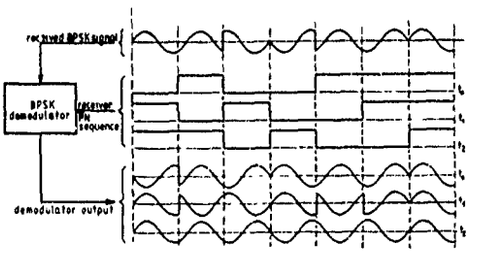


Fig. 2 Principle of path delay measurement using PN sequences

Actually, the measurement system uses a direct-sequence spread-spectrum technique. When correlation occurs, the received BPSK signal is 'despreaded' by the BPSK demodulator, resulting in a signal with a much smaller bandwidth.

3.2 Transmitter configuration

The transmitting station (Fig. 3) transmits PN sequences, which are repeated continuously. The PN sequences are BPSK modulated on an HF carrier frequency, which is transmitted by a 400 W transmitter. The transmission bandwidth is 6 kHz. The transmit antenna is a vertical whip antenna. A personal computer (PC) controls the transmitter unit.

The PN sequences are generated by a PN sequence generator implemented as an add-on card in the PC. The length of a PN sequence is 255 bit and the chip rate is 3 kbit/s, so the duration of a PN sequence is 85 ms. A DCF77 receiver, which is also implemented as an add-on card in the PC, serves as time synchronization device for both receiving stations at the beginning of each measurement. Time signals are received from the standard time transmitting station DCF77, which transmits at a frequency of 77.5 kHz from Mainflingen in Germany.

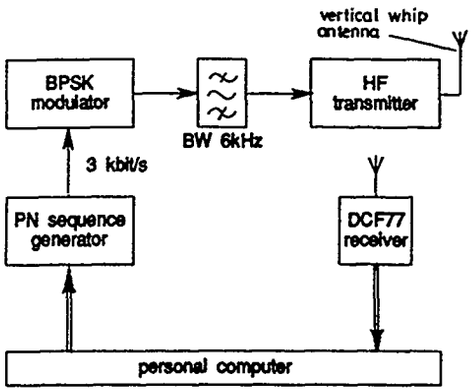


Fig. 3 Transmitter configuration

3.3 Receiver configuration

The receiving stations in Dongen and Kjeller are identical (Fig. 4). An active rod antenna is used as the receiver antenna. The receiver is a modified HF communications receiver. A BPSK demodulator is placed following the 455 kHz intermediate frequency (IF) band-pass filter with 6 kHz bandwidth. This IF signal is demodulated using a local PN sequence generated by the receiver PN sequence generator. The demodulated signal is converted further down in frequency.

Like the transmitter unit, the receiver units are also controlled by a PC. The PC is equipped with a DCF77 card for time synchronization with the transmitting station and a PN sequence generator card to generate the same PN sequence as does the transmitter PN sequence generator. The difference with the transmitter PN sequence is that from the beginning of a measurement the receiver PN sequence is retarded half a bit every 256 ms. So after N times 256 ms from the beginning of a measurement, the total delay of the receiver PN sequence in comparison with the transmitter PN sequence is N times 0.1667 ms (half bit time). When the total delay of the receiver PN sequence equals the path delay, correlation will occur and the output of the BPSK demodulator will be a sine-wave.

The bandwidth reduction factor (time scaling) of our measurement system is: 256 ms / 0.1667 ms = 1536. So when correlation occurs the received BPSK signal with 6 kHz bandwidth is 'despreaded' by the BPSK demodulator, resulting in a signal with 3.9 Hz bandwidth. The processing gain of the measurement system - the ratio between bandwidth of the transmitted signal and that of the 'despreaded' signal - is 32 dB. During calibration we found that BPSK signals down to 20 dB below the noise level were detectable with a negligible false alarm probability.

The duration of one measurement is 130.56 seconds. During this time, the path delay of all modes of propagation between transmitting and receiving station with path delays between 0 and 85 ms and with a sufficient signal-to-noise ratio will be measured with a resolution of 0.17 ms. The absolute accuracy of the path delay measurement depends on the time synchronization between the transmitting and receiving stations at the beginning of each measurement.

The frequency of the sine-wave which appears when correlation occurs, depends on the Doppler shift of the received signal due to motions in the ionosphere. Because more than one echo may appear at the same time, we must be able to deal with the possibility of more carriers showing up simultaneously on slightly different frequencies. As no one should be missed a detector that in fact is a low frequency (LF) spectrum analyzer, is used. In doing so, we also deal with frequency drift problems that occur in practice. As a consequence of using a spectrum analyzer for detection, we do not obtain a solid correlation peak, but samples of it. From these samples the original correlation peak can be reconstructed, however.

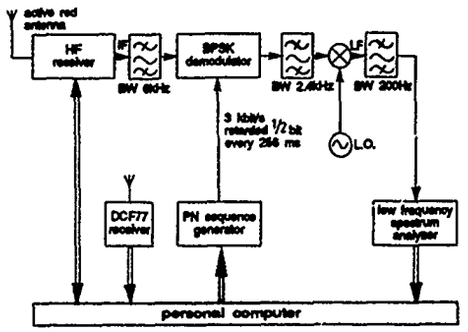


Fig. 4 Receiver configuration

3.4 Signal processing

Sine-waves which appear at the point of output of the BPSK demodulator when correlation occurs, are detected by the LF spectrum analyzer. From the output data of the spectrum analyzer - amplitude and frequency - correlation peaks are reconstructed. The path delay can be derived from the time interval between the beginning of a measurement and the appearance of a correlation peak as described in paragraph 3.1. Each correlation peak represents a signal received from the transmitting station with a specific path delay. The signal strength of the received signal is derived from the amplitude detected by the spectrum analyzer and the automatic gain control (AGC) voltage of the receiver. Since the gain of the active rod receive antenna is known, the field strength of the received signal can be calculated from the received signal strength. The frequency of the restored carrier is measured by the spectrum analyzer.

Measurement data like: time interval, strength and frequency of the detected signal and receiver AGC voltage are stored in the PC at the receiving station during the measurements. After the measurements these data are, at the request of the PC at the transmitting station in The Hague, send to this PC through modem and telephone line. In The Hague, the measurement results are further processed to obtain path delays and signal strengths for the given operating frequency on which the measurements were taken as a function of time of day.

4. MEASUREMENT RESULTS

Between June 1989 and June 1990, measurements were taken over the 66 km path in The Netherlands and, simultaneously, over the 958 km path between The Netherlands and Norway. Measurements were taken at 3 minute intervals during 16 or 24 hours a day. Operating frequencies close to 2.0, 3.5, 5.7, 7.5, 20.0 and 29.8 MHz were used.

For each measurement period of 2 hours, a plot was made of the field strength at the receiving station as a function of time of day and path delay. Each trace in these plots represents the received field strength as a function of path delay for one measurement. In these plots the different modes of propagation, with corresponding path delays, can be distinguished. In figures

6 through 11, some measurement results containing backscatter signals are shown for different frequencies in the HF band. Note that the field strengths are scaled differently in the various plots. In the box in the upper left corner of the plots, the figures of the predicted propagation mode - by IONCAP [7] - are summarised.

Figure 6 shows results of measurements taken on a frequency of 2.0 MHz during night time in late summer. On the path between The Hague and Dongen a groundwave mode of propagation is present - approximately 0.2 ms path delay - as well as backscatter signals with path delays between 6 and 16 ms (Fig. 6a). Most of the backscatter signals are dispersed between 6 and 10 ms path delay. The predicted 1F2 mode (one hop via the F2-layer) is not present. On the path between The Hague and Kjeller, a one hop mode with a path delay around 4 ms is present (Fig. 6b).

In figure 7, results are shown of measurements on 3.5 MHz taken between 3.00 and 5.00 universal time (UT) in autumn. On the path to Dongen groundwave, one hop and multi-hop signals are present (Fig. 7a). The signals with path delays between 10 and 15 ms are backscatter signals. Due to time synchronization errors, the absolute path delays of the different traces in figure 7a are somewhat dislocated. The path to Kjeller shows two modes of propagation with slightly different path delays as well as some weaker signals with longer delays (Fig. 7b).

Results of measurements taken in autumn during the evening on 5.7 MHz are shown in figure 8. There are a one hop and a two hop mode and also some backscatter signals with path delays around 12 ms on the path to Dongen (Fig. 8a). The path to Kjeller shows a one hop mode and also signals with longer delays, with a maximum of 13 ms (Fig. 8b).

Figure 9 shows results of measurements taken in summer between 21.00 and 23.00 UT on a frequency of 7.5 MHz. On the path to Dongen one hop and multi-hop modes are present and also a few backscatter signals with longer path delays (Fig. 9a). On the path to Kjeller a one hop and possibly a two hop mode are present (Fig. 9b).

Results of measurements taken on 20.0 MHz during daytime in autumn are shown in figure 10. The groundwave mode and the 1F2 mode are not present, the latter being omitted since the frequency of 20.0 MHz is above the maximum usable frequency (MUF) for the link between The Hague and Dongen. Only backscatter signals with a delay between 6 and 10 ms are present (Fig. 10a). On the path to Kjeller there is a one hop mode present, so apparently the MUF is higher in frequency than predicted (18.9 MHz) (Fig. 10b).

In figure 11, results from measurements on 29.8 MHz are shown for the path between The Hague and Dongen taken on a winter morning. Between 7.00 and 8.00 UT there are no backscatter signals, only groundwave signals are present (Fig. 11a). After 8.00 UT backscatter signals appear with path delays around 15 ms (Fig. 11b).

The measurements on one operating frequency were repeated several times each month. For example, in October 1989 measurements on 7.539 MHz were taken on seven days. For this month the cumulative field strength was derived for each measurement period of three hours. In figure 5, the cumulative field strength is shown for the measurements taken in October 1989 between 1.00 and 3.00 UT for the path between The Hague and Dongen. For a given path delay on the horizontal axis, the corresponding value on the vertical axis represents the cumulative field strength of all signals with a longer path delay than the given value. The contribution of NVIS signals - around 3 ms - and of backscatter signals - around 10 ms - is visible in figure 5. For the given operating frequency, month and time of day period, there was no contribution of groundwave signals.

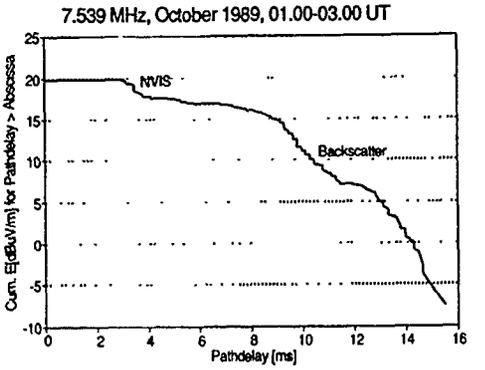


Fig. 5 Cumulative field strength (The Hague - Dongen)

On the short path a groundwave mode, a NVIS mode or both modes of propagation were present. When backscatter occurred, the difference in path delay between one of these modes and the backscatter signals was relatively large. During our measurement period, backscatter signals with path delays up to 20 ms were encountered, while the path delay of groundwave signals was approximately 0.2 ms and that of the NVIS signals was around 3 ms. Due to this large difference in path delay, the time dispersion of signals could cause intersymbol interference even on low data rate links. For a 75 Baud data link, for instance, the symbol period is 13 ms, so two signals with a difference in path delay in the order of this period will cause intersymbol interference, provided they have comparable signal strengths.

On the longer path, the principle propagation mode was one hop skywave. In general, the difference in path delay between this skywave mode and side-scatter signals was smaller than that between backscatter and the other modes of propagation on the short path. Our first impression is that the occurrence of side-scatter on the longer path is less common than the occurrence of backscatter on the short path.

The frequencies on which backscatter occurs are dictated by the frequency window for longer range links. So in the lower end of the HF band, most backscatter is observed during the night and in the higher end during the day. Most backscatter is expected in winter and our first impression of the measurement results is that this is viable. However, the analysis of measurement data to obtain statistics of path delays and signal strengths of backscatter signals as a function of season has not finished yet. Also, the statistics of backscatter as a function of solar activity for our measurement period of one year still have to be derived from the measurement data.

5. CONCLUSIONS

On a fixed frequency, path delays can be measured using a direct-sequence spread-spectrum technique by transmitting a BPSK signal modulated with a PN sequence and demodulating the received signal with the same - but retarded - PN sequence. This provides an alternative for the use of a high power pulse transmitter for measuring path delays. With the measurement system, path delays between 0 and 85 ms can be measured with a resolution of 0.17 ms and an absolute accuracy depending on

the time synchronization between the transmitting and receiving stations.

The measurement results show that particularly on short range links (<100 km), long delayed backscatter signals appear. Interference of backscatter signals with groundwave or NVIS signals may cause intersymbol interference, even on low data rate links. In the lower end of the HF band, most backscatter is observed during the night and in the higher end during the day. The occurrence of backscatter is dictated by propagation conditions for longer range links, thus depends on frequency, time of day, season and solar activity.

6. RECOMMENDATIONS

The interference on short range HF data links caused by backscatter can be avoided by technical or operational solutions. The technical solution is to design a modem that discriminates between the different components, which can have differences in path delay up to 20 ms. Backscatter signals are then used in a positive way in signal demodulation. If possible, antennas can be used which have a radiation pattern such that the strength of the received backscatter signals is suppressed. The operational solution is to make a proper choice of operating frequency. Such a choice could be based on statistics on the occurrence of backscatter.

Further analysis of the measurement data is necessary to obtain sufficient statistical background about the occurrence of backscatter. Recommendations can then be made concerning operating frequency for a given link range, time of day, season and solar activity.

7. REFERENCES

[1] CCIR, "Ground and ionospheric side- and back-scatter", (1978 - 1986), CCIR Report 726-1.

[2] Hayden, E.C., "Ground scatter in review", AGARD Conf. Proc. No. 37, August 1968, pp. 1-1 - 1-26.

[3] Eckersley, T.L., "Analysis of the effect of scattering in radio transmission", J. IEE, Vol. 86, June 1940, pp. 548-563.

[4] Dieminger, W., "The scattering of radio waves", Proc. Phys. Soc., Vol. 64-B, February 1951, pp. 142-158.

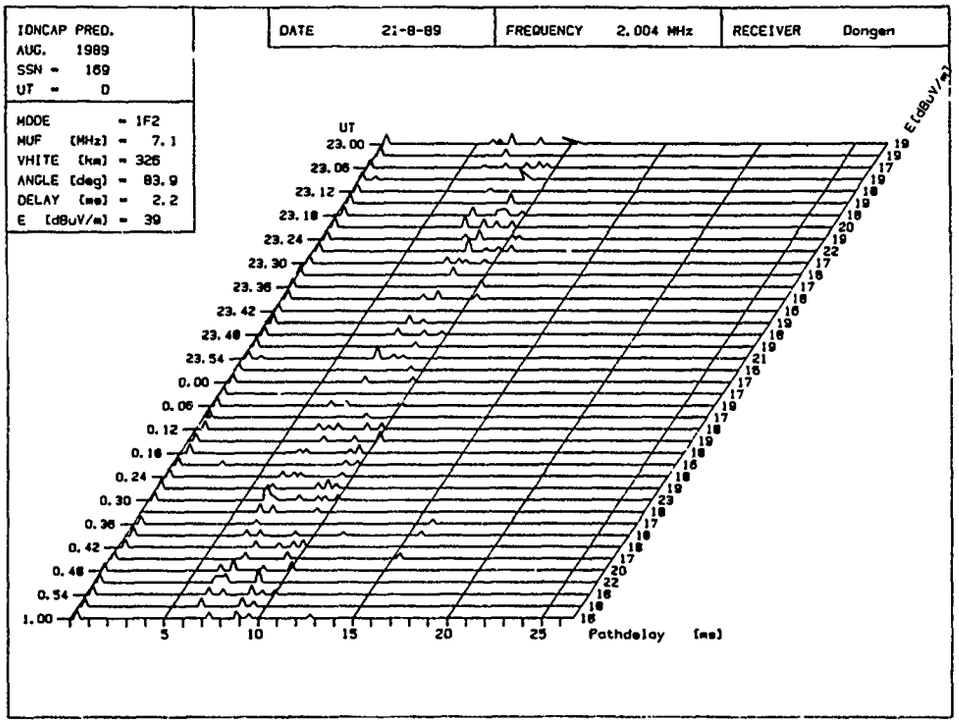
[5] Peterson, A.M., "The mechanism of F-layer propagated back-scatter echoes", J. Geophys. Res., Vol. 56, June 1951, pp. 221-237.

[6] CCIR, "The operational use of side-scatter and back-scatter", (1982 - 1986), CCIR Report 890-1.

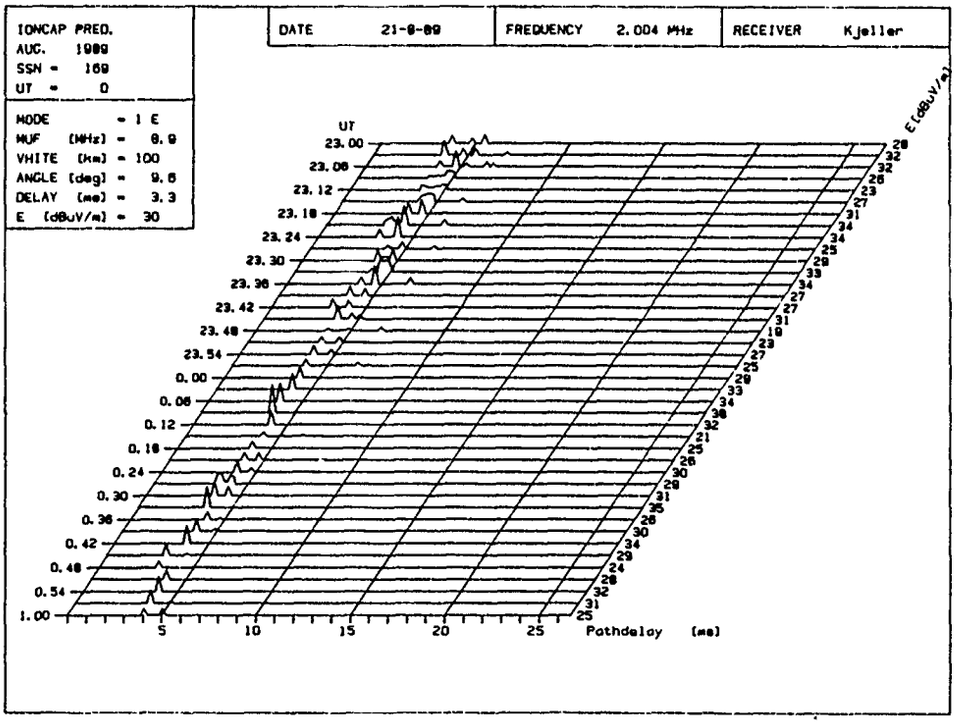
[7] Teters, L.R., Lloyd, J.L., Haydon, G.W., Lucas, D.L., "Estimating the performance of telecommunication systems using the ionospheric transmission channel", July 1983, NTIA Report 83-127.

ACKNOWLEDGEMENTS

The authors are grateful to their colleagues from the Communications Group, for their contribution to the project. In particular, we acknowledge the efforts of P.Ch. Hoefsloot, J. Niemantsverdriet, R.J.G.M. Langeveld and J. van den Oever, who played major roles in the development of hard- and software and the processing of measurement data. We also wish to thank the Norwegian Defense Research Establishment (NDRE) and the Royal Netherlands Army (RNLA) for their cooperation and hospitality, which made it possible to locate our receiving stations in Kjeller and Dongen. The research reported in this paper was supported by the RNLA, under contract number A87KL134.

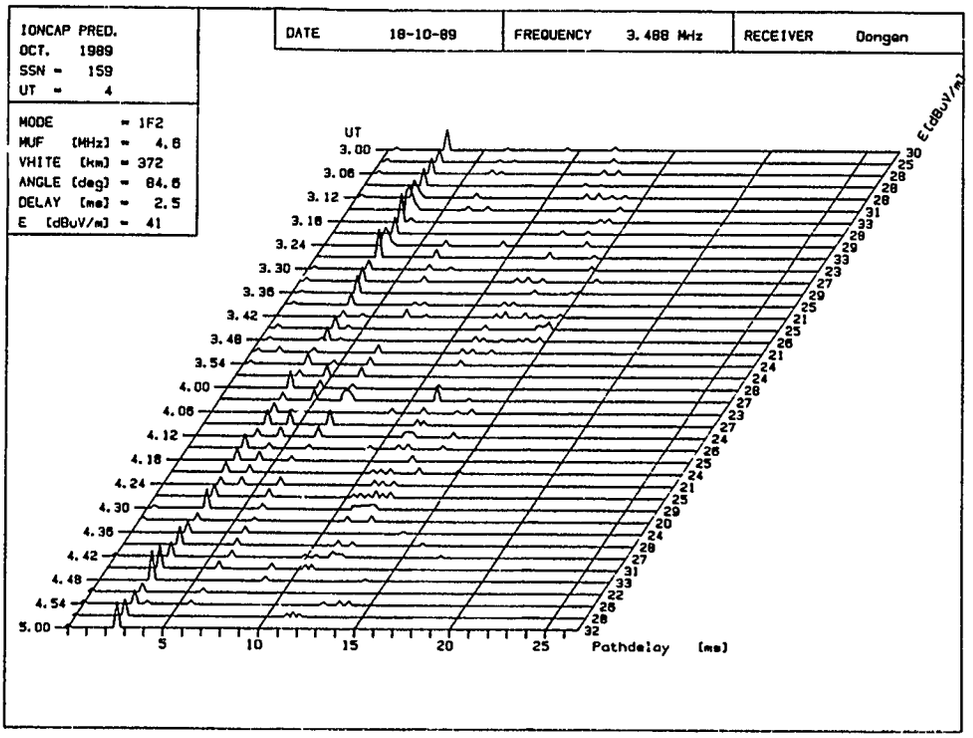


a) The Hague - Dongen

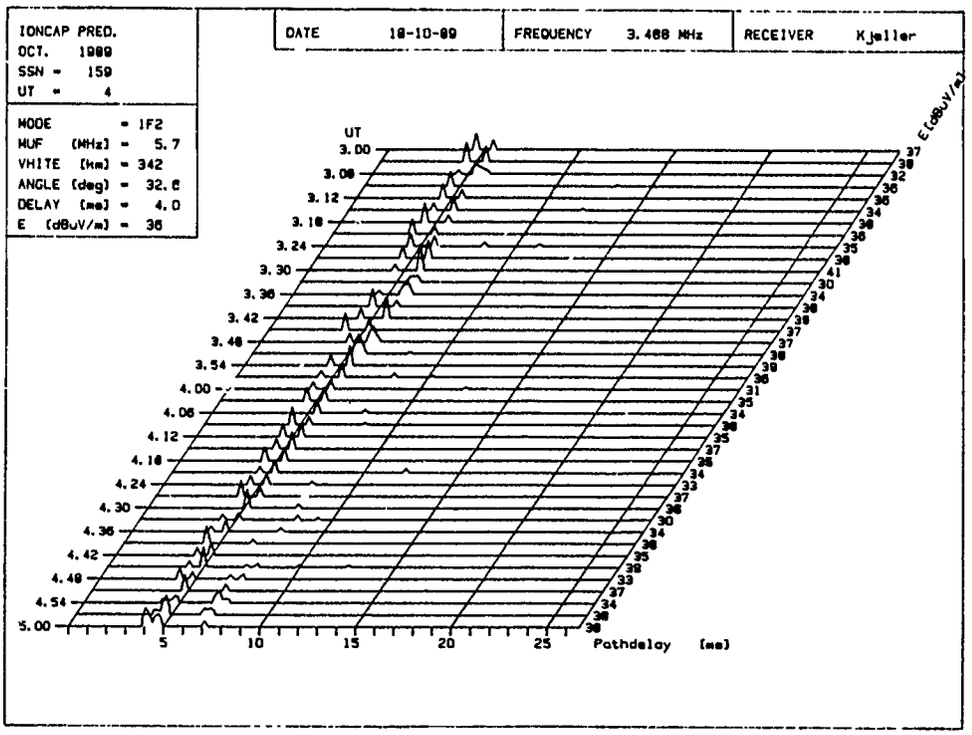


b) The Hague - Kjeller

Fig. 6 Path delay and field strength as function of time of day; frequency 2.0 MHz

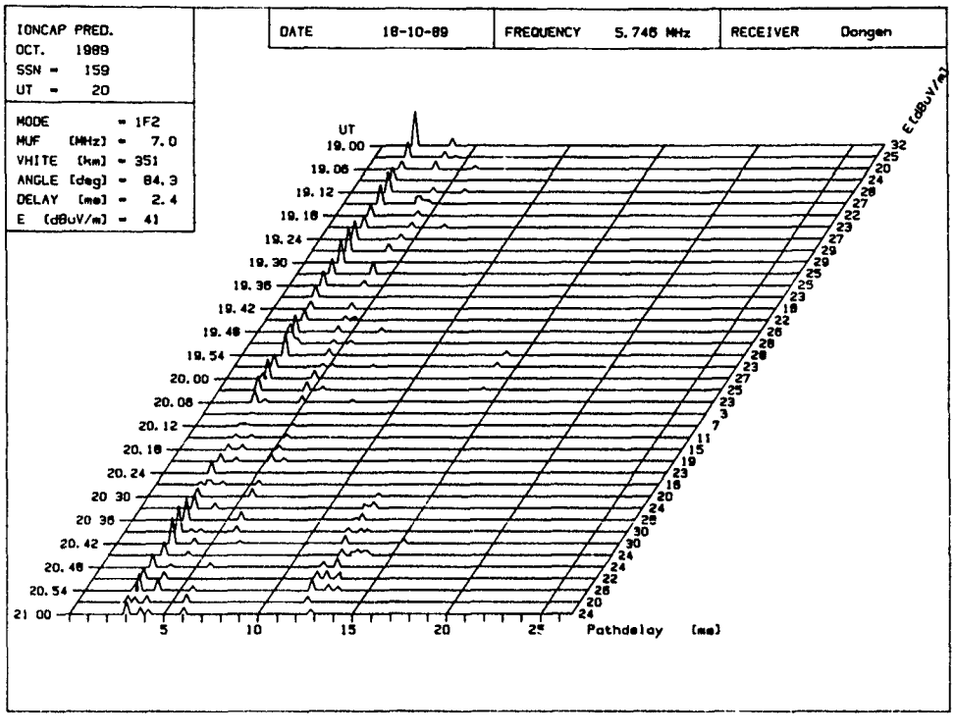


a) The Hague - Dongen

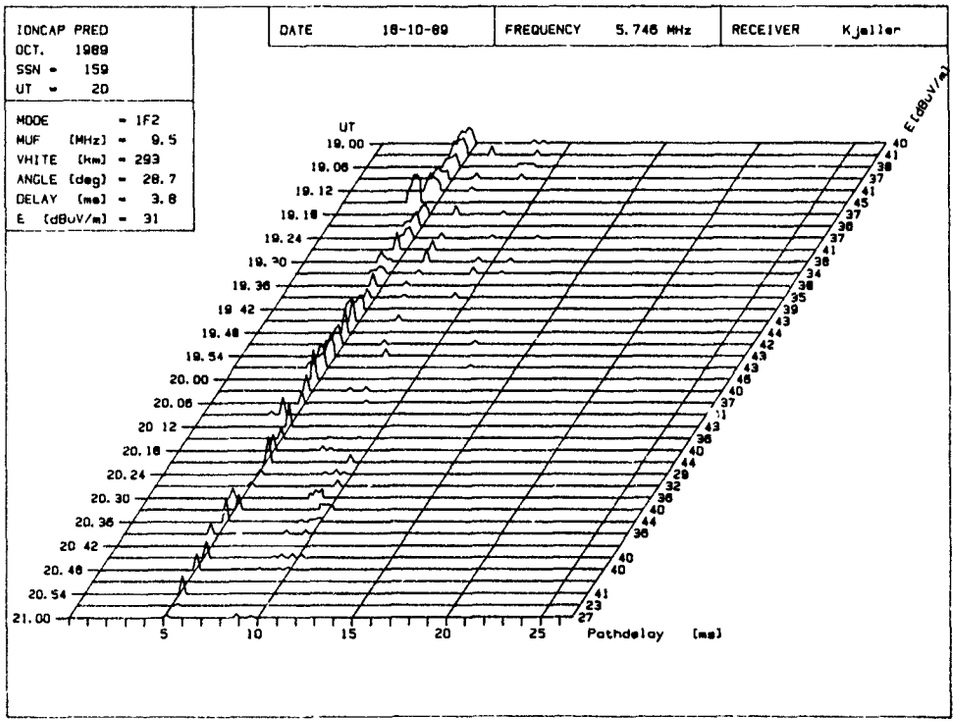


b) The Hague - Kjeller

Fig. 7 Path delay and field strength as function of time of day; frequency 3.5 MHz

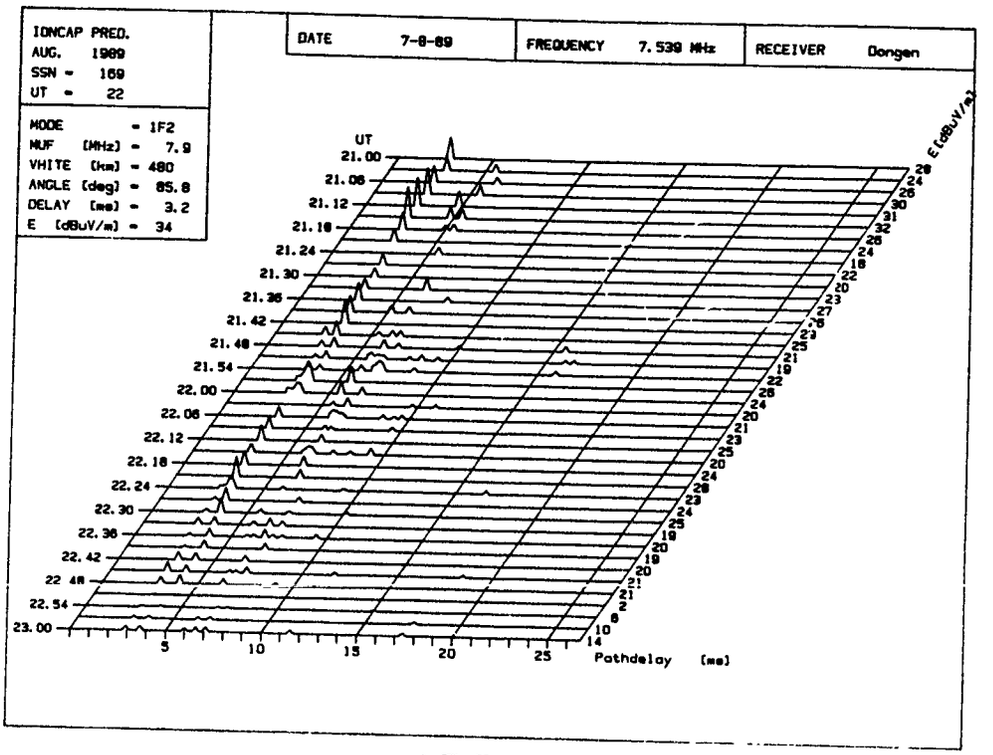


a) The Hague - Dongen

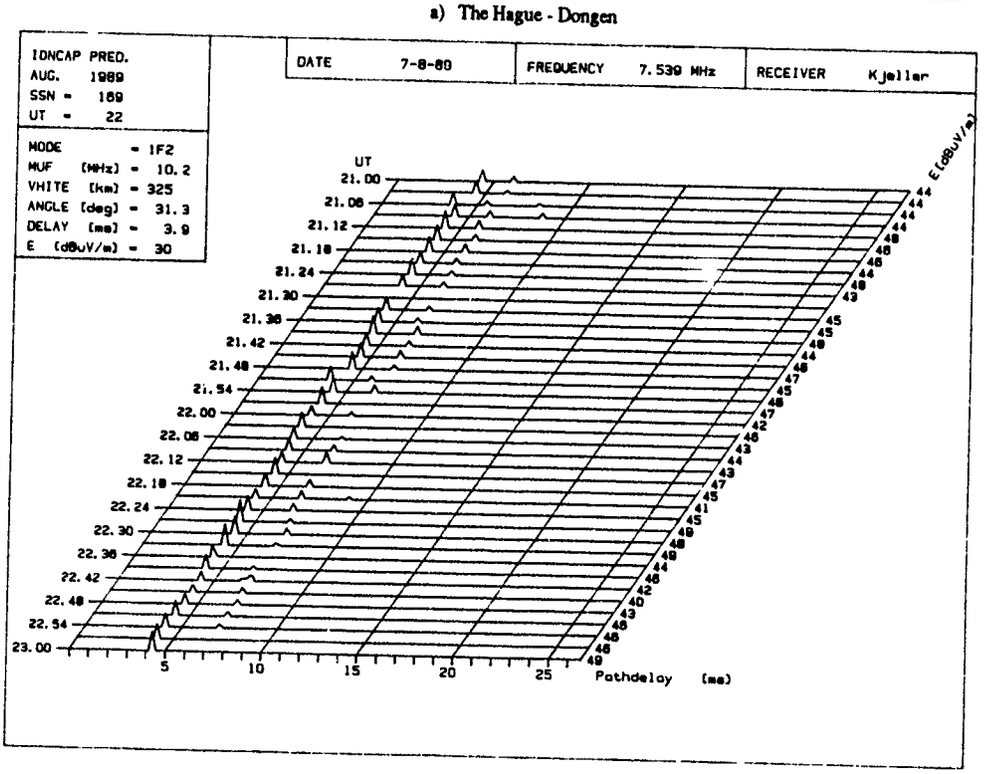


b) The Hague - Kjeller

Fig. 8 Path delay and field strength as function of time of day; frequency 5.7 MHz

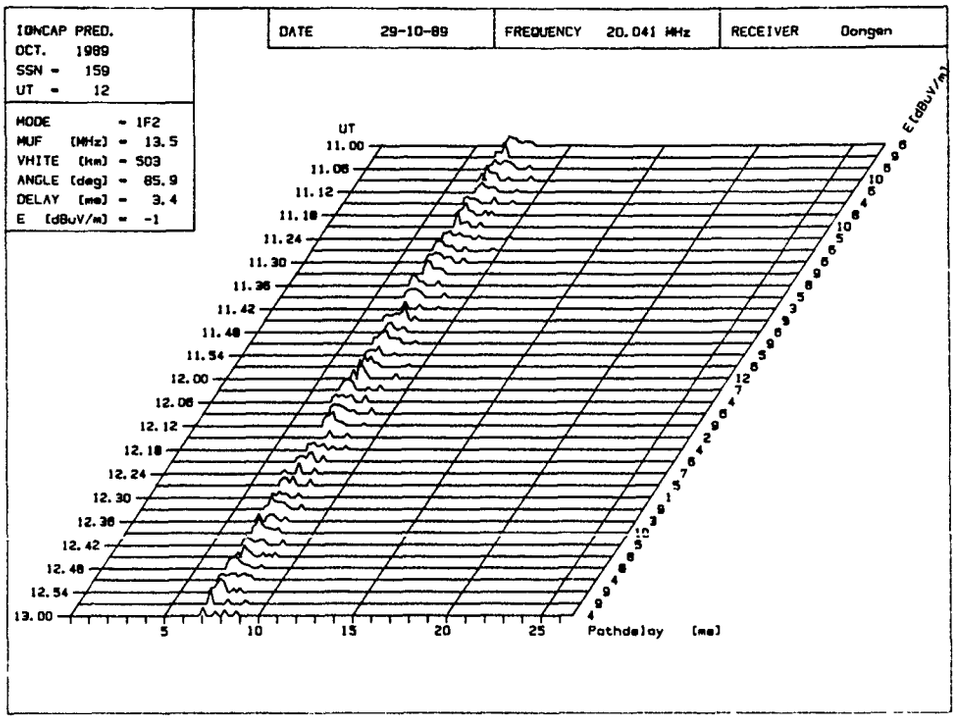


a) The Hague - Dongen

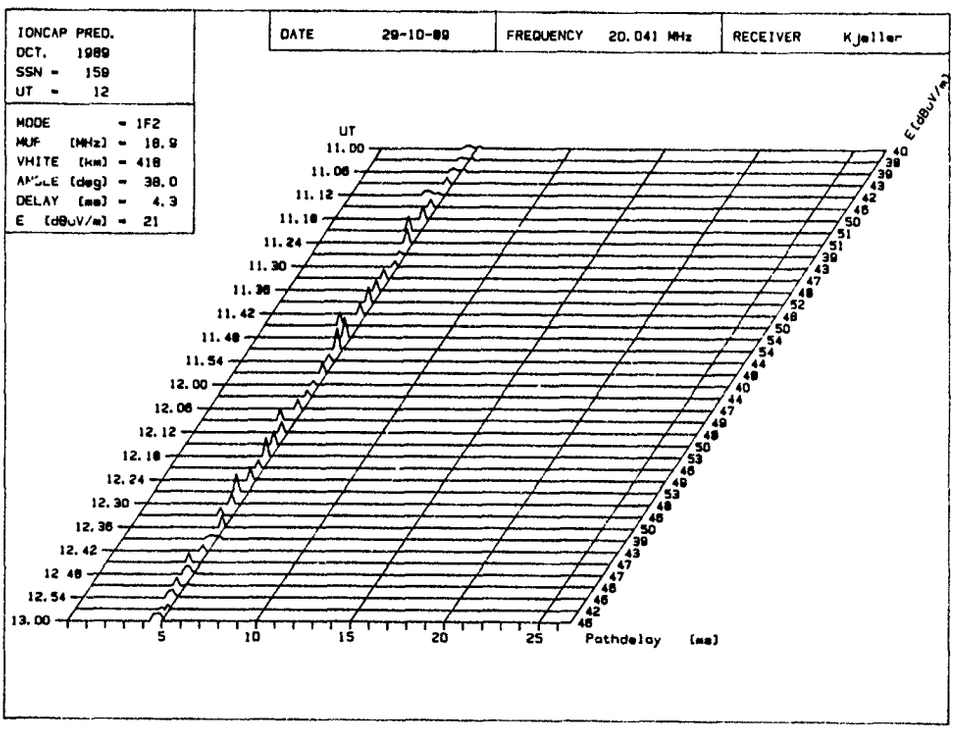


b) The Hague - Kjeller

Fig. 9 Path delay and field strength as function of time of day; frequency 7.5 MHz

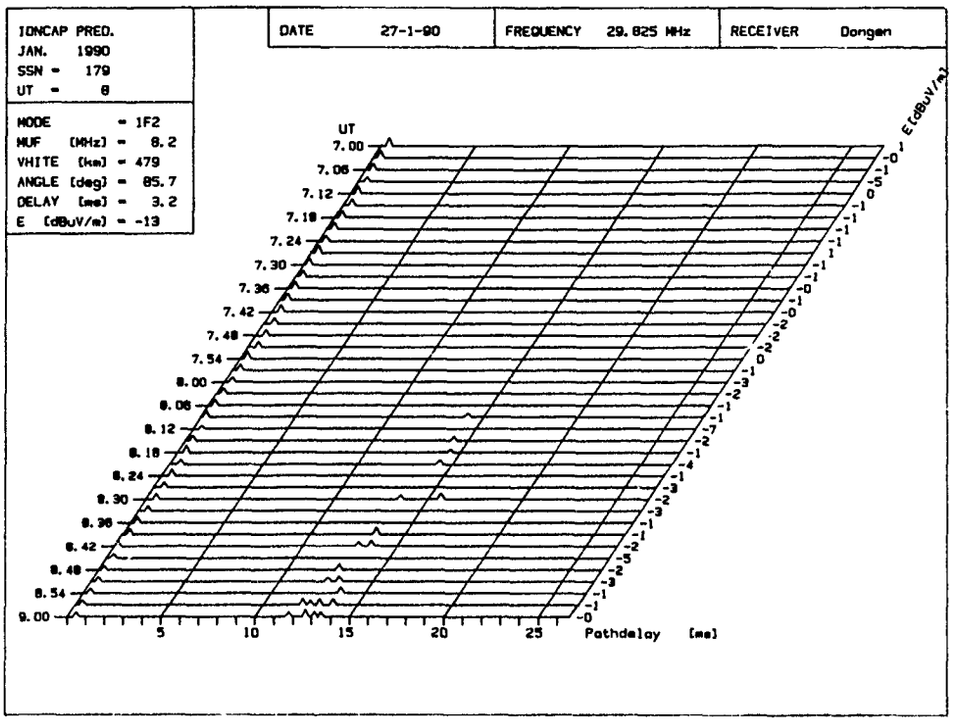


a) The Hague - Dongen

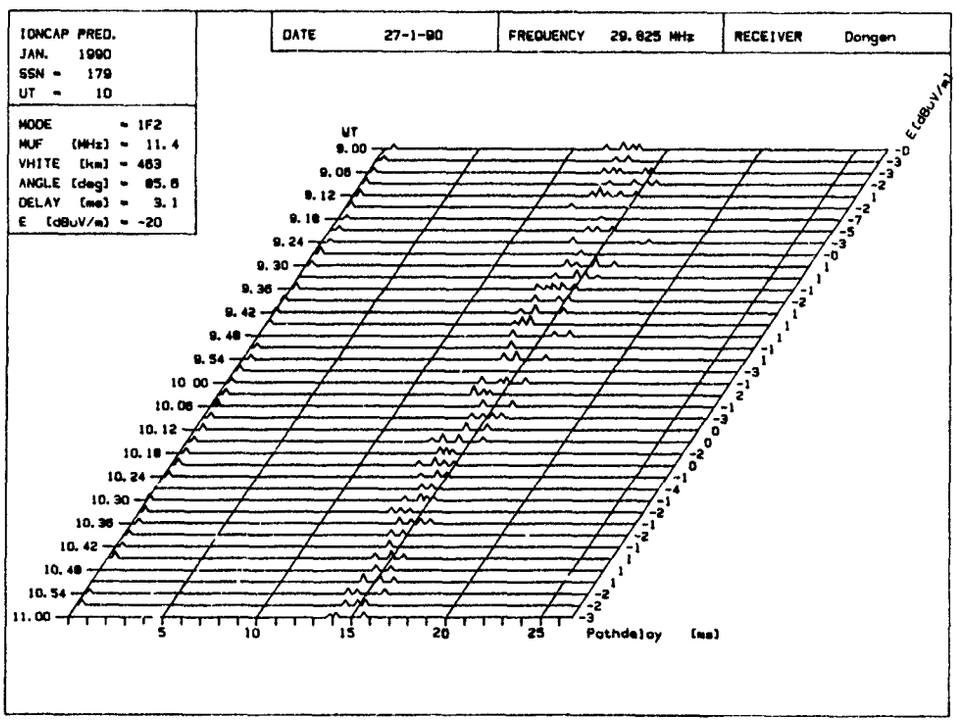


b) The Hague - Kjeller

Fig. 10 Path delay and field strength as function of time of day; frequency 20.0 MHz



a) The Hague - Dongen, 7.00 - 9.00 UT



b) The Hague - Dongen, 9.00 - 11.00 UT

Fig. 11 Path delay and field strength as function of time of day; frequency 29.8 MHz

DISCUSSION

M. DARNELL

Could you say something about the relative levels of the various received signal components, i.e. what was the level of the backscatter component as compared with the other components ?

AUTHOR'S REPLY

It is difficult to say something about the relative level of backscatter components in general. Sometimes the backscatter signal levels were comparable in strength with the groundwave or nearly vertically skywave signals. At other times, for instance, we encountered only backscatter signals, without other components. The received signal level of each component depends on the specific propagation conditions for that component, for a given operating frequency and time of day.

John S. BELROSE

I was rather astonished by your response to the previous questioner that the NVIS signal could at times be comparable with the groundwave signal for the shore path over which you have made measurements. The vertical whip should help to reduce such a problem, due to the fact that it should have an overhead null (-15 to -25dB).

Could you give more detail on the specific transmitting antenna employed.

AUTHOR'S REPLY

The aim of the measurements was to invertigate HF backscatter, in order to assess the interference caused on operational links. For this reason we used a 4½ metre vertical whip antenna, which is similar to the antennas used by the Royal Netherlands Army.

In fact, sometimes the received level of backscatter signals was comparable with the received level of groundwave or NVIS signals. Since the analysis of our measurement results has not finished yet, it is not possible to give statistical data on the difference in received field strength, between the different components, yet.

OPTIMUM ANTENNA SPACING FOR DIVERSITY IN METEOR BURST COMMUNICATIONS SYSTEMS

by

Paul S. Cannon
Anil K. Shukla
Applied Ionospheric Physics, Flight Systems Department
Royal Aerospace Establishment, Farnborough,
Hampshire, GU14 6TD, United Kingdom

Mark Lester
Ionospheric Physics Group
Leicester University
Leicester, LE1 7RH, United Kingdom

SUMMARY

Temperate latitude meteor burst (MB) space diversity measurements, using cross correlation analysis of 37 MHz signals scattered over an 800 km path, are reported. Signals of duration ≥0.75 s, received on antennas separated by 5λ, 10λ and 20λ, are investigated using 6 days of data collected over a 9-day period during February 1990. Signal decorrelation is shown to be achieved by an antenna separation of 5λ and there is no apparent variation in average cross correlation coefficient, for antenna separations between 5λ and 20λ, for any of the signal categories examined. We discuss in some detail previous diversity studies by Bartholomé[1] and Ladd[2].

1 INTRODUCTION

It is mathematically convenient, when first investigating meteor burst (MB) propagation, to assume that the incident wave has been scattered from idealised straight columns of ionisation created by incoming meteors[3]. Implicit in this formalisation is a single, planar wave incident on the receiving antenna.

Mesospheric wind shears and turbulences can, however, radically alter the initial shape of the trail and may, as a consequence, cause the formation of multiple scattering regions (or glints)[4]. These glints may move independently and each scattered signal will experience a different Doppler shift. Fading will then be observed due to the superposition of two or more discrete Doppler shifted frequency components. The aggregate wavefront is now clearly no longer planar and a different time dependent signal fading pattern will be observed on appropriately spaced antennas. Consequently, it may be possible to exploit diversity, *eg* Ref 5, techniques, similar to those used in the HF (high frequency) band, to enhance the performance of MB communications systems. A pre-requisite for the implementation of diversity is the existence of uncorrelated fading at two, or more, of the receiving antennas. Under these circumstances a faded signal from one antenna may be compensated for by in phase addition of the signal components at the second antenna. One method of obtaining these uncorrelated signals is by reception at two appropriately spaced antennas (space diversity).

Investigation of MB space diversity on long duration (~4 s) echoes was performed by Ladd[2]. An operating frequency of 49 MHz and a 1 kW transmitter were used with signals received on two antennas separated by either 22λ or 60λ. These signals were recorded and digitised using a 100 ms sampling interval. The correlation coefficients were computed for underdense, specular overdense and non-specular overdense signals. The high values obtained (0.9946, 0.9819 and 0.7970 respectively) led Ladd[2] to conclude that little space diversity gain could be obtained. However, fading periods on signals in this frequency band range typically[6] between 100 ms and 1 s and the 100 ms sampling period adopted would, therefore, have been too infrequent. The cross correlation coefficients were also computed over complete trail durations. Similar analyses performed by Cannon *et al*[7] showed the correlation coefficient to be a function of the time into the trail decay. In this latter work high correlation coefficients were reported, as expected[8] at the start of the trail decay, with the coefficients decreasing towards the end of the trail lifetime. The conclusions appear to place those of Ladd[2] in question.

A cross correlation coefficient of 0.6, or below, was adopted by Ladd[2] as a 'practical' indication of useful decorrelation[9]. Such a condition is, however, only applicable to Rayleigh fading conditions. Extensive studies have convinced these authors that Rayleigh fading between multiple glints is only one of many situations that can occur. Interference between two, or a small number of glints, resulting in a deep periodic fading envelope which approximates a rectified sine wave is more likely. Greenhow[10] suggested that approximately 400 ms are required for the formation of the first glint. This then restricts the applicability of the Rayleigh fading model, requiring the formation of multiple (>6) glints, to long enduring meteor trails. It is misleading, therefore, to base MB diversity conclusions on simply Rayleigh fading assumptions. The deep periodic fading patterns resulting from two or three glints[8] also relies on the formation of glints of similar dimensions, a situation which may not prevail. Superposition of signals from smaller scattering regions and the main scattering body will still, however, cause an amplitude fluctuation as opposed to deep periodic fading.

Bartholomé[1] suggested that the conclusions of Ladd[2] were pessimistic. He performed experiments in the 38 to 41 MHz band with 400 watt transmissions and an antenna separation of 4λ in a line perpendicular to the great circle path from the transmitter to the receiver. The diversity effect reported was small both for underdense and overdense specular reflections. However, when the trails were long (>2 s) and wind distortion had led to the creation of several glints, the resulting multiple wavefronts added

vectorially to produce quasi-independent time varying signal fields at each of the separated antennas.

The only fully developed MB communications system reported in the open literature which has incorporated diversity is COMET[11]. The COMET system, which implemented antenna space diversity, together with diversity combining of signals from antennas mounted at two different heights, showed a remarkably high duty cycle compared with other MB communications systems[12]. The reasons for this high throughput were never formally proven but Bartholomé and Vogt[11] believed that diversity contributed a considerable degree. If so, this is in conflict with the conclusions of Ladd[2].

In summary, the literature of the 1960s is confused with regard to the benefits of space diversity in MB systems. MB data communications are generally regarded as low data rate systems, with bit rates varying from a few, to generally only tens of bits per second[12]. The possibility of improving this by using spaced diversity techniques cannot be ignored and as a consequence further work has been pursued more recently. Cannon *et al*[7] and Shukla *et al*[13] have clarified the situation on a temperate latitude 37 MHz path by showing that although the correlation coefficient of the total signal envelope is high, as found by Ladd[2], the correlation coefficient falls as a function of time into the trail. Shukla *et al*[13] categorised the received signals into three groups, ie underdense, overdense and not known (NK). Of the three signal types NK signals were observed to be less correlated than underdense or overdense trails. At an antenna spacing of 10λ, when all three categories were combined, 40% of all signals exhibited cross correlation coefficients of <0.6 after 0.25 s of signal decay. In this paper we will address an aspect of the work reported by Shukla *et al*[13] namely the optimum antenna spacing in such a diversity system.

Whilst this paper addresses meteor burst signals it must also be recognised that three other sources of uncorrelated signal exist in the MB frequency band; namely ionoscatter, sporadic-E and propagation via the F2 layer during periods of high sunspot activity. The magnitude and occurrence of signals propagated via these modes is dependent on mechanisms with various temporal scales and on the radio frequency of operation. These other modes can add vectorially with MB signals providing the possibility of diversity gain. Similar conclusions were reached by Bartholomé[1].

2 EXPERIMENTAL TECHNIQUE

The experiments performed consisted of the transmission of a 37 MHz, 400 W continuous wave (cw) signal over an 800 km path. The transmitter was located (Figure 1) in Wick, Scotland (58.56 N, 3.28 W), and signals were received at Cobbett Hill Radio Station in Southern England (51.27 N, 0.63 W). In order to minimise the reception of sporadic-E signals which might confuse the study of MB signals, the experiment was performed in February. Morning and afternoon data were collected for three antenna separations of 5λ, 10λ and 20λ (Table 1). These separations compare to 22λ and 60λ used by Ladd[2] and the 4λ separation implemented by Bartholomé[1]. The receiving antennas were sited in a line perpendicular to the great circle path from transmitter to the receiver. Both the transmitting and receiving antennas were horizontally polarized 4-element Yagis, mounted at a height of 1λ which is suitable for centre path illumination at an altitude of 95 km.

The received signals from the two antennas were fed, without amplification, down two lengths of loss equalised cable to the receiving and signal monitoring system, Figure 2. The two down-converter[14] output signals were fed into separate RACAL 1792 HF receivers operated with a bandwidth of 3 kHz. The two 100 kHz intermediate frequency output from the receivers were recorded on two FM channels of a RACAL STORE 4DS tape recorder after passing via a unit described as a 'Log Detector'[14]. The latter units output the logarithm of the detected input signal level. These signals were recorded in a 1.25 kHz bandwidth using a tape speed of 3.75 inches per second.

In order to derive absolute signal levels, each one hour data tape was calibrated every 28 minutes with signals ranging in voltage equivalent power levels from -133 dBm to -76 dBm in 3 dBm steps.

3 DIGITAL ANALYSIS TECHNIQUE

Data capture

A computer program, configured for the Hewlett Packard HP 9836 desk top computer, was written to store and analyse the analogue recorded signals. The program enables two signal voltage channels (eg from spaced antennas in this case), to be sampled and stored by an 8 bit analogue to digital converter. The two adjacent signal channels are sampled within 40 μs. This then approximates simultaneous sampling. The anti-aliasing filter is a variable low pass filter with a 48 dB per octave roll off, Figure 3.

These digitised data points represent received signal power in dBm and are derived from the digitised analogue signal voltages using the calibration data. The data files can be displayed to show signal strength (dBm, ordinate axis) against time (seconds, abscissa) for a selection of time windows. The software suite is described in detail by Shukla[15]. By using a vertical sliding time cursor meteor signal envelopes, above a fixed power threshold, can be visually identified and easily characterised. The software also calculates the cross correlation coefficient of the two signals over a time segment specified by the user. Calculated correlation values can then be stored to disc with an identifying label which best describes the category of signal currently under investigation.

Data analysis

Data collected at the three antenna separations, were analysed according to the preceding description using a channel sampling rate of 100 Hz (10 ms) and an anti-aliasing filter bandwidth of 25 Hz. This sampling rate is ten times higher than that used by Ladd[2]. A threshold of -120 dBm (approximately 10 dB above the noise floor) was used for signal analysis. All signals above the analysis threshold were categorised as 'overdense', 'underdense' or 'not known (NK)'. The category 'NK' refers to that group of signals which may have propagated via sporadic-E, ionoscatter, or via non specular meteor trails. Cross correlation coefficients for zero lag ($\tau = 0$) were then evaluated for these signals. Since meteor decay is a non-stationary process cross correlation values of $|\tau| > 0$ are invalid.

In analysing the underdense and overdense trails the correlation start time corresponds to the end of trail formation. This can be identified by the discontinuity on the rising edge of both underdense and overdense signal envelopes. Those two channel data within the first 0.25 s of formation were correlated and the cross correlation value recorded along with the trail category (Figure 4a). Successive 0.25 s segments were windowed and correlated until the end of the trail was reached. The latter is taken to be that time when both channels fall below the pre-set analysis threshold. The decay period is, therefore, determined by the longer duration channel opening time. This definition incorporates the important period when only one channel is above the threshold, during which time diversity is expected to give its greatest improvement in channel availability.

Greenhow[6] has shown that the frequency range of deep periodic fading from meteor trails is 1-10 Hz and an analysis segment size of 0.25 s is, therefore, sufficient to encompass most fades. Correlation values from the first 0.25 s form the statistics for segment 1, correlation values from the second 0.25 s form the statistics for segment 2, and so on.

NK signals show no discontinuity near their start and consequently start times were identified by the first signal crossing of the analysis threshold. Otherwise the analysis was the same as underdense and overdense trails.

In data analyses successive and grouped segments were compared. Implicit in this approach is the assumption of an average ionisation height and location. This averaging ignores the differing geometrical factors which affect the decay of meteor trails occurring within the antenna common volume. In addition to time segmentation of the signal, cross correlation analysis was also performed over the total signal decay period (Figure 4b) to compare our results with those of Ladd[2].

Trail categorisation

Figure 4 is illustrative of the three trail categories to which the digitised signals were assigned. Similar trail categorisation was performed by Østergaard *et al*[16] and this was used as a basis for ordering our data.

An underdense trail is characterised by a short signal rise time and a slower linear decay time (in dB) which starts immediately after the envelope has reached its maximum amplitude. An underdense trail is illustrated in Figure 4a where the decay duration on channel 1 is greater than channel 2 by an amount Δ. In this example selection diversity would be advantageous.

A specular overdense trail is depicted in Figure 4b. It is characterised by a fast rising increase in signal strength until the Fresnel zone is formed. This is followed by a further period of slowly increasing signal strength, resulting in a rounded top to the received envelope. Overdense decay durations and received signal strengths are usually greater than underdense trails. In the example Figure 4b, selection diversity would show little advantage.

The 'NK' category of signal have an undefined format and may be associated with signals scattered from sporadic-E, ionoscatter signals, or non-specular meteor signals. Figure 4c shows short duration, highly uncorrelated waveforms. There is a delay of Δ2s between the signal crossing the analysis threshold on channel 1 and the signal crossing on channel 2. There are also a number of NK signals which fall into a group suffering from echo overlap decorrelation[17]. These signals occur when two trails form within the lifetime of each other and within the antenna common volumes. Figure 5 illustrates such a case. Between times T_1 and T_2 a single underdense trail exists but at T_2 a second meteor trail forms. The signals add or subtract at the antennas, depending on their relative path difference, resulting in the envelopes seen in Figure 5. These trails are grouped in the NK category due to their unusual and rare nature. Signals scattered from remnants of trails which were not previously in the antenna common volume, or trails which have been severely distorted and which are not easily recognised are also grouped in this category.

Figure 6a&b illustrate the deep periodic fading which derives from [Ref 4] two (or more) Doppler shifted components beating together to produce a |sin t| modulation pattern. The total decay duration of the underdense trail (Figure 6a) is ~2.5 s. In a conventional communications system, however, this single trail would be interpreted as three different trails due to the three deep (~10 dB) fades in the trail decay. The potential data throughput from such a trail is then not fully exploited. In this particular example of an underdense trail, diversity would not be advantageous due to the phase coherency of the deep fades on the two antennas. Conversely, although a similar fading pattern is observed for the overdense trails in Figure 6b, there is a fade time delay between the two channels of approximately 10 ms. Diversity systems can exploit this fade delay by combining the signals to reduce the fade depth. The combined signal would enable a MB communications system to utilise the total decay duration, thereby increasing the system data throughput.

4 RESULTS: SPATIAL CORRELATION VARIATION

An important question relating to the implementation of any space diversity receiving system is the optimum distance required between the antennas. A large antenna separation may be required to obtain sufficient signal decorrelation. In order to minimise the amount of land required to deploy the antennas the minimum separation commensurate with signal decorrelation is, however, considered to be the optimum separation.

In order to determine this unknown optimum spacing, investigations were performed on signals received on antennas separated by 5λ, 10λ, and 20λ (Table 1). Six days of data were collected spread over a 9-day period. At each antenna spacing the three signal categories of duration $\geq$ 0.75 s were identified and analysed. All data for each antenna separation were combined in order to increase the trail count.

Underdense trails

Between 20 and 40 trails of duration $\geq$0.75 s were identified over the 2 days at each antenna separation. Cumulative correlation distributions values, from underdense trails calculated over the first 0.25 s (segment 1), at the three antenna spacings are plotted in Figure 7a. Similar distributions are also plotted (Figure 7b&c) for segments 2 and 3.

The three piecewise linear distribution curves plotted for segment 1 (Figure 7a) are similar in shape and no consistent order to the distribution curves is observed as a function of antenna separation. The maximum distribution spread between any of the three curves is approximately 10%; this represents a difference of only three trails in the 5λ data

set. Consequently, within the limits of our data set we conclude that in segment 1 the decorrelation of underdense trails shows no increase or decrease with antenna separations between 5λ, 10λ and 20λ. That is antenna separations of 5λ, 10λ and 20λ should provide very similar space diversity gains.

The absence of spatial variability is also noted for segments 2 and 3 (Figure 7b&c). Although an increase in decorrelation occurs with increasing segment number, there is again no consistent order in the correlation distribution curves obtained from underdense signals at the three antenna spacing.

Overdense trails

Cumulative distributions for the three segments at the three antenna spacings are plotted in Figure 8. The top two frames show that the correlation values of the signals received at 5λ , 10λ, and 20λ are closely distributed in both segments 1 and 2.

In segment 3 (Figure 8c) approximately 20% of the trails at 5λ have correlation values less than 0.6. This contrasts with 15% at 10λ and 32% at 20λ. The maximum distribution spread occurs at a correlation value of 0.8 between the 20λ and 10λ curves and is 25%. This represents 11 trails of the 10λ data set which is a fairly significant trail count.

We conclude that, within the limits of our data set, the decorrelation of overdense trails shows no variation with antenna spacing between 5λ, 10λ and 20λ for segments 1 and 2. Correlation values for segment 3, however, appear to favour the 20λ antenna spacing, although, it is unclear if this is due to the limited data sets involved.

NK signals

Preliminary studies of the 5λ data revealed that NK signals analysed on day number 44 were significantly less correlated than NK signals analysed 9 days later on day number 53. Within the 6days of data considered, those signals recorded on day 44 showed an uncharacteristic amount of multipath. Cross correlation statistics from these 2 days were not, therefore, combined but kept separate and are thus presented in the following diagrams. The multipath on day 44 had little impact on the analysis of underdense and overdense trails since, by our definition, for selection these were required to meet certain exacting requirements relating to the signal envelope with all other signals categorised as NK.

In our discussions we firstly exclude data from the 5λ spacing on day 44 where the correlation values are influenced by severe multipath. For any given segment the correlation distributions (Figure 9a-c) at the three antenna spacings are broadly similar to each other. It is also apparent that NK values are more uncorrelated than the underdense or overdense signals. The absence of a systematic variation in signal correlation with antenna spacing, in any of the three time segments, essentially mirrors the results from the other two signal categories.

Returning now to day 44 we observe that the effects of the multipath are dramatic. The probability of finding decorrelated signals on this day in segments 1, 2 or 3 is significantly greater than on the other 5 days of the experiment. On day 44 the space diversity gain should be high.

5 CONCLUSION

This paper has investigated space diversity of 37 MHz scattered signals. The spatial variation of correlation values, a parameter important to system designers, was addressed for antenna spacings of 5λ, 10λ, and 20λ. All of the following conclusions relate to trails of duration ≥0.75 s.

The spatial variation of received signal correlation values was investigated at three antenna spacings of 5λ, 10λ, and 20λ. Little spatial variation was observed for any of the signal categories and it appears that decorrelation is achieved at very modest spacings of 5λ or 10λ. The 20λ separation is less practical and unnecessary if compact MB systems are required. On the basis of our measurements the 4λ diversity antenna separation implemented by Bartholomé and Vogt[11] in COMET would probably have been as successful as a much larger antenna spacing.

Our measurements were made with a fixed transmitter power of 400 watts. Variation of this power will change the relative numbers of underdense, overdense and NK trails present. Given, however, that all three trail categories show no significant decorrelation dependence with antenna spacing, we expect that the above conclusion regarding the optimum antenna spacing will not vary as a function of transmitter power.

An important caveat to this work relates to the latitude of signal measurements. At high latitudes in particular, where anomalous propagation modes are more common, different conclusions may be obtained.

In conclusion, the results presented in this paper suggest that space diversity will be advantageous, especially when signals propagate via NK trails. A diversity contribution can, however, also be expected from overdense and underdense trails particularly towards the end of their trail lifetimes.

Table 1

Timetable of Data Collected from Cobbett Hill

Day number	Antenna separation (λ)	Hours of data collected
44	5	5
45	10	4
46	20	5
47	20	4
51	10	5
53	5	4

REFERENCES

1 Bartholomé, P.J.; The fine structure of meteor-burst signals, Technical Memorandum, TM-36, Shape Air Defence Centre (1962)

2 Ladd, A.W.; Diversity reception for meteoric communications. *IRE Trans. Com. Sys.* ,9(2): 145-148 (1961)

3 Sugar, G.R.; Radio propagation by reflection from meteor trails. *Proc. IEEE.*, 52(February): 116-136 (1964)

4 Manning, L.A., Villard, O.G.J. and Peterson, A.M.; Double Doppler study of meteoric echoes. *J. Geophys. Res.*, 57, 387-403 (1952)

5 Schwartz, M., Bennett, W.R. and Stein, S.; 'Part III'. Communications Systems and Techniques. McGraw-Hill. New York (1966)

6 Greenhow, J.S.; The fluctuation and fading of radio echoes from meteor trails. *Phil. Mag.*, 41, 682-693 (1950)

7 Cannon, P.S., Tyler, J. N.,Shukla, A.K. and Dickson, A.H.; Space Diversity Reception in Meteor Burst Communications System. *IEE Conference on HF Radio Systems and Techniques*, 284, 110-114 (1988)

8 Manning, L.A.; Air motions and the fading, diversity, and aspect sensitivity of meteoric echoes. *J. Geophys. Res.* , 64(10), 1415-1425 (1959)

9 Staras, H. ; Diversity reception with correlated signals. *J. of Appl. Physics*, 27(1) (1965)

10 Greenhow, J.S.; Characteristics of radio echoes from meteor trails:III The behaviour of the electron trails after formation. *Proc. Phys. Soc.* , 65B(3), 169-181 (1952)

11 Bartholomé, P.J. and Vogt, I. M.; COMET - A new meteor burst system incorporating ARQ and diversity reception. *IEEE Trans on Comm.* , Com-16(2), 268-278 (1968)

12 Cannon, P.S. and Reed, A.P.C.; The evolution of meteor burst communications systems. *J. Inst. Elec. Rad. Eng.* , 57(3), 101-112 (1987)

13 Shukla, A.K., Cannon, P.S. and Lester, M.; Spaced antenna diversity in temperate latitude meteor burst systems: variation of signal cross correlation coefficient at 37 MHz. To be submitted to Rad. Sci. (1991)

14 Dickson, A. H., Cannon, P.S., Tyler, J.N. and Bowker, I.H.; Technical description of the RAE, BLOSSOM-A, meteor burst communications system. Unpublished RAE Technical Report (1987)

15 Shukla, A.K.; Improved analysis of meteor burst data using cross correlation techniques, RAE Report (to be published)

16 Østergaard, J.C., Rasmussen, J.E., Sowa, M.J. , Quinn, J. M. and Kossey, P.A.; Characteristics of high latitude meteor scatter propagation parameters over the 45-104 MHz band. *AGARD Conference on Propagation effects on military sytems in the high latitude region.* Vol.382 (1985)

17 Berry, J.B.J., James, J.C. and Meeks, M.L.; Studies of meteor propagation at 49 MHz and 74 MHz. *IRE Trans Ant. & Prop.* AP-9(4), 395-403 (1961)

Figure 1. Location of the transmitting and receiving sites.

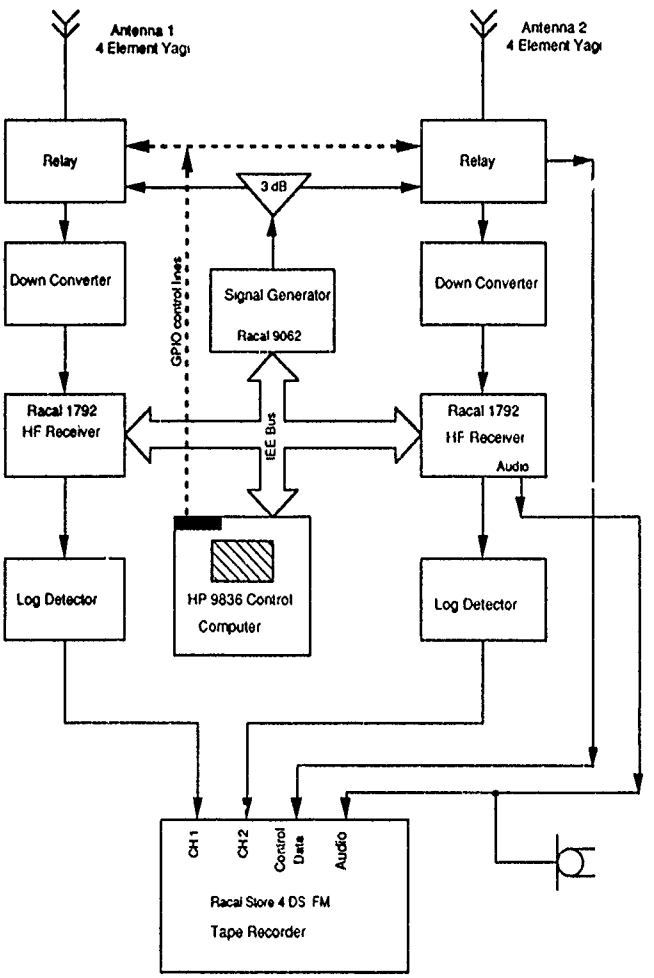


Figure 2. Block diagram of the receiving system.

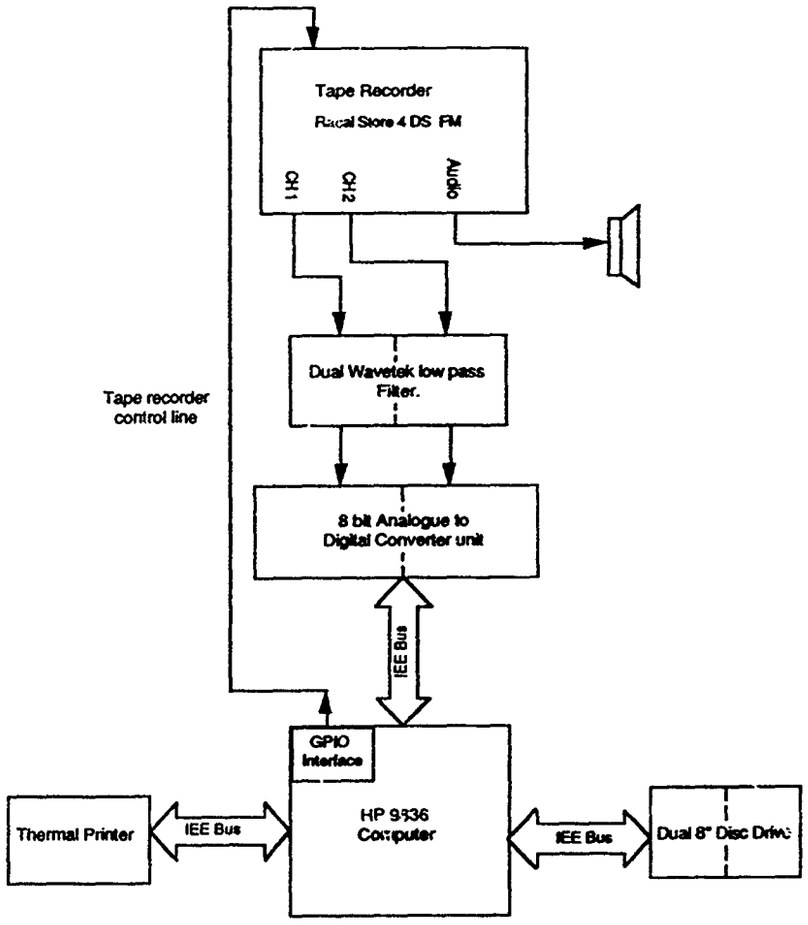


Figure 3. Block diagram of the digital analysis system.

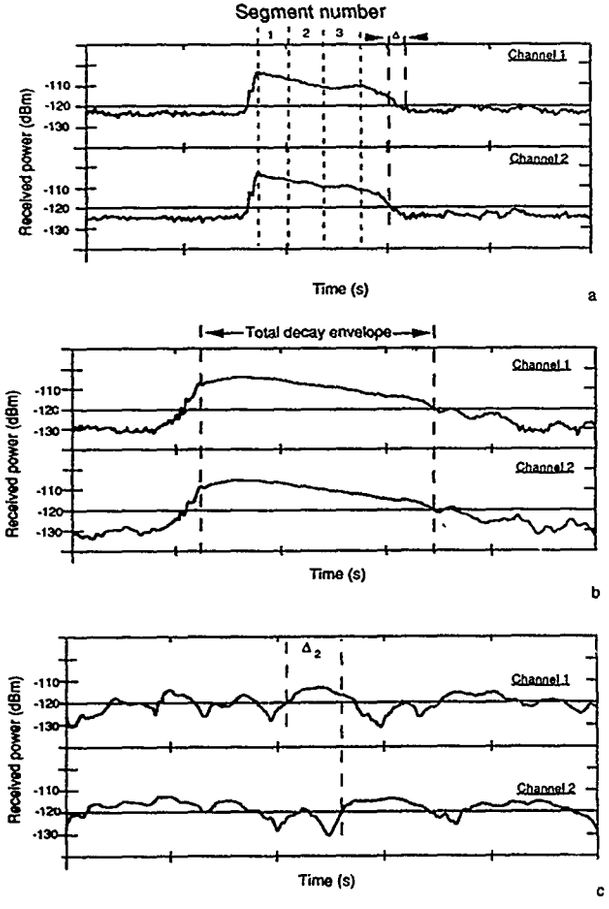


Figure 4. Signal classification:- a) Underdense signals, b) Overdense signals, and c) Not Known (NK) signals.

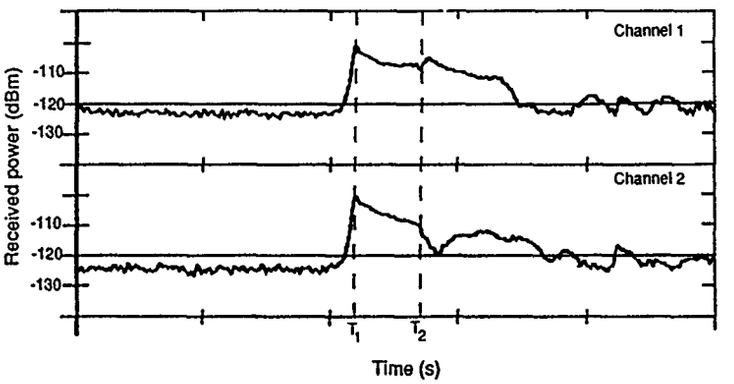


Figure 5. An example of echo overlap trails

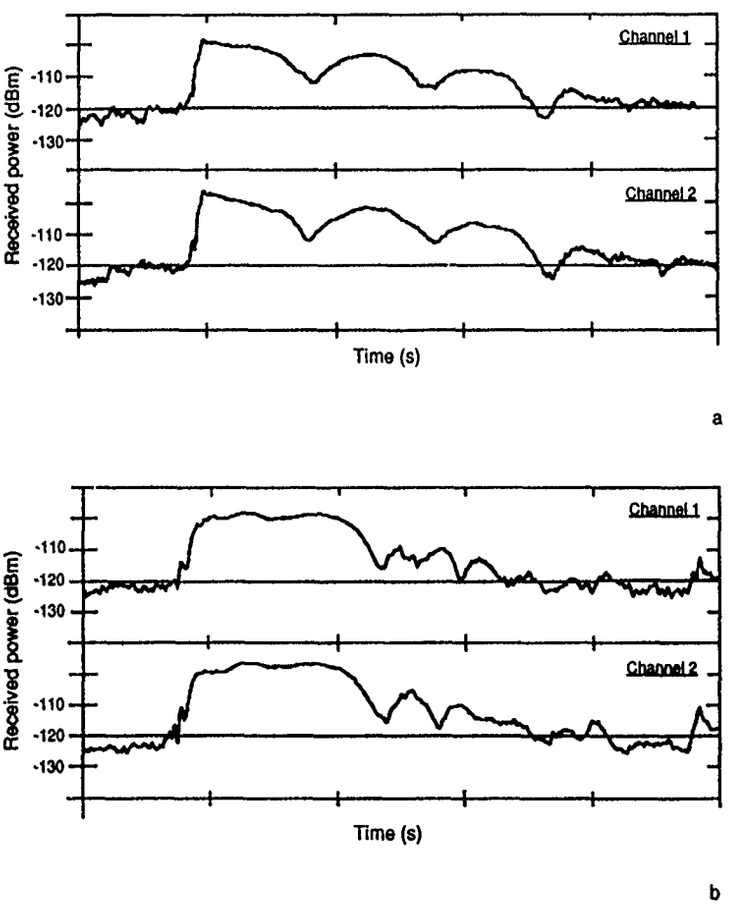
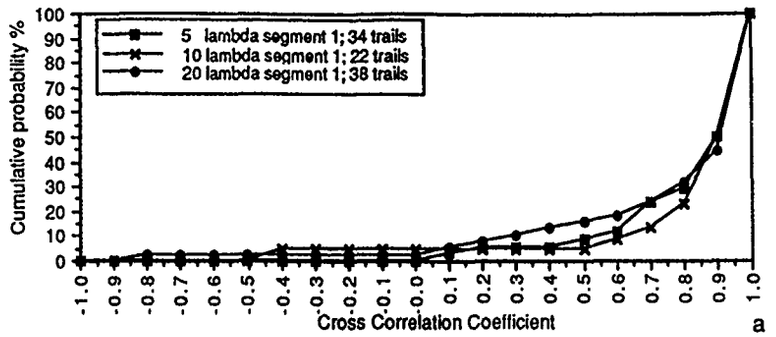
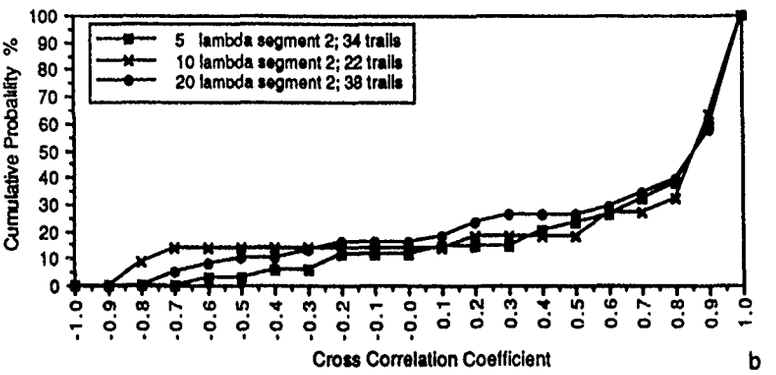


Figure 6. Deep periodic fading observed on a) underdense trails, and b) overdense trails.





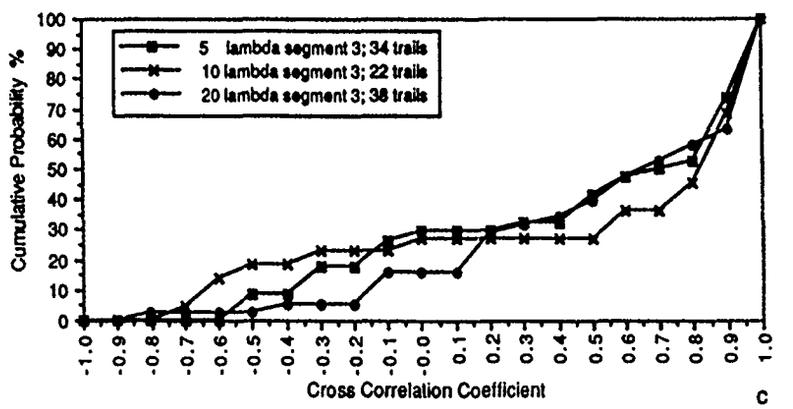
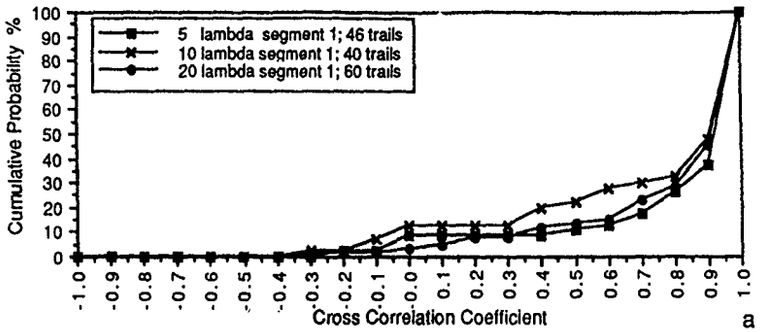
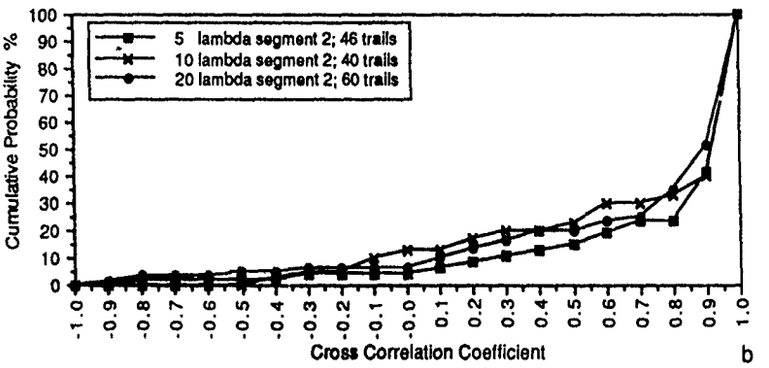


Figure 7. Cumulative correlation probabilities applicable to underdense signals of duration ≥ 0.75s received at antenna spacings of 5λ, 10λ, 20λ :- a) segment 1, b) segment 2, c) segment 3.





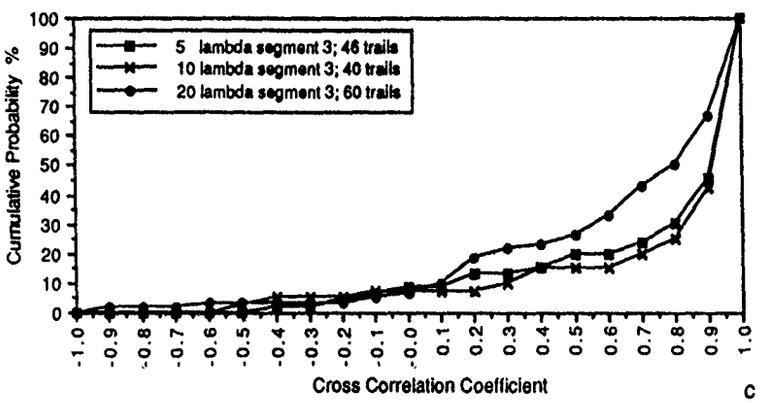
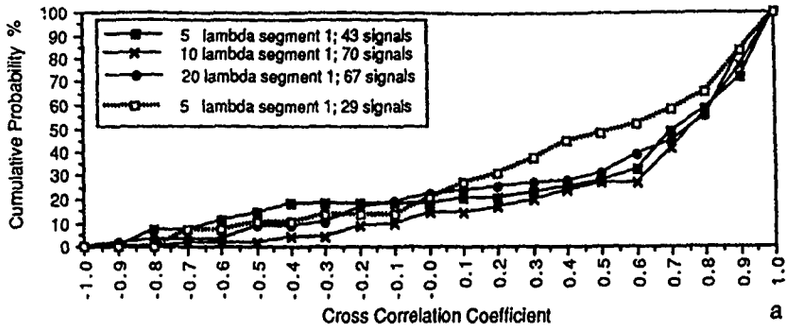
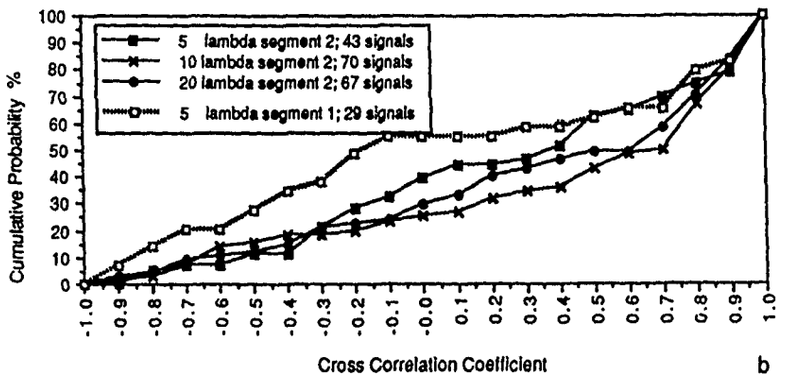


Figure 8. Cumulative correlation probabilities applicable to overdense signals of duration ≥ 0.75s received at antenna spacings of 5λ, 10λ, 20λ :-
a) segment 1, b) segment 2, c) segment 3.





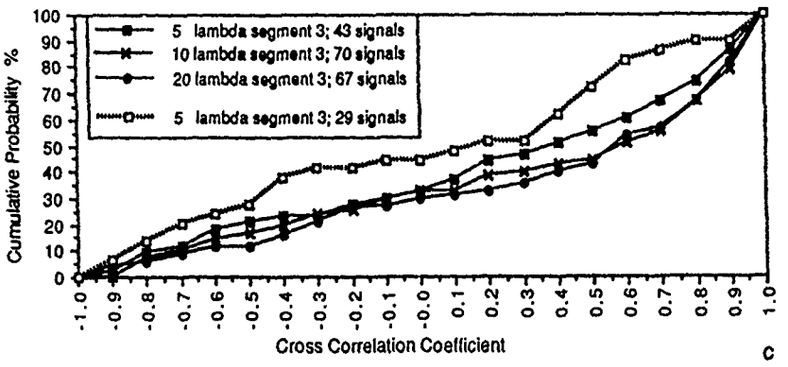


Figure 9. Cumulative correlation probabilities applicable to NK signals of duration ≥ 0.75s received at antenna spacings of 5λ, 10λ, 20λ :- a) segment 1, b) segment 2, c) segment 3. 5λ data is given for day 53 (—■—) and day 44 (—□—).

DISCUSSION

D. YAVUZ

Can you relate this work to variable rate coding (Eg. Miller & Milstein of Univ. of Californie, San Diego) method of extending useful trail duration.

AUTHOR'S REPLY

Since coding is a form of diversity - time diversity - I would espect it to extend the useful trail direction. The code rate would, however, have to be matched to the trail signal-to-noise ratio to extrac its maximum advantage.

C. GOUTELARD, FR

1 - Il ne m'apparaît pas absolument nécessaire d'utiliser des systèmes ARQ dans le cas d'utilisation de codage. En effet, le taux d'erreur varie en fonction du temps, augmentant dans la "queue" du météore. L'entrelacement ajouté au codage répartit de façon uniforme les erreurs et peut éviter les systèmes ARQ dont l'emploi s'avère lourd.

2 - Ne pensez-vous pas que d'autres dispositions des antennes, non placées systématiquement dans le plan du grand cercle, pourraient être envisagées compte tenu des directions, des lieux d'apparition des météores et de l'effet des vents de cisaillement.

AUTHOR'S REPLY

1 - I agree that interleaving is approprite technique to counteract burst errors due to fading but in the end we need ARQ to obtain retransmission of the lost data.

2 - This is a very interesting point. We have separeted our 4 antennas perpendicular to the great circle path but depending on the wind direction greater diversity gain may be obtained by reorientating the antennas out of the perpendicular plane.

M. DARNELL

1 - Can you state precisely what you mean by "equal gain" diversity ? Is it selection diversity ?

2 - May I point out that error control coding is also a form of time diversity : therefore, one might expect a similar order of improvement from this.

3 - We seem to have a similar conflict of interest here to the HF case : for receiving antenna gain to be effective, a plane wavefront is required ; for

diversity to be effective, an irregular wavefront is necessary. I believe that you have shown that diversity is more effective than an increase in transmitter power ; hence, the irregular wavefront dominates.

AUTHOR'S REPLY

1 - We have added the amplitudes of the signals on the two channels.

2 - I agree.

3 - Certainly the irregular wavefront dominates at the end the trail, although at the start of the trail the regular wavefront will probably dominate.

ANALYSE DES RESULTATS EXPERIMENTAUX DE LIAISONS METEORIQUES REALISEES POUR DIFFERENTES DISTANCES ET FREQUENCES

D. SORAIS

THOMSON-CSF/RGS
66, rue du Fossé Blanc - 92231 GENNEVILLIERS CEDEX

O. RAVARD, L. BERTEL

LABORATOIRE RADIOCOMMUNICATIONS
UNIVERSITE DE RENNES I
CAMPUS DE BEAULIEU - 35042 RENNES CEDEX
FRANCE

RESUME

Dans cet article, nous donnons les résultats expérimentaux de liaisons météoriques réalisées entre 100 et 1000 km. Une interprétation physique de ces résultats est obtenue en exploitant les distributions statistiques des paramètres du canal, ainsi que les enregistrements de la fonction de transfert du canal. La base de données constituée complète et amende utilement les informations fournies pour les modélisations classiques du canal.

ABSTRACT

This paper describes the experimental results of meteor burst communications recorded at ranges from 100 to 1000 km. A physical interpretation of these results was obtained by using a statistical distribution of the meteor burst intervals and of the transfert function measurements. The resulting data base completes and improves the information provided by the traditionnal models.

1 INTRODUCTION

Le flux permanent de météores incident sur l'atmosphère terrestre crée, vers 100 km d'altitude, des traînées ionisées permettant de réaliser des liaisons radioélectriques intermittentes. Le canal météorique ainsi défini a été étudié par de nombreux chercheurs depuis le début des années 50 [1] , [2]. Malgré cela, on ne dispose pas actuellement d'outils de prévisions à partir desquels les performances et la fiabilité d'un système de transmission quelconque pourraient être évaluées.

Pour une liaison et une période de l'année données, les modèles de prévisions permettent, au mieux, d'estimer les valeurs moyennes des paramètres caractéristiques d'une liaison intermittente. Une expérimentation de l'ordre de 80 heures étant nécessaire pour mesurer ces valeurs moyennes , il est difficile de constituer une base de données représentative de différentes configurations de liaisons (distances, antennes, périodes de l'année, fréquences, ...). De fait, la plupart des expérimentations réalisées entre stations terrestres ont été effectuées pour des distances proches de 1 000 km durant une période de l'année. Il existe une expérimentation réalisée par NES [3] pour laquelle des mesures simultanées des paramètres du canal ont été effectuées pour différentes distances.

On décrit dans cet article les résultats de l'expérimentation THEOREME (Transmission HErtzienne par Ondes REfléchies sur traînées MEtéoriques), réalisée durant l'année 1989. Pour chacune des six distances comprises entre 100 et 1 000 km, les paramètres du canal ont été mesurés durant au moins 80 heures pour des périodes de l'année distinctes. Afin d'interpréter correctement les résultats expérimentaux, il est nécessaire d'identifier, et si possible de quantifier les contributions des modes de propagation non météoriques relativement aux performances globales mesurées. Pour atteindre partiellement ce but, des expérimentations complémentaires associées à chacune des distances précédentes ont permis de mesurer le module de la fonction de transfert du canal. Un logiciel, non décrit dans cet article, permet de classer automatiquement les réponses temporelles mesurées parmi les différents modes de propagation identifiés.

Pour l'essentiel, on fournit ici, d'une part les paramètres moyens du canal, et d'autre part, une interprétation physique des résultats obtenus à l'aide des distributions statistiques et des fonctions de transfert mesurées. Ces résultats fournissent des éléments pour évaluer la faisabilité de liaisons de distances quelconques inférieures à 1 000 km réalisées à partir d'une station principale. Un effort particulier d'analyse a été effectué pour les distances inférieures à 500 km pour lesquelles les résultats disponibles dans la littérature semblent peu cohérents.

2 DESCRIPTION DES EXPERIMENTATIONS

Une liaison intermittente est principalement caractérisée par trois paramètres :

- la durée de service d_s
- la durée moyenne d'ouverture $\bar{\tau}$
- la durée moyenne de fermeture $\bar{\delta}$

Pour une période d'observation donnée, la durée de service du canal est égale au pourcentage du temps de cette période pour lequel le rapport signal à bruit reçu est supérieur au seuil de détection. La durée d'une fermeture élémentaire est définie par l'intervalle de temps séparant deux instants d'ouverture consécutifs du canal . Ces trois paramètres sont dépendants :

$$d_s = \frac{\bar{\tau}}{\bar{\tau}+\bar{\delta}}$$

Les ouvertures du canal sont mesurées par la transmission permanente d'une courte séquence codée reconnue en réception. Le banc de mesure permet donc d'obtenir les valeurs moyennes τ et δ ainsi que les fonctions de répartitions des variables aléatoires associées. Le module de la fonction de transfert du canal est mesuré avec un autre banc de mesures permettant de réaliser un enregistrement numérique de la réponse temporelle.

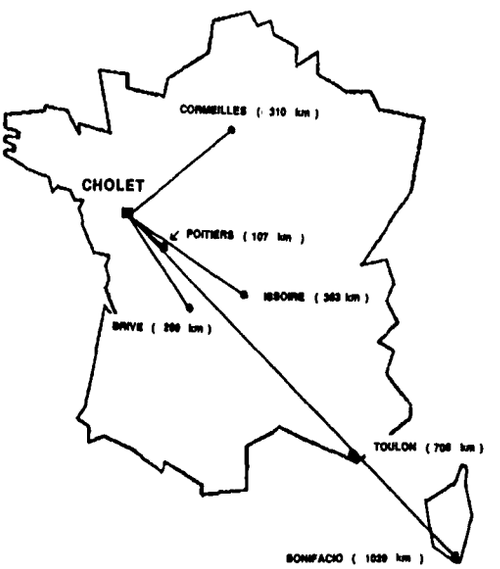
Les six liaisons expérimentales ont été réalisées sur le territoire français entre une station d'émission située à Cholet et une station de réception déplacée sur les six lieux de mesures (voir carte, figure 1 et tableau 1). On décrit ci-après les différentes expérimentations liées à ces liaisons.

2.1 *Mesures des ouvertures et des fermetures du canal*

2.1.1 Expérimentations principales

Les caractéristiques radioélectriques de ces liaisons sont présentées dans le tableau 2. Ces expérimentations ont été effectuées à une fréquence de 42 MHz en utilisant, à l'émission et à la réception, des antennes log-périodiques horizontales orientées suivant le grand cercle de la liaison (tableau 3).

Les antennes sont placées à une hauteur au-dessus du sol telle que les axes liés aux maximums des lobes principaux s'intersectent en un point situé à environ 100 km d'altitude à la verticale du point milieu de la liaison . Pour les courtes distances (sauf Poitiers), les mesures ont également été réalisées pour des hauteurs d'antennes élevées, à priori non optimales pour le canal météorique, afin de favoriser la propagation par diffusion troposphérique pour laquelle le gain aux angles d'élévations faibles est essentiel. Les valeurs calculées du gain de l'antenne log-périodique en fonction de l'angle d'élévation sont représentées figures 2 et 3 pour différentes hauteurs de l'antenne au-dessus du sol.



POSITIONS GEOGRAPHIQUES DES LIAISONS METEORIQUES
LOCATIONS OF THE METEORS LINKS CARRIED OUT

Figure 1

STATIONS	COORDONNEES GEOGRAPHIQUES	DISTANCES STATION/CHOLET
CHOLET (Emission)	47.1N - 0.9E	
POITIERS	46.6N 0.3E	107 KM
BRIVE	45.1N 1.5E	289 KM
CORMEILLES	49.1N 2E	310 KM
ISSOIRE	45.6N 3.3E	363 KM
TOULON	43.1N 6.1E	706 KM
BONIFACIO	41.4N 9.3E	1029 KM

TABLEAU 1 - CARACTERISTIQUES GEOGRAPHIQUES DES LIAISONS

Puissance émetteur	2 000 W
Pertes coaxiaux émission (Identiques en réception)	2.2 dB
Antennes émission et réception :	Gain/espace libre
Log-périodiques placées à une hauteur h au-dessus du sol	7 dBi
Facteur de bruit récepteur	3.4 dB
Rythme binaire	16 Kbit/sec
E_b/N_o minimum nécessaire	6 dB

TABLEAU 2 - CARACTERISTIQUES RADIOELECTRIQUES DES LIAISONS

STATION DE RECEPTION	HAUTEURS ANTENNE/SOL (m)	NOMBRE D'HEURES DE MESURE	DATE MEDIANE DES MESURES
POITIERS (107 km)	3	52	15-5
BRIVE (289 km)	3	23	25-5
	4	141	2-6
	11.5	68	31-5
CORMEILLES (310 km)	3	223	20-4
	4	99	7-4
	11.5	17	5-4
ISSOIRE (363 km)	4	165	15-6
	11.5	40	17-6
TOULON (706 km)	6.5	84	29-6
BONIFACIO (1029 km)	11.5	229	26-9

TABLEAU 3 - MESURES DES OUVERTURES ET DES FERMETURES EXPERIMENTATIONS PRINCIPALES

2.1.2 Expérimentations complémentaires

Une série d'expérimentations complémentaires a permis de faire des mesures à 42 MHz en polarisation verticale ainsi qu'à 70 MHz en polarisation horizontale. Les performances d'une liaison réalisée avec une antenne d'émission formée de deux dipôles horizontaux alimentés en quadrature ont également été évaluées. Les principales caractéristiques de ces expérimentations sont résumées dans le tableau 4. Les gains de l'antenne log-périodique verticale et des dipôles croisés sont respectivement représentés sur les figures 4 et 5.

2.2 *Mesure du module de la fonction de transfert du canal*

Le sensibilité de la chaîne de réception utilisée est de 4 dB meilleure que celle des expérimentations du paragraphe 2.1. Les autres caractéristiques radioélectriques sont identiques.

3 INFLUENCE DU PROTOCOLE DES EXPERIMENTATIONS SUR LES VALEURS MESUREES

Pour la mesure des ouvertures et des fermetures une interruption de la liaison supérieure à 5 ms environ sera identifiée comme une fermeture. La quasi-totalité des coupures liées aux évanouissements associés à une trainée ionisée détermineront donc des ouvertures et des fermetures de courtes durées. L'influence du protocole sur la mesure de la durée moyenne d'ouverture a été estimée en supposant que les durées d'ouverture étaient exponentiellement distribuées (figure 6).

Les coupures de liaisons supérieures à 1 ms seront détectées lors de la mesure du module de la fonction de transfert. La bande passante du filtre d'entrée de ce banc de mesures est égale à 7.5 KHz.

STATION DE RECEPTION	FREQUENCE	ANTENNES	NOMBRE D'HEURES DE MESURES	DATE MEDIANE DES MESURES
CORMEILLES (310 km)	42	Log-pério. verticales :		
		4 m/sol	22	26-4
		11.5 m/sol	1°	25-4
	70	Log-pério. horizontales :		
		3 m/sol	92	11-4
		11.5 m/sol	25	15-4
BONIFACIO (1029 km)	42	Dipôles croisés 5 m/sol ↓ Log-pério. horizontales 5 m/sol	60	21-9
		Log-pério. verticales 11.5 m/sol	18	26-9
	70	Log-pério. horizontales 11.5 m/sol	20	22-9

TABLEAU 4 - MESURES DES OUVERTURES ET DES FERMETURES EXPERIMENTATIONS COMPLEMENTAIRES

4 ANALYSE DES RESULTATS DES EXPERIMENTATIONS PRINCIPALES

On discerne deux classes homogènes de résultats respectivement associées aux courtes distances (d < 500 km) et aux moyennes distances (500 < d < 1000 km).

4.1 *Liaisons moyennes distances*

Dans un premier temps, on fournit les résultats de la liaison Cholet - Bonifacio. Chaque point du graphique de la figure 7 représente la valeur de la durée de service d_s (h, j) pour une heure h du jour j de l'expérimentation. Pour les 229 heures de mesures, le rapport du maximum au minimum horaire est égal à 17. Pour une heure donnée, les fluctuations de jour à jour peuvent être filtrées en moyennant sur les N_j jours de mesures :

$$\bar{d}_s(h) = \frac{1}{N_j} \cdot \sum_{j=1}^{N_j} d_s(h,j)$$

Les variations de $\bar{d}_s$ (h) fournissent la signature diurne sinusoïdale attendue dont le rapport des extremums est égal à 3.3 . Les modélisations du canal permettront, au mieux, de prévoir les valeurs de la durée de service $\bar{d}_s$ (h) liée à une semaine donnée de l'année. La moyenne de $\bar{d}_s$ (h) sur les heures d'un jour représente la valeur moyenne d_s de la durée de service pour la totalité de l'expérimentation.

$$\bar{d}_s = 4.1\ \%$$

Les durées moyennes d'ouverture et de fermeture du canal correspondantes ont les valeurs suivantes :

$$\bar{\tau} = 270\ \text{ms} \ ; \ \bar{\delta} = 6.3\ \text{sec}$$

Il est intéressant de comparer ces résultats avec ceux de l'expérimentation COMET [4] réalisée dans la même zone géographique que l'expérimentation THEOREME et sur une distance quasi-identique (995 km). Une partie des expérimentations COMET a été effectuée vers la mi-novembre 1967 alors que notre expérimentation a été réalisée vers la fin du mois de septembre 1989, qui est la période de l'année pour laquelle l'activité du canal météorique est maximale (météores sporadiques) . Les variations autour de ce maximum étant peu rapides, la comparaison des performances mesurées est réalisable en prenant uniquement en compte les caractéristiques radioélectriques respectives de ces deux expérimentations. On a utilisé la description et les résultats de l'expérimentation COMET fournis par BROWN [5] tels que :

$$\bar{d}_{s_{COMET}} = 14\ \% \qquad \bar{\tau}_{COMET} = 470\ \text{ms}$$

Si P_e désigne la puissance d'émission, P_{min} la puissance minimale nécessaire en réception et f la fréquence de l'expérimentation (36.6 MHz pour COMET) on admettra que la durée de service suit la loi :

$$d_s \propto \left[\frac{P}{P_{min}}\right]^{0.6} \times f^{-2.4}$$

Afin d'estimer la valeur de la durée de service THEOREME que l'on obtiendrait avec les paramètres de COMET, cette durée de service doit être mutipliée par le facteur 0.8 dû aux termes de la relation précédente. Par ailleurs, le facteur multiplicatif lié au gain du système antennaire de COMET par rapport au notre est de l'ordre de 2.1. Finalement, on obtient un facteur global égal à 1.7 :

$$7\ \% = \bar{d}_{s_{THEOREME}} \times 1.7$$

Il est possible de justifier l'écart entre la valeur extrapolée à partir de THEOREME (7 %) et la valeur effectivement mesurée (14 %) par le gain de diversité du système de réception COMET qui permettrait de réduire la perte sur la durée totale d'ouverture produite par les évanouissements associés à un certain nombre de trainées ionisées [6] .

La fonction de répartition complémentaire des intervalles de temps δ séparant les instants d'ouvertures du canal (c'est-à-dire les fermetures), doit être exponentielle si ces instants ont une distribution poissonnienne. Pour un jour donné de l'expérimentation on constate en effet que la droite attendue est correctement établie pour δ supérieur à une seconde environ (figure 8). On obtient 40 % (x_p %) d'évènements liés à des instants d'arrivée poissonniens et 60 % d'évènements qui, à priori, peuvent être attribués, d'une part, à des courtes fermetures produites par des évanouissements et, d'autre part, à d'éventuels modes de propagation non météoriques.La fermeture moyenne globale $\bar{\delta}$ mesurée est évidemment plus faible que la fermeture moyenne $\bar{\delta}_p$ des distributions exponentielles. Pour l'ensemble de cette expérimentation, on obtient :

$$\bar{\delta} = 6.5\ \text{sec} \ ; \ \bar{\delta}_p = 16\ \text{sec} \quad \text{et} \quad x_p = 45\ \%$$

L'analyse des réponses temporelles enregistrées pour la même configuration de liaison, permet d'interpréter ces résultats. Les évanouissements des réponses temporelles du canal météorique produiront un ensemble d'ouvertures et de fermetures lorsque le niveau de ces réponses fluctue autour du seuil de détection (figures 9 et 10). Il existe également des réponses temporelles liées à un mode de propagation par diffusion ionosphérique (figure 11).

Pour un enregistrement de 110 minutes de la fonction de transfert du canal, la fonction de répartition complémentaire des fermetures a été construite pour différents-traitements des données : (figure 12).

- La fonction de répartition pour les données brutes est telle que 50 % des évènements sont poissonniens.

- La suppression des ouvertures liées à une diffusion ionosphérique produit une légère remontée de la droite de Poisson (57 % d'évènements poissonniens).

- On a enfin supprimé les fermetures dues aux évanouissements des réponses du canal météorique. Les réponses liées à la diffusion ionosphérique étant toujours supprimées, on obtient une droite proche de celle due à un phénomène totalement poissonnien.

Il est donc certain que les évanouissements réduisent simultanément les durées moyennes d'ouvertures et de fermetures par rapport aux valeurs prévues par les modélisations classiques du canal météorique.

Les distributions des ouvertures du canal sont beaucoup plus difficiles à interpréter. Pour une journée de notre expérimentation, la fonction de répartition complémentaire des ouvertures représentée figure 13 ne peut être que très approximativement approchée par une distribution exponentielle classiquement utilisée dans les simulations du canal météorique. En particulier la probabilité d'obtenir des durées d'ouvertures élevées est notablement plus forte que celle fournie par une loi exponentielle. WEITZEN [7] a donné des interprétations physiques intéressantes sur les distributions observées.

L'analyse des résultats de l'expérimentation Cholet - Toulon, conduit à des conclusions identiques à celles de la liaison Cholet - Bonifacio précédemment examinée.

4.2 *Liaisons courtes distances*

On examinera tout d'abord les mesures des stations de Brive, Cormeilles et Issoire pour lesquelles les interprétations physiques des résultats sont semblables. Les antennes étant placées à 3 m au -dessus du sol, on observe, pour Cormeilles, une signature diurne de la durée de service caractéristique du canal météorique (figure 14) qui n'est plus obtenue lorsqu'on élève les antennes au-dessus du sol.

La valeur du pourcentage x_p d'évènements poissonniens diminue lorsque l'on élève les hauteurs des antennes corrélativement avec l'accroissement de la durée de service d_s . Ce phénomène est clairement représenté figure 15 par des résultats partiels des distributions de fermeture de la liaison Cholet → Cormeilles. La représentation de d_s en fonction de x_p , pour les trois hauteurs de l'expérimentation (3, 4 et 11,5 m), permet de filtrer les modes de propagation présents (figure 16). Le déplacement des points dans le plan (x_p , d_s) est dû au mode de propagation par diffusion troposphérique qui est favorisé par l'élévation des aériens. Les gains sur le bilan de liaison par diffusion troposphérique par rapport à la configuration des antennes placées à 3 m, sont respectivement de 6 et 22 dB pour les hauteurs de 4 et 11,5 m. On montre, par simulation, que l'influence des hauteurs des aériens sur la durée de service du canal météorique, n'est pas en rapport avec les évolutions observées. L'accroissement de la hauteur électrique des antennes produite par l'augmentation de la fréquence (40 → 70 MHz) accroît la valeur de la composante troposphérique du canal, contrairement à la composante météorique qui décroît fortement .

L'interprétation des résultats de la station de Brive est semblable à celle de Cormeilles. La composante du mode de propagation par diffusion troposphérique devient très faible pour Issoire (363 km).

Pour la station de Poitiers (107 km), la diffusion troposphérique est dominante même pour des antennes placées à 3 m de hauteur. On peut obtenir des durées de service supérieures à 50 % en plaçant les antennes à 11,5 m.

Une réponse temporelle due à la diffusion troposphérique (figure 17) montre les ouvertures et les fermetures induites par le seuil de détection en présence d'un évanouissement. Compte tenu de la vitesse de cet évanouissement, on obtient des durées moyennes d'ouverture comparables à celles du canal météorique.

4.3 *Analyse globale des expérimentations principales*

Le tableau 5 résume les principaux résultats de ces expérimentations.

Pour les courtes distances la contribution de la propagation par diffusion troposphérique est négligeable pour les liaisons de Brive et Cormeilles réalisées avec des antennes placées à 3 m au-dessus du sol, ainsi que pour Issoire, avec des antennes situées à 4 m de hauteur. Pour les autres configurations courtes distances, la propagation troposphérique, toujours présente dans nos expérimentations, permet d'obtenir de meilleures performances que celles dues au seul canal météorique.

L'évaluation de la pente δ_p de la droite de la distribution exponentielle montre que l'intervalle moyen entre météores se maintient à 45 sec environ lorsque la distance décroît de 310 à 100 km (expérimentations réalisées durant les mois d'avril et mai 1989).

Afin d'accroître la fiabilité des liaisons courtes distances, il est donc possible d'exploiter simultanément le canal météorique et le canal de propagation par diffusion troposphérique en utilisant des mats tactiques de faibles hauteurs (≅ 10 m). Les valeurs minimales et maximales de la durée de service moyenne du canal météorique sur une année, pour toutes distances comprises entre 300 et 1 000 km, sont à priori représentées par les deux configurations suivantes :

Cormeilles 3 m/sol
Avril 1989 - 310 km → $\bar{d}_s \cong 0.9$ %

Bonifacio 11,5 m/sol
Septembre 1989 - 1029 km → $\bar{d}_s \cong 4.1$ %

Les mois d'avril et de septembre correspondent en effet aux extrémums de l'activité météorique annuelle (météores sporadiques). Compte tenu du rapport attendu de ces fluctuations (environ 3), on doit conclure que la décroissance de d_s avec la distance est notablement plus faible que celle prévue par les modèles classiques de prévisions [3], [8]. La valeur moyenne de la durée d'ouverture du canal météorique est comprise entre 250 et 300 ms quelle que soit la configuration de liaison. Les traînées de longues durées qui déterminent un évanouissement du signal reçu produisent de courtes ouvertures qui diminuent fortement la durée moyenne d'ouverture. Le pourcentage d'évènements non poissonniens (100 - x_p) %, déterminés partiellement par ces évanouissements est plus élevé aux moyennes distances (~ 1000 km), qu'aux courtes distances (~ 300 km) (tableau 5). Ceci est cohérent avec le fait que la probabilité d'évanouissement est d'autant plus forte, que les durées d'ouvertures sont importantes.

5 ANALYSE DES RESULTATS DES EXPERIMENTATIONS COMPLEMENTAIRES

Le tableau 6 fournit la durée de service d_s, la durée d'ouverture moyenne et la durée de fermeture moyenne pour les différentes configurations retenues.

L'utilisation des antennes log-périodiques verticales placées à 4 mètres au-dessus du sol (Cormeilles 310 km) détermine une faible valeur de la durée de service (d_s = 0,3 %). Ceci peut être imputé en partie à la réduction du gain due à l'incidence pseudo-brewstérienne ainsi qu'à la décroissance rapide de ce gain pour des angles d'élévations supérieurs à 50° (Figures 4).

Les pertes de polarisation sans doute notables pour cette distance [S] , sont difficilement séparables des pertes dues aux gains des antennes. La propagation par diffusion troposphérique permet cependant d'obtenir de bonnes performances lorsque les antennes sont élevées à 11,5 m au-dessus du sol. Conformément aux prévisions théoriques, la polarisation verticale fournit des performances comparables à celles de la polarisation horizontale pour Bonifacio (1023 km).

Pour cette dernière distance, une expérimentation relativement longue (60 heures), a été réalisée avec, en émission, une antenne formée de deux dipôles horizontaux croisés alimentés en quadrature et, en réception, une antenne log-périodique horizontale ($\bar{d}_s \cong 1.9$ %). Une diffusion radioélectrique omnidirectionnelle en azimut peut être obtenue avec cette station d'émission.

On notera enfin que pour une fréquence de 70 MHz, les performances demeurent excellentes pour les courtes distances (composante troposphérique) et se dégradent fortement pour les distances moyennes (composante météorique).

6 CONCLUSION

Les évanouissements des réponses temporelles liées aux traînées ionisées agissent fortement sur les paramètres moyens qui caractérisent le canal météorique ainsi que sur les distributions statistiques associées à ces paramètres. Bien que les évènements dus aux météores surdense soient relativement peu nombreux, ils interviennent notablement sur les distributions des ouvertures et des fermetures. Ces différents phénomènes,non correctement pris en compte dans les modélisations classiques, déterminent les écarts constatés entre les mesures et les prévisions. Les modélisations sont cependant utiles pour définir les aériens d'un système de transmissions ainsi que certaines variations paramétriques des performances du canal météorique.

Actuellement les résultats fournis par ces modèles doivent être nécessairement corrigés et complétés par les informations expérimentales disponibles.

7 REMERCIEMENTS

Cette étude a été réalisée grâce à un contrat de la Délégation Générale pour l'Armement - Service Technique de l'Electronique et de l'Informatique (D.G.A. - S.T.E.I.).

Les auteurs remercient Philippe BALLEUX pour sa collaboration active dans l'élaboration des logiciels, ainsi que la section "Radiocommunications" du Centre d'Electronique de l'Armement (CELAR), qui a réalisé une partie des expérimentations.

STATIONS DE RECEPTION	HAUTEURS ANTENNES (m)	d_s %	$\overline{\tau}$ (ms)	$\overline{\delta}$ (sec)	$\overline{\delta p}$ (sec)	xp %
POITIERS	3	11.8	260	2	46	4.2
BRIVE	3	0.9	263	28	48	59.7
	4	1.7	209	12	39	32.8
	11.5	22.1	225	0.8	25	1.9
CORMEILLES	3	0.9	253	29	45	67
	4	1.9	219	11	45	28
	11.5	22.2	186	0.7	10	2.4
ISSOIRE	4	1.8	248	14	25	57
	11.5	1.3	211	16	52	38
TOULON	6.5	3.9	298	7.4	16	48
BONIFACIO	11.5	4.1	280	6 5	16	45

TABLEAU 5 - SYNTHESE DES RESULTATS DES EXPERIMENTATIONS PRINCIPALES

STATIONS DE RECEPTION	f MHz	CONFIGURATION	d_s %	$\overline{\tau}$ (ms)	$\overline{\delta}$ (sec)
CORMEILLES (310 km)	42	LPV 4 m/sol	0.3	176	56
		LPV 11.5 m/sol	3 4	143	4 1
	70	LPH 3 m/sol	2 2	190	8.4
		LPH 11.5 m/sol	19.7	182	0 7
BONIFACIO (1029 km)	42	DIC 5 m/sol LPH 5 m/sol	1.9	240	12.7
		LPV 11.5 m/sol	3.7	225	5 8
	70	LPH 11.5 m/sol	0 8	190	23

TABLEAU 6 - SYNTHESE DES RESULTATS DES EXPERIMENTATIONS COMPLEMENTAIRES

8 BIBLIOGRAPHIE

[1] CANNON - REED
"The evolution of meteors communications systems". J. of the Institution of Electronic and Radio Engineers. Vol. 57, n° 3, may-june 1987.

[2] WEITZEN - RALSTON
"Meteors scatter : an overview". IEEE Trans. on Antennas and Propagation. Vol. 36, n° 12, dec. 1988.

[3] NES
"Dimensioning technique for meteor-burst communication systems". IEE Proceedings. Vol. 132, n° 6, oct. 1985.

[4] BROWN - WILLIAMS
"The performance of meteor-burst communications at different frequencies". AGARD CP 244 : "Aspects of electromagnetics waves scattering in radio communications". Cambridge UK, sept. 1978.

[5] BROWN
"A physical meteor-burst propagation model and some signifiant results for communications system design". IEE J. on selected areas in communications. Vol. SAC-3 n° 5, sept. 1985.

[6] BARTHOLOME - VOGT
"COMET - A new meteor-burst system incorparating ARQ and diversity reception". IEEE Trans. on Communication Technology. Vol. COM-16, april 1968.

[7] WEITZEN - HIBSHOOSH - SCHILLING
"Some observations on the distributions of amplitude and duration of under dense meteor and its application to design of meteor scatter protocols". MILCOM 88 - San Diego, California.

[8] WEITZEN
"Communicating via meteor burst at short ranges". IEEE Trans on Communications. N° 11, nov. 1987.

[9] WEITZEN
"Effects of polarization coupling loss mechanism on design of meteor scatter antennas for short and long range communication". Radio Sciences. Vol. 24, n° 4, july-august 1989.

[10] CANNON - DICKSON
"BLOSSOM A. An air to ground meteor burst communications system". MILCOM, oct. 1986, Monterey, California.

[11] FORSYTH - VOGAN - HANSEN - HINES
"The principles of JANET - A meteor-burst communication system". Proc. IRE, Vol. 45, dec. 1957.

[12] SINNOT - PLACE - THOMSON - SUNKENBERG - KILLIAN - BENHAM
"Meteor burst communication with a buried antenna". MILCOM 85, Boston

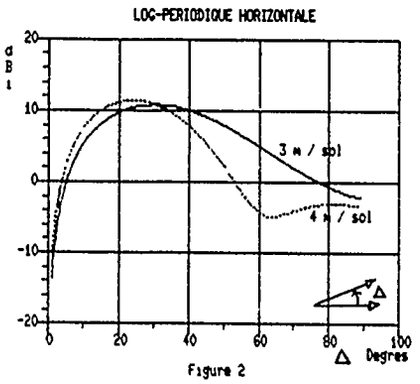


Figure 2

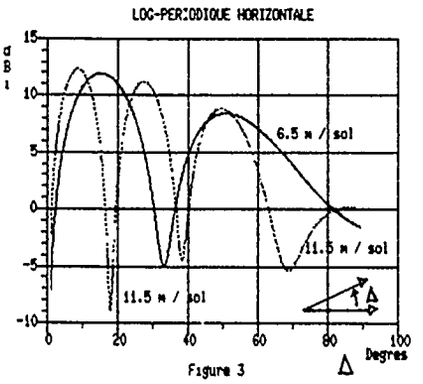


Figure 3

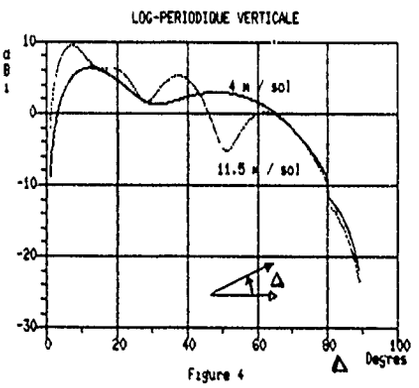


Figure 4

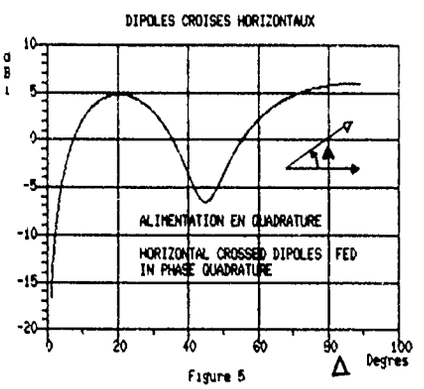


Figure 5

Directivités d'antennes placées au-dessus du sol.
Fréquence : 42 MHz - ε_r=12 - σ=0.003 S/m

Directivities of antennas above ground.

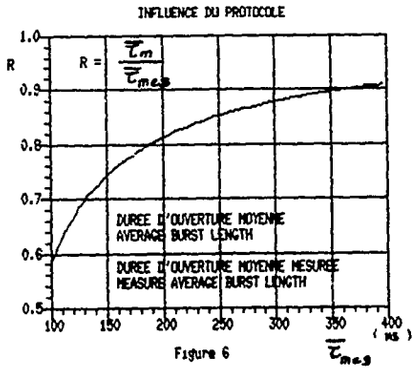
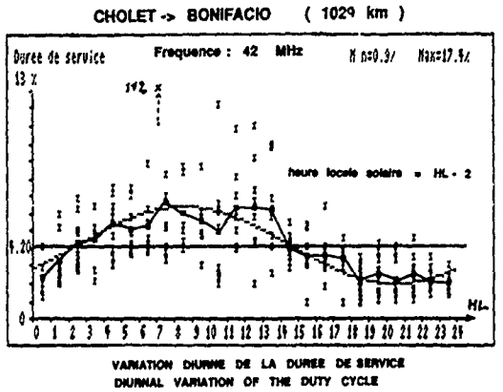
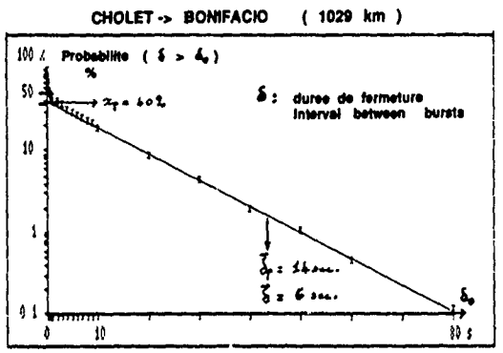


Figure 6



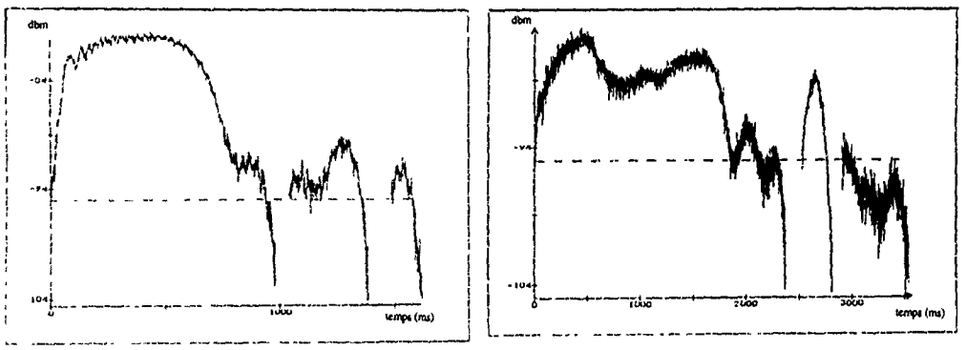
VARIATION DIURNE DE LA DUREE DE SERVICE
DIURNAL VARIATION OF THE DUTY CYCLE

Figure 7



DISTRIBUTION DES DUREES DE FERMETURES
BURST INTERVALS DISTRIBUTION

Figure 8



FIGURES 9 et 10

Exemples de traînées météoriques surdenses.

Examples of overdenses meteors trails

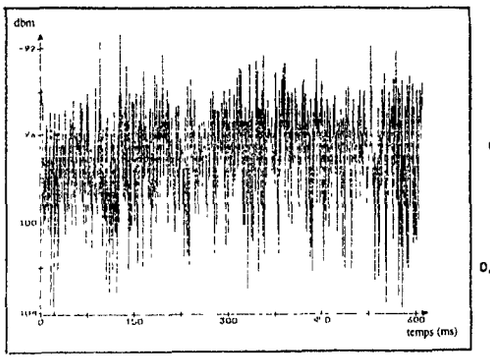
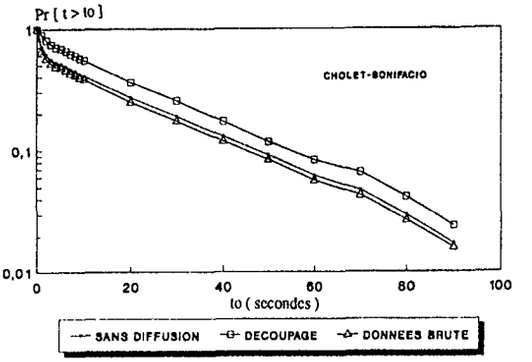
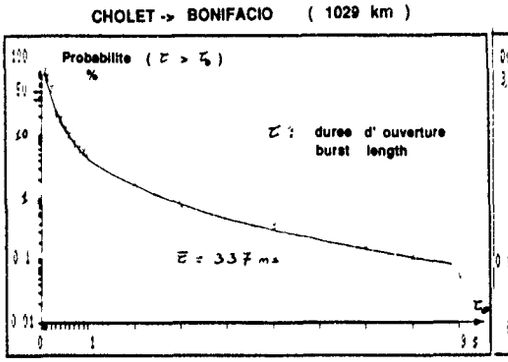


FIGURE 11
Exemple de diffusion ionosphénque
Example of ionospheric scattering



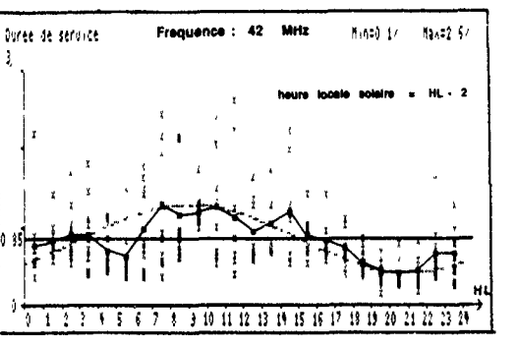
DISTRIBUTION DES DUREES DE FERMETURES
BURST INTERVALS DISTRIBUTION

Figure 12



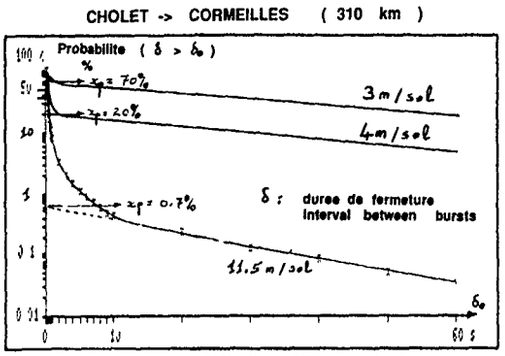
DISTRIBUTION DES DUREES D' OUVERTURE
BURST LENGTHS DISTRIBUTION

Figure 13



VARIATION DIURNE DE LA DUREE DE SERVICE
DIURNAL VARIATION OF THE DUTY CYCLE

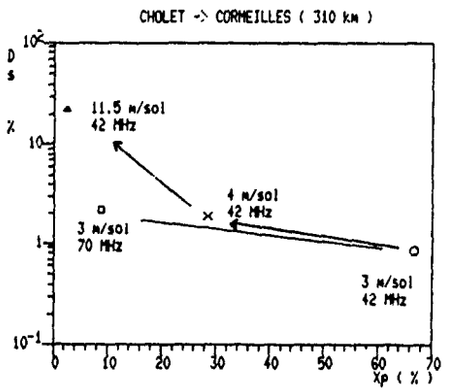
Figure 14



VARIATIONS DES DUREES DE FERMETURE EN FONCTION DE LA HAUTEUR D' ANTENNE AU-DESSUS DU SOL

DISTRIBUTIONS VARIATIONS OF BURST INTERVALS AS A FUNCTION OF ANTENNA HEIGHT ABOVE GROUND

Figure 15



Representation de la durée de service ds en fonction du nombre d'evenements poissonniens xp pour differentes configurations

Representation of the duty cycle ds as function of the percentage xp of poisson events for various configuration

Figure 16

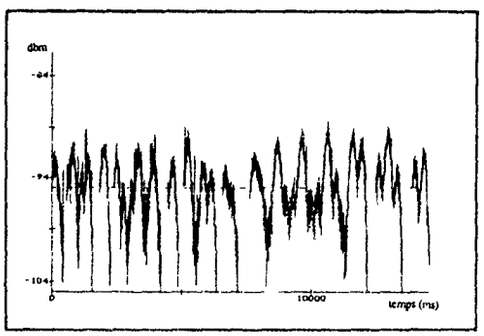


FIGURE 17

Exemple de propagation troposphérique

Example of tropospheric propagation

DISCUSSION

Paul S. CANNON

Has your modelling, and have your comparisons between data, taken at different times of the year, allowed for the diurnal and annual variation in noise level ?

AUTHOR'S REPLY

Le bruit externe n'a pas été enregistré durant l'expérimentation. Pour les différents sites de réception nous avons cependant vérifié que le bruit industriel n'était pas dominant. Les variations diurnes et annuelles du bruit galactique peuvent effectivement biaiser légèrement les résultats obtenus. La modélisation du bruit galactique reçu par l'antenne de réception n'a pas été réalisée.

M. DARNELL

I should like to confirm your results for short paths. We have been making measurements over a 300 km path for 1 year at a frequency of 47MHz. We have found that, for all receiving antenna heights, troposcatter is dominant. From a practical communications viewpoint, at short ranges, troposcatter would seem to be more useful than meteor-burst.

C. GOUTELARD, FR

Les propagations par traînées météoriques et par diffusion troposphériques sont facilement distinguées dans les expériences menées avec les radars par rétrodiffusion. En effet, les échos des traînées sont parfaitement localisés alors que les phénomènes de diffusion occupent des espaces notablement plus étendus. On peut donc, dans ce cas, mener une étude simple. En transmission on ne peut observer ces phénomènes et l'étude des distributions proposée ici est une solution intéressante.

ETUDES SPATIO-TEMPORELLE DES TRAINEES METEORIQUES PAR RETRODIFFUSION

Space and time analysis of meteor trails using VHF backscatter

G. COURJAULT, B. PIEROT
Société NARDEUX
37600 LOCHES, FRANCE

C. GOUTELARD
Laboratoire d'Etudes des Transmissions Ionosphériques
Université de Paris-Sud
Avenue de la Division Leclerc
94230 CACHAN, FRANCE

I - INTRODUCTION

Les communications utilisant les traînées ionisées créées par les météorites suscitent un intérêt permanent depuis les années 1950 durant lesquelles un premier système a été proposé [1] [2]. D'autres expérimentations ont été effectuées dans les décénnies 1960 - 1970. - COMET, METEORFAX - qui devaient fournir des résultats de grand intérêt [3] [4] et des systèmes opérationnels comme SNOTEL, AMBS, FEBA ont été implantés aux Etats-Unis d'Amérique. En France l'étude THEOREME [5] et l'étude présentée dans ces pages ont été suscitées par l'administration militaire française.

En effet, bien que des systèmes aient été conçus, il demeure des problèmes non résolus autour desquels de nombreuses études ont été effectuées et sont toujours en développement. Deux raisons principales semblent finaliser ces études :

-L'intérêt opérationnel de ces systèmes qui permettent d'établir des communications au delà de l'horizon avec de faibles coûts et opérationnels même en cas de conflit nucléaire.

-La seconde raison, plus fondamentale, tient aux études qui peuvent être effectuées sur les problèmes d'ionisation par les météores et la physique de la haute atmosphère où les mouvements des particules neutres peuvent être observés.

Les études effectuées ont d'abord porté sur la puissance transmise, la durée des signaux et les lois d'apparition des traînées [6] [7] [8] puis des analyses plus fines ont été effectuées pour caractériser les composantes aléatoires des réponses [9] et l'influence des fréquences utilisées [4]. Des modélisations, basées souvent sur des études antérieures, ont été proposées [10] [11] [12] rendant compte des phénomènes principaux. Récemment la prise en compte des effets de dépolarisation des ondes [13] [14] a été faite en vue, notamment, d'accroître l'efficacité des systèmes. Avec le même objectif l'évolution des traînées est étudiée dans l'optique de réalisations de systèmes adaptatifs [15] [16] [17].

Des études effectuées avec des radars dans la gamme VHF ont permis d'effectuer des études de vents de particules neutres dans la partie supérieure de la mésosphère [18] [19] [20] [21].

L'étude décrite dans cet article concerne une expérimentation effectuée avec un radar utilisant des impulsions longues codées dont le principe a été utilisé dans les sondages ionosphériques en rétrodiffusion [22].

Le système, fonctionnant en détection cohérente, permet d'étudier l'effet doppler et de remonter une part du mécanisme de formation de la traînée.

Le système réalisé pour cette expérience est décrit et ses performances sont précisées. Un ensemble de mesures sélectionnées sur plusieurs mois, effectué à deux fréquences différentes, avec les trois configurations de polarisation possibles, sert de base de données à une étude statistique. Elle confirme des résultats connus sur les variations journalières, saisonnières, de l'occurence des échos, mais elle fait également apparaître des résultats nouveaux sur l'effet de polarisation, la durée des échos et l'occurence de leur apparition.

Une étude plus fine de la forme des échos obtenus permet de donner une interprétation en bon accord avec une théorie proposée.

Le grand nombre de mesures traitées donne aux résultats une validité des plus intéressantes.

II - SYSTEME EXPERIMENTAL - PERFORMANCES

II.1. PRINCIPE

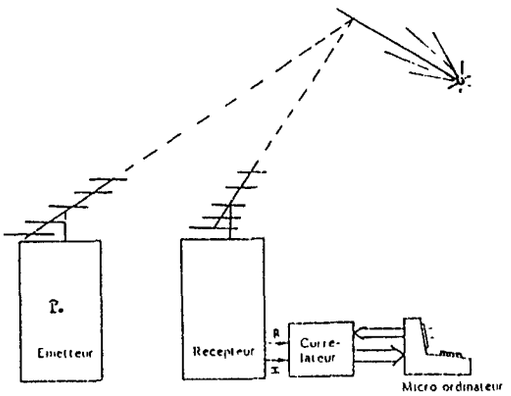


Figure 1 Système radar VHF

Le système expérimental est représenté sur la figure 1.

Si Po est la puissance disponible à l'émission, la puissance Pr, reçue est donnée par la relation

$$Pr = \frac{Po\, Gt\, Gr\, \lambda^2\, \sigma}{(4\pi)^3\, Ro} \qquad (1)$$

où Gt, Gr : gain des antennes à l'émission et la réception

σ : Surface radar équivalente

Ro : Distance de la traînée au système

λ : Longueur d'onde

Soit Tc la durée du signal émis. L'énergie reçue par mesure sera alors Wr = Pr Tc.

Le rapport signal sur bruit après démodulation est directement proportionnel à cette énergie.

$$\frac{S}{N} = \frac{Wr}{No}$$ avec No : Densité spectrale du bruit

a) Dans le cas d'un fonctionnement monostatique, les deux aériens peuvent être alors groupés (Gt = 2 Go = Gr) et être utilisés alternativement pour l'émission puis pour la réception.

Dans ce cas Tc = Tm < 1/2 Tm où Tm est la durée du cycle et l'énergie reçue par mesure Wrm vaut :

$$Wrm = \alpha\, \frac{Po\, 4\, Go^2\, \lambda^2\, \sigma}{(4\pi)^3\, Ro^4}\, Tm\, \underbrace{Ac^2/F}_{*} \qquad (2)$$

* Terme correctif tenant compte des pertes Ac dans le couplage des antennes ainsi que du facteur de bruit F du module Tr.

b) Dans le cas bistatique, l'émission est continue. Tc = Tm. Les aériens sont dédiés à l'émision où à la réception.

$$Wrb = \frac{Go^2\, \lambda^2\, \sigma}{(4\pi)^3\, Ro^4}\, Tm \qquad (3)$$

La comparaison de ces deux cas nous conduit à considérer le rapport

$$r = \frac{Wrb}{Wrm}$$

$$r = \frac{F}{4\,\alpha\, Ac^2} \# \frac{1}{2\,\alpha} \quad (F = 1{,}5\ dB,\ Ac = 1\ dB)\ (4)$$

La solution bistatique offre un bilan énergétique toujours plus favorable ($\alpha < 1/2$). De plus elle a l'avantage d'une sensibilité indépendante de la distance d'observation due à la fenêtre de réception.

C'est donc cette solution qui a été retenue.

II . 2 LE SIGNAL

On a choisi d'effectuer une modulation de phase bipolaire par un code à enchainement dont la fonction de corrélation est à pointe unique. Cette technique est parfaitement adaptée à l'émission continue et permet d'obtenir le meilleur bilan énergétique en équipant le récepteur d'un démodulateur synchrone et d'un corrélateur numérique à deux voies (réelle et imaginaire).

Les codes utilisés sont caractérisés par :

La durée d'un moment du code : Tr

Le nombre de moments : $Nr = 2^n - 1$

$n \in \{4, 5, 6, 7, 8\}$

La durée de la séquence : Tc = Nr.Tr.

Le gain de corrélation G : $G^2 = 2^n - 1$ à condition que la bande passante du récepteur soit adaptée au code (2B = 1/Tr).

II.3. LIMITE DU SYSTEME BISTATIQUE

Le couplage entre l'émetteur et le récepteur introduit deux principales limitations au système.

a) Sur la sensibilité
b) Sur la dynamique

a)Le signal E à la sortie du démodulateur synchrone est : E(t) = B(t) + b(t) + s(t).

où
B : Signal brouilleur provenant directement de l' émetteur.
b : Bruit supposé gaussien d'origine galactique, industrielle et électronique.
c : Signal utile (écho).

De la compression (par le correlateur) des différents signaux, résulte sur chacune des deux voies (figure 2).

- Pour B(t) : Un signal périodique A(t) à pointe unique de gain de traitement G^2
- Pour b(t) : Un signal gaussien (Gg(t)
- Pour s(t) : Un signal périodique y(t) à pointe unique de gain de traitement G^2.

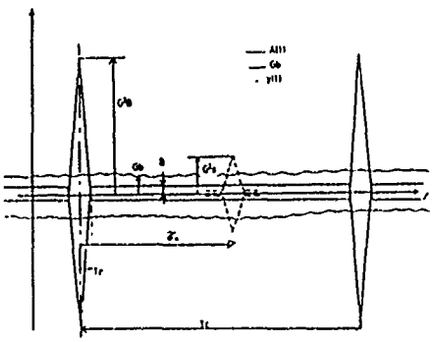


Figure 2

On prend alors le critère de détection suivant:G s > B + (Gb + 10 dB) où s, B et Gb représentent les niveaux indiqués sur la figure 2.

Le seuil minimum de détection vaut alors:

$$Smin = \frac{B}{G^2} + \frac{3,16}{G} \qquad (6)\ \alpha = 3,16\ G$$

-Si le brouillage est fort, B >> α b, la sensibilité du système est dégradée.

-Sinon, B << α b, la sensibilité du récepteur ne dépendra que du bruit ambiant.

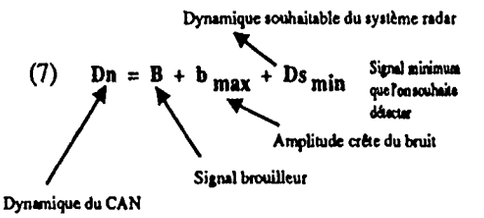
Ces performances sont résumées dans le Tableau 1 et les courbes de la Figure 3.

CODE DE LONGUEUR n	15 4	31 5	63 6	127 7	255 8
B > α b s min (dB)	B-23,5	B-30	B-36	b-42	B-48
α (dB)	22	25	28	31	34
B < α b s min (dB)	b-1,8	b-5	b-8	b-11	b-14

Tableau 1

b) Compte tenu de la qualité souhaitée pour les mesures, aucun des étages ne doit être susceptible d'être saturé. La dynamique du système est limitée par la partie numérique.

L'énumération des différents signaux se présentant à l'entrée du convertisseur analogique numérique conduit à l'équation.



$$\text{D'après (6) : } s\ min = \frac{B}{G^2} + \frac{3,16}{G}$$

On prendra b max = b + 10 dB = 3,16 b.

D'autre part, si on veut utiliser le corrélateur au niveau théorique de ses capacités, il faut minimiser le bruit de quantification issu de la conversion analogique numérique. s min est alors quantifié sur 3 bits. Les courbes de la figure 4 donnent la dynamique du système en fonction du niveau B. (Pour Po = 1 KW, Dn = 512, B < α b).

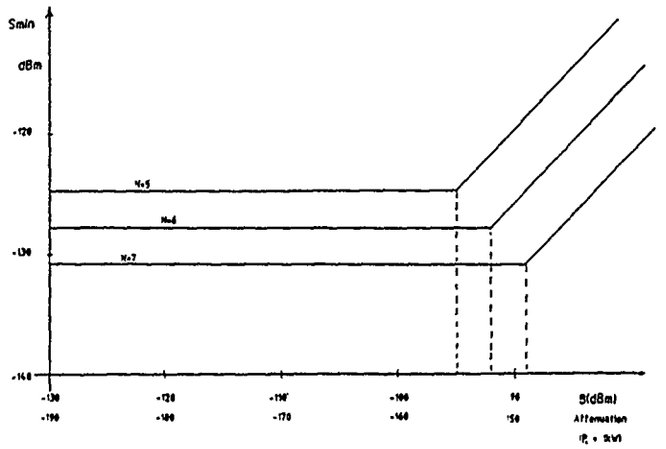


Figure 3

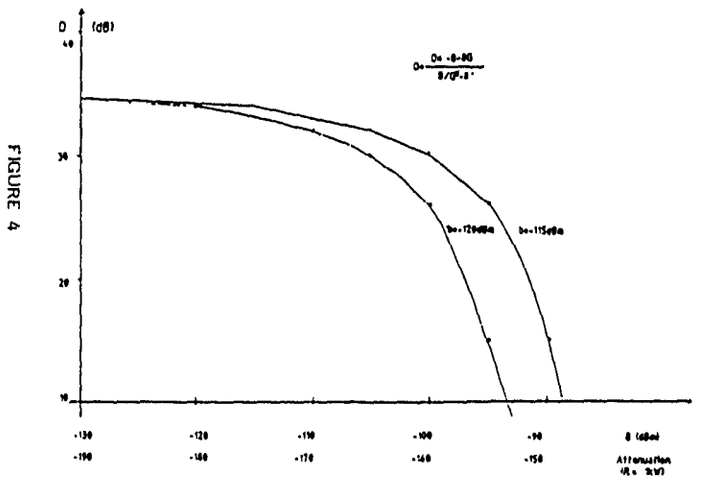


Figure 4

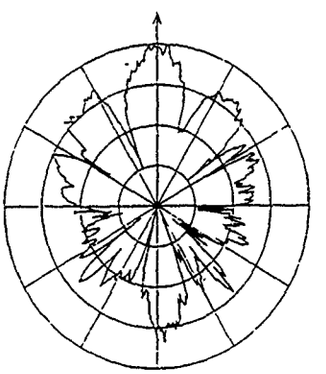


Figure 5(a) 9/03/89 Mesure en élévation angle 5°

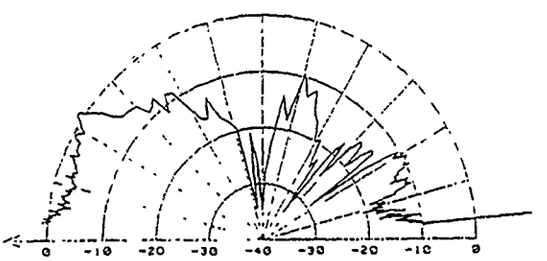


Figure 5(b) 12/07/89 Mesure dans l'axe de l'antenne: $\varphi = 0$

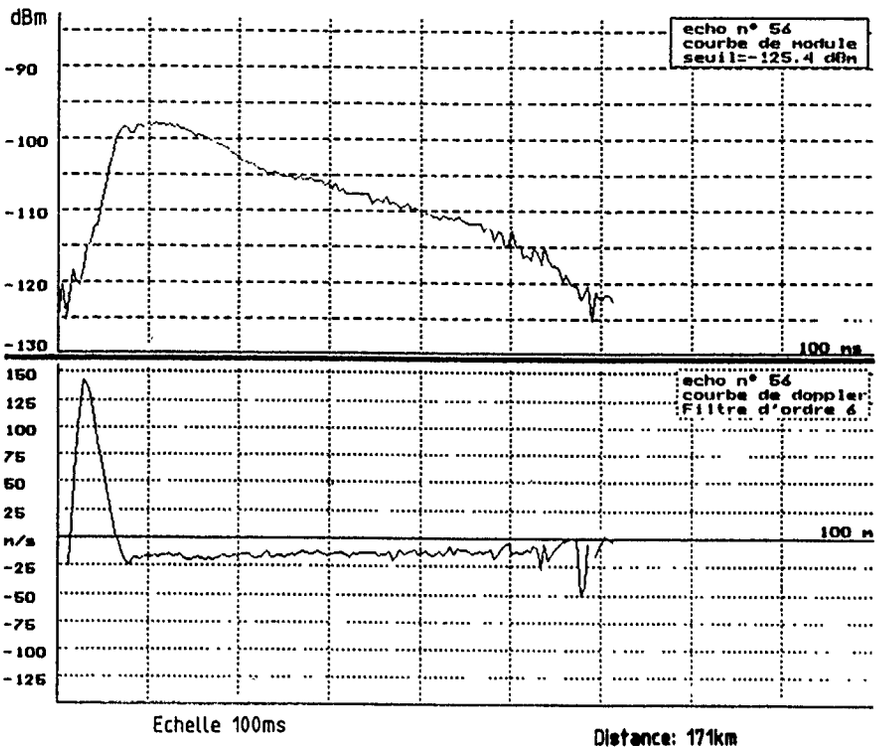


Figure 6(a)

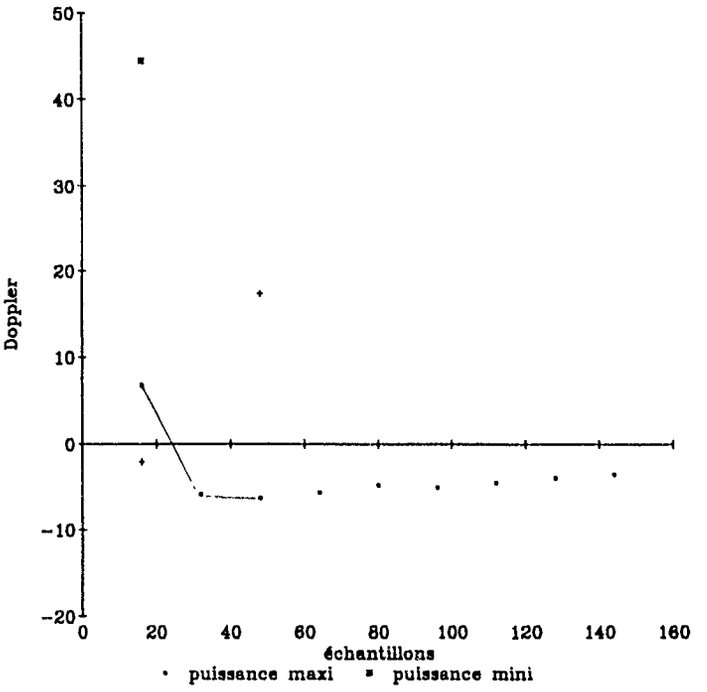


Figure 6(b) Analyse spectrale haute résolution
Fichier Nardeaux, écho n°56, 16 ech.

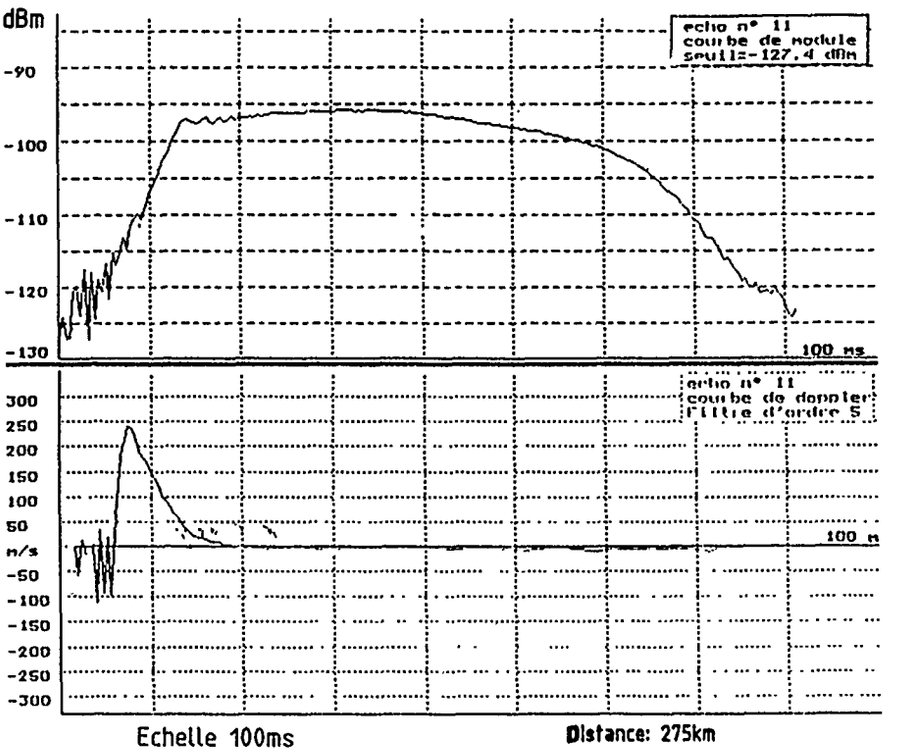


Figure 7(a)

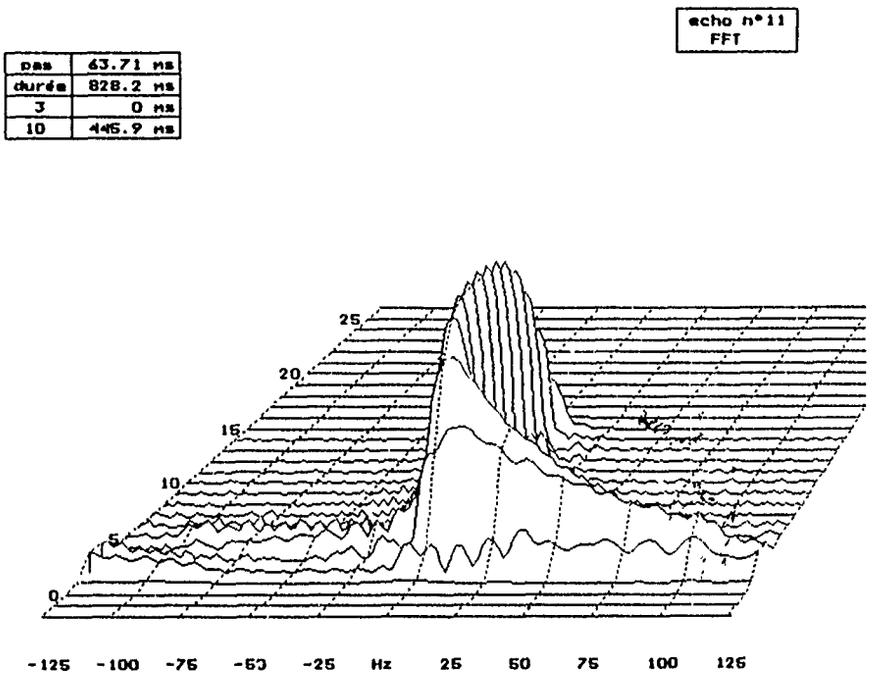


Figure 7(b)

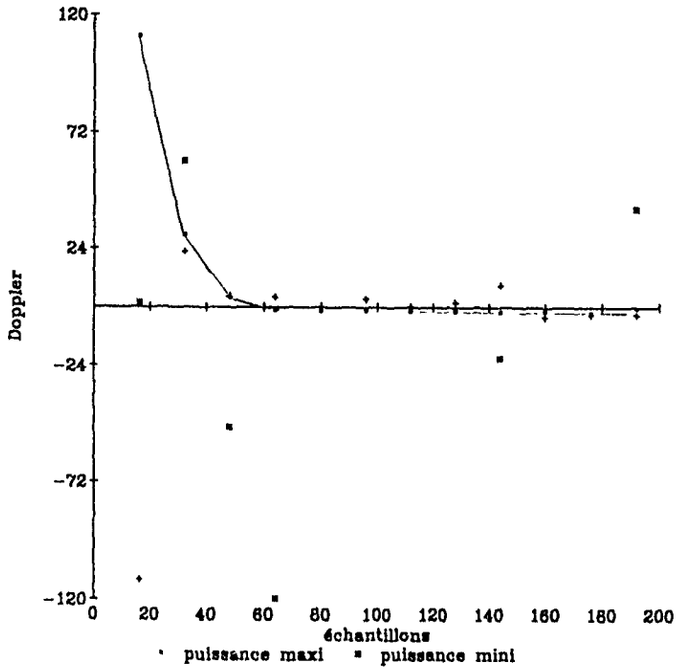


Figure 7(c) Analyse spectrale haute résolution
Fichier Nardeaux, écho n°11, 16 ech

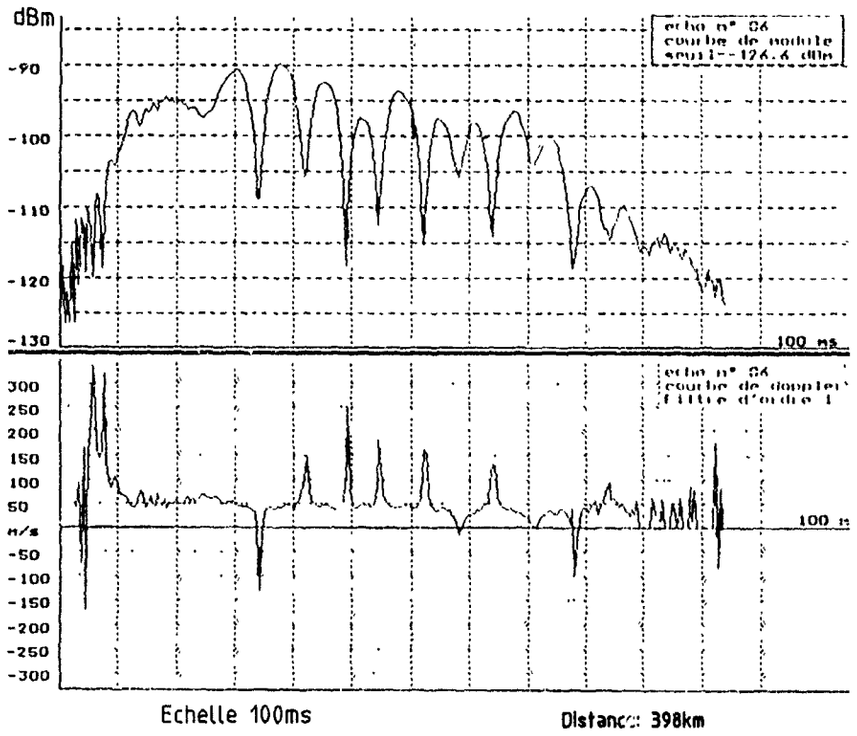


Figure 8(a)

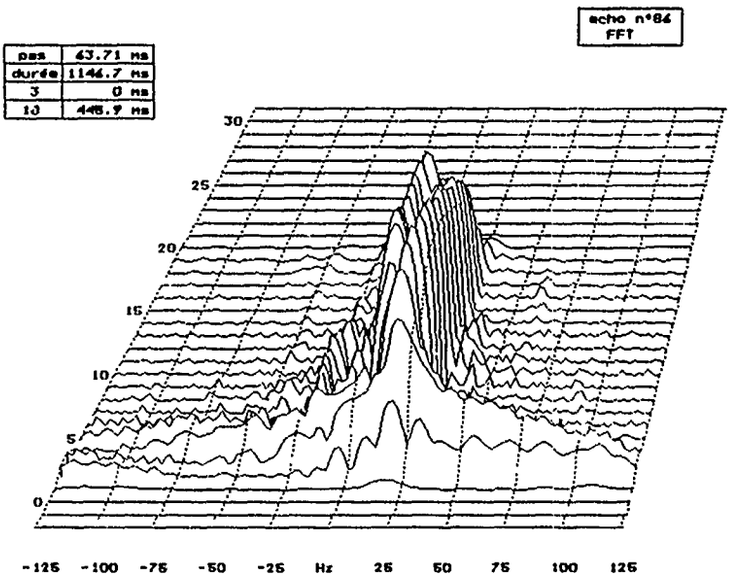


Figure 8(b)

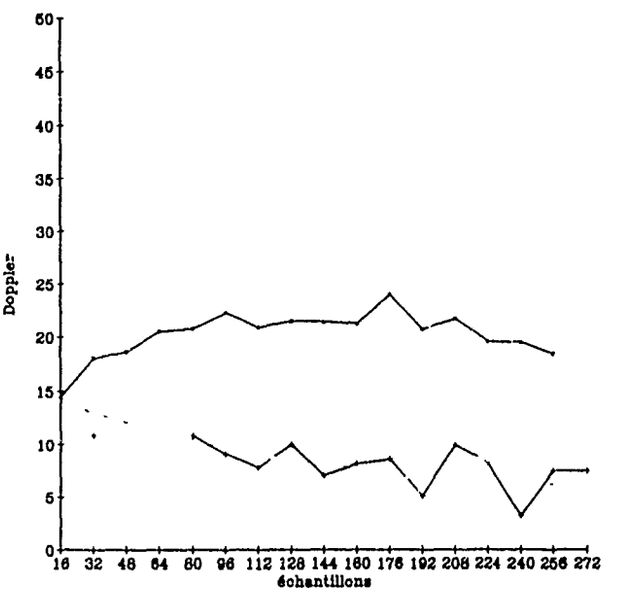


Figure 8(c) Analyse spectrale haute résolution
Fichier Nardeaux, écho n°86, 16 ech.

En conclusion, la faisabilité d'un radar bistatique est fortement conditionnée par le découplage entre l'émetteur (E) et le receteur (R).

Pour cela, les deux stations ont été éloignées de 7 km. Chaque aérien est constitué d'un mini réseau de deux antennes yagi 5 éléments séparées de 3 λ/2 toutes parallèles (dirigées vers la zone du ciel à observer) et perpendiculaire à une ligne passant par les stations E et R, afin de conjuguer la directivité des antennes avec la mise en opposition de phase des ondes, issues de chaque antenne, dans l'axe E-R.

La précision de l'installation des antennes à 3 λ/2 n'étant pas suffisante, le système d' émission, comme celui de réception, est spécialement construit pour accorder les phases et les niveaux des signaux des deux antennes. On atteint ainsi des découplages, entre les deux stations, de l'ordre de - 160 dB malgré la présence de trajets multiples et de variations de propagations dans la journée.

c) Synchronisation des deux stations

Le calage des deux codes (le code émis et la réplique du corrélateur) s'effectue en temps réel. La visualisation sur un oscilloscope du signal corrélé (voir figure 2) permet en "volant" ou en ajoutant un moment du code réplique, de déplacer la pointe de B(t) à l'origine des temps (point de corrélation n°1).

Une synchronisation des stations a été obtenue en asservissant les pilotes sur une émission radiophonique (France-Inter) à l'aide d'un système à verrouillage de phase, afin de pouvoir garantir un déphasage constant entre les deux pilotes.

I I..4 -LE MATERIEL - SES PERFORMANCES

Les aériens décrits plus haut, sont constitués d'antennes Yagi 5 éléments de gain/isotrope de 10 dB, d'angle d'ouverture à 3 dB de l'ordre de 50°. Deux campagnes de mesures à l'aide d'un petit avion ont permis d'assurer la bonne association des antennes deux par deux. Nous avons pu également tracer un certain nombre de diagrammes de rayonnement des mini réseaux, des sites émission et réception, aussi bien en élévation qu'en azimut.

Quelques exemplaires de ces courbes sont représentés figure 5.

L'émetteur est constitué d'une partie commune (alimentation, excitateur, synthèse de fréquence) et de deux chaînes d'amplification restituant 500 Watts, sur chaque antenne.

Chacune des chaînes d'amplification est séparément asservie en puissance. Un déphaseur variable sur une voie permet de rattraper les écarts de phase d'une chaîne par rapport à l'autre.

La stabilité en puissance et en phase sont respectivement inférieures à 5% et 2° après une heure de fonctionnement.

La partie analogique du récepteur est constituée de deux sous-ensembles :

-Une tête HF dont le but est de transposer le signal reçu à une fréquence intermédiaire de 100 KHz.

-Un démodulateur synchrone restituant la partie réelle et imaginaire du signal en bande de base.

Le facteur de bruit de l'ensemble est de 7 dB. Le bruit galactique et industriel a été mesuré à différents moments de la journée. Il culmine à 16 H entre - 156 et - 145 dBm/Hz pour décroître la nuit et au petit matin jusqu'à - 160/- 158 dBm/Hz (soit 14 à 16 dB au dessus du bruit thermique.

La partie numérique du système de réception se décompose en deux sousensembles:

a)Le corrélateur ou ensemble de traitement numérique du signal est construit autour de deux cartes calcul (pour les parties réelles et imaginaires du signal). Une carte séquenceur (AM 2910) assure leur gestion. La conversion analogique-numérique de chacune des voies est effectuée sur 10 bits signés, avec des temps de conversion pouvant descendre jusqu'à 5 µs. La fréquence d'horloge du corrélateur $1/\tau$ est de 5 MHz. Il est capable de calculer ne points de corrélation (ne = 16, 32, 64, 128, 256) en un temps:

$$T = (ne + 2)\, p\, Nr\, \tau$$

La durée Tr d'un moment du code est un multiple entier p (1 à 8) de la période d'échantillonnage Te. La longueur du code Nr est ajustable (15, 31, 63, 127, 255)

La plupart des mesures ont été effectuées pour Nr = 63, p = 2, Tc = Tr x Nr = 3,9816 ms, ne = 128, ce qui est le meilleur compromis entre la fenêtre en distance instantanée d'observation (0 à 600 km) et le nombre de mesures par seconde (250).

Ainsi la plage doppler instantanée d'observation est :

$$\pm Fd = \pm \frac{1}{2\,Tc} = \pm 125 \text{ Hz.}$$

Ce qui signifie qu'il sera possible de mesurer des vitesses radiales jusqu'à :

$$\pm \frac{Fd\lambda}{2} = \pm 350 \text{ m/s}$$

La résolution en distance du radar est

$$\Delta D = \frac{C}{2} Te = 4{,}74 \text{ km}$$

La largeur de bande du signal codé peut être réduite à 1/Tr., ce qui correspond à la bande passante optimale du récepteur analogique soit 16 kHz. C'est finalement un filtre de 20 kHz qui a été utilisé.

b)Un micro-ordinateur (HP 9836) relié au corrélateur via une carte interface programme le système et saisit les résultats des cartes calculs à une cadence de 60 000 mots de 16 bits par seconde. Le stockage sur disquette est effectué dès que la mémoire vive du calculateur est remplie (cela représente entre 1/2 heure et 3 heures de mesures suivant l'importance du nombre des échos).

Une carte supplémentaire (carte reconnaisance de formes) améliore le critère de détection des échos par une analyse du passé de chaque point de corrélation: l'enregistrement des données n'est autorisé que si le dépassement du seuil d'un point de corrélation persiste sur plusieurs prises successives.

Le radar ainsi contitué permet d'observer est échos de faible amplitude jusqu'à - 130 dBm dans les cas les plus favorables (lorsque le bruit et le brouilleur sont faibles) mais presque toujours jusqu'à - 120 dBm.

La surface équivalente d'une traînée sous-dense en rétrodiffusion est donnée par la relation :

$$\sigma = 4\pi r_e^2 L^2 q^2$$

où

$r_e = 2{,}8.10^{-15}$ m rayon effectif de l'électron

2L = Longueur de la première zone de Frénel (m) = $\sqrt{2\lambda Ro}$

q = densité linéique d'électrons de la traînée (m^{-1})

De l'équation (1) on tire

$$\sigma = \frac{Pr\,(4\pi)^3\,Ro^4}{Po\,Go^2\,\lambda^2}$$

Ainsi on peut estimer la sensibilité du radar:

R_o(km)	200	300	400	500	600
σ (m^2)	60	310	1000	2500	5000
L (m)	750	900	1050	1200	1300
q	1.10^{-12}	2.10^{-12}	3.10^{-12}	$4{,}2.10^{-12}$	$5{,}4.10^{-12}$

Tableau 2

III -RESULTATS

III.1.ANALYSE DES SIGNATURES DES TRAINEES IONISEES DE METEORITES

Le traitement à postériori des données, à l'aide d'un micro-ordinateur (PC AT/286), nous a permis de tracer un certain nombre de courbes caractéristiques.

-Variation du niveau en fonction du temps A (t) (courbe module).

-Evolution de la phase du signal $\frac{d\phi\,(t)}{dt}$ (courbe doppler).

-Transformée de Fourier du signal

. FFT du signal complet (sur 512 points de mesure soit 2 ms)

. FFT glissantes (sur 32 points soit 127 ms au pas de 16 Tc soit 63,7 ms).

-Analyse haute résolution du signal.

Quelques exemplaires typiques sont représentés figures 6 à 8.

On distingue alors trois grandes classes de signatures caractérisées par leur courbe de module.

Classe 1 : (Figure 6) il y a apparition brutale d'une région ionisée (traînée) et évanouissement progressif suivant la loi exponentielle $e^{-t/\tau}$.

Ces traînées correspondent parfaitement au modèle des traînées sous denses (q : densité linéique d'électron 2.10^{14} m^{-1}), élaboré par MMs. Mc Kinley (1961) et Sugar (1964).

Dans ce cas

$$\tau = \frac{\lambda^2}{32\pi^2 D}$$

où D est la constante de diffusion.

Classe 2 : (Figure 7)comme précédemment il y a apparition brutale d'une région ionisée, puis le pouvoir de réflexion de cette perturbation varie très peu (augmente ou diminue faiblement) pendant un temps indéterminé, pour enfin décroître rapidement, ce qui caractérise une traînée sur-dense ($q > 2.10^{14}$ m^{-1})

Classe 3 : (Figure 8) dans cette classe sont regroupées les signatures particulières traduisant l'existence de phénomènes plus complexes. L'analyse des spectres de ces échos met en évidence la présence de plusieurs raies, créant ainsi des interférences. L'origine de ces subdivisions semble être causée, essentiellement, par un cisaillement de la traînée par les vents violents des hautes altitudes.

La plupart des études consacrées au sujet font généralement abstraction du temps d'établissement des traînées jugé inutilisable à cause des variations très grandes du doppler et de l'amplitude du signal reçu. Cependant, l'analyse de la phase de formation de la traînée nous a paru intéressante, car elle peut être susceptible de prévoir l'évolution de la phase de persistance et donc d'évaluer la capacité du canal de transmission.

III.1.1.Modélisation de la traînée

Les échos radars obtenus sur les traînées ionisées créés par les météores permettent d'interpréter, au moins partiellement, le mécanisme de formation et le développement de la traînée ainsi que le mouvement des particules neutres dans la haute atmosphère.

III.1.1.a Modélisation du processus de persistance de la traînée

Des relations relatives au développement des traînées après leur formation ont été développées par différents auteurs et on peut trouver une excellente synthèse dans l'article de D.W. BROWNS et H.P. WILLIAMS [4] ainsi que dans celui de G.M. MILLMAN [8].

Dans le cas particulier de la rétrodiffusion, telle qu'elle existe dans les systèmes radar, la première zone de Fresnel [8] (figure 9) est donnée par la relation :

$$L = \left[\frac{\lambda Ro}{2} \right]^{1/2}$$

où

L est la longueur de la demi-zone de Fresnel

λ la longueur d'onde

Ro la distance minimale de la trajectoire du météore au radar.

Dans le cas des traînées sous-denses, qui constituent la majeure partie des traînées observées, la puissance reçue Pr, après formation de la traînée est liée à la puissance émise Po par la relation :

$$\frac{Po}{P_t} = KG_r G_t \frac{\lambda^3}{Ro^3} q^2 \cos^2\mu \exp\left(\frac{-8\pi^2 ro^2}{\lambda^2} \right) \exp\left(\frac{-32\pi^2}{\lambda^2} D(t-to) \right) \quad (9)$$

où

K (= $2{,}48.10^{-32}$ m^2) est une constante déduite des paramètres physiques du processus de création de la traînée et de l'équation du radar.

ro le rayon initial de la traînée donnée par la relation empirique [4]
$\log ro_{(m)} = 0{,}035\ h - 3{,}45$.

D la constante de diffusion ambipolaire donnée par la relation empirique [23]
$\log D_{(m^2/s)} = 0{,}067\ h - 56$.

μ l'angle entre les axes de polarisation supposées rectilignes aux antennes d'émission et de réception.

$t\text{-}to \geq \frac{L}{V}$ t: le temps mesuré à partir de l'instant où le météore est à la distance Ro de l'émetteur et lorsqu'il a traversé totalement la première zone de Fresnel.

Ces relations, comparées à celles valables en liaisons bistatiques, montrent que les durées des traînées Te mesurées lorsque la puissance reçue est réduite dans le rapport $1/e^2$ par rapport à sa valeur maximale et donnée par :

$$Te = \frac{\lambda^2}{16\pi^2 D \cos^2\phi} \qquad (10)$$

sont notablement plus faibles en rétrodiffusion qu'en transmission, d'une part à cause du facteur $\cos^2\phi$ ($\cos^2\phi = 1$ en rétrodiffusion, $\cos^2\phi = 0{,}05$ pour une liaison de 1000 km) et d'autre part à cause de la réduction de la longueur de la première zone de Fresnel.

III.1.1.b Modélisation de la formation de la traînée

L'étude des traînées météoriques à l'aide d'un radar permet d'observer la formation des traînées, c'est-à-dire le développement de la traînée lorsque le météore traverse la première zone de Fresnel (voire les suivantes).

Si un mobile se déplace à une vitesse V uniforme sur une trajectoire rectiligne (figure 10), la fréquence doppler mesurée par un radar s'exprime par :

$$Fd = Fd_{max} \sin\left[\mathrm{Arctg}\,\frac{V}{Ro}(t\text{-}to)\right] \qquad (11)$$

où

$Fd_{max} = 2\,V\,fo/c$

fo: fréquence d'émission

c: célérité de la lumière

to: instant où le mobile passe à la distance Ro du radar.

Cette relation conduit directement à

$$Vr = V \sin\left[\mathrm{Arctg}\,\frac{V}{Ro}(t\text{-}to)\right] \qquad (12)$$

vitesse radiale de la cible.

L'accélération apparente de la cible :

$$\frac{dVr}{dt} = \frac{V^2}{Ro}\,\frac{1}{1+\frac{V^2(t\text{-}to)^2}{Ro^2}}\cos\left[\mathrm{Arctg}\left(\frac{1}{Ro}(t\text{-}to)\right)\right] \qquad (13)$$

L'angle φ sous lequel est vue la cible par rapport à la normale à la trajectoire est donnée par

$$\varphi = \mathrm{Artg}\left(\frac{V}{Ro}(t\text{-}to)\right) \qquad (14)$$

L'angle 2φ max sous lequel elle est vue la première zone de Fresnel est définie par (15).

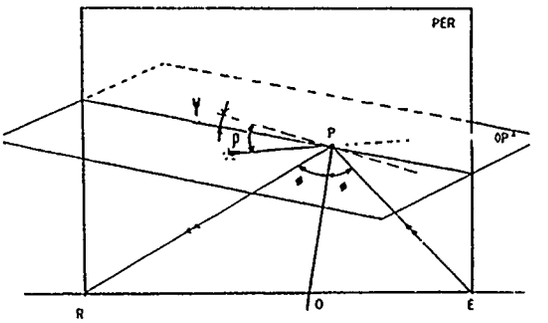


Figure 9 Géométrie d'une prodiffusion

ϕ Angle d'incidence
β Angle entre l'axe de la traînée et le plan de propagation PER
ψ Angle entre l'axe de la traînée et le plan perpendiculaire à OP: OP' ou déviation spéculaire

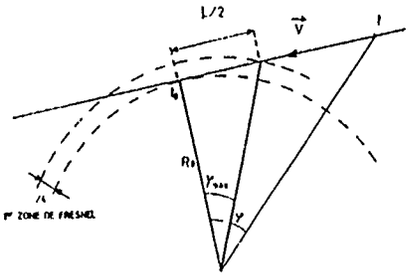


Figure 10

$$\mathrm{tg}\,\varphi_{max} = \frac{L}{Ro} = \sqrt{\frac{\lambda}{2Ro}} \qquad (15)$$

A 53 MHz, pour un objet se déplaçant à 25 km/s on obtient le tableau ci-dessous où T_{2L} : temps de traversée de la zone de Fresnel.

Au voisinage de la zone de Fresnel

$\varphi \ll 1$rd.

$$\text{donc} \quad Vr \# \frac{V^2}{Ro}(t\text{-}to) \qquad (16)$$

$$\frac{dVr}{dt} \# \frac{V^2}{Ro} \qquad (17)$$

Ro (km)	200	300	400	500	600
L (m)	750	900	1050	1200	1300
φ max (mrd)	3,75	3,0	2,62	2,4	2,15
T_{2L} (ms)	60	72	84	96	104

Tableau 3

La vitesse radiale est une fonction lineal-

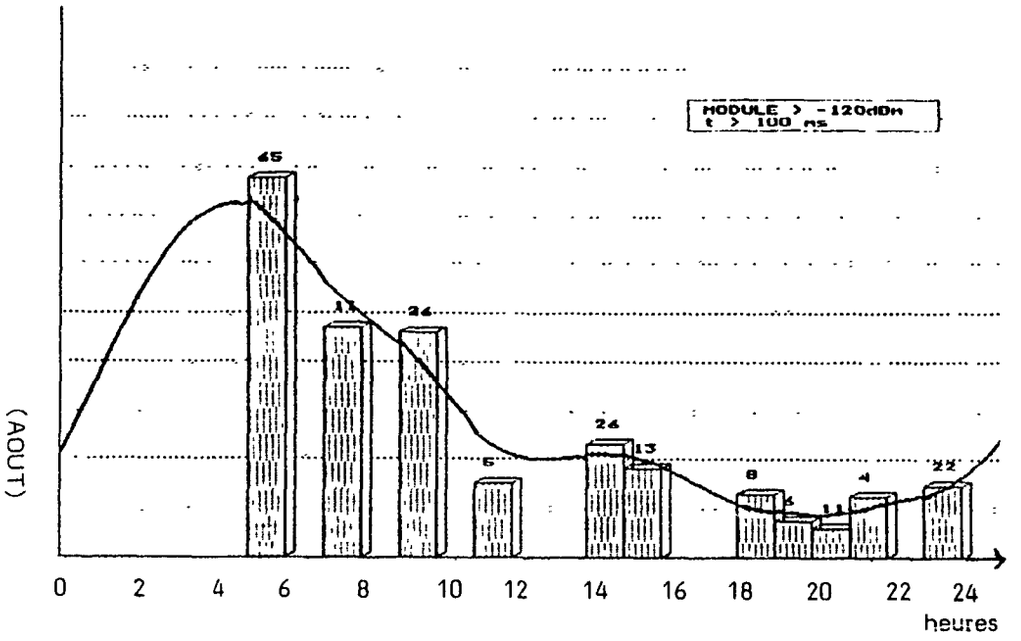


Figure 10(a) Nombre de traînées

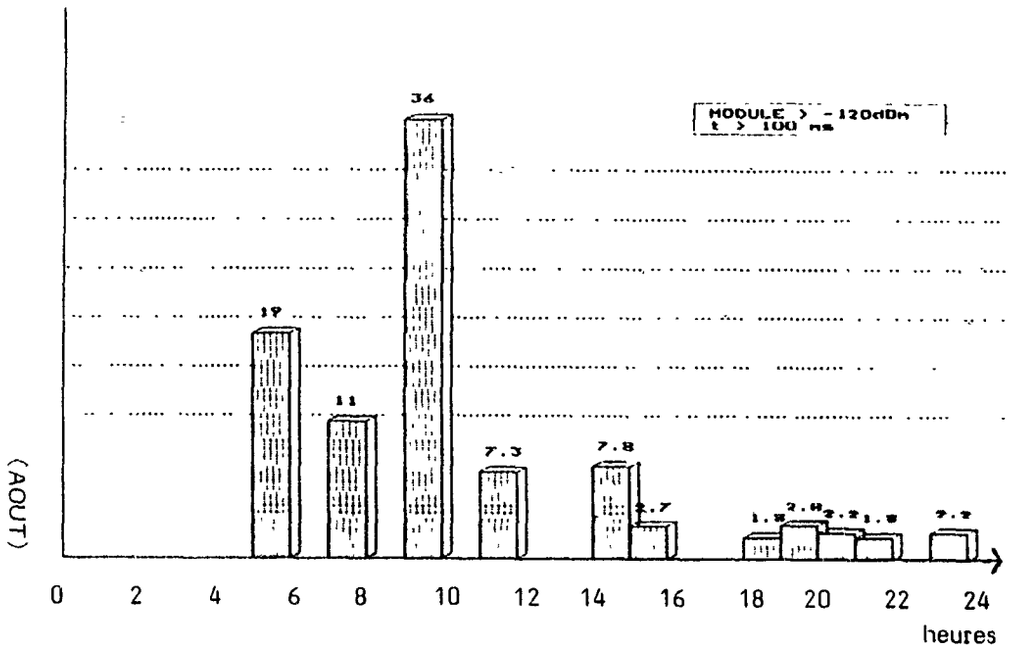


Figure 10(b) Temps d'ouverture

re du temps, l'accélération est constante et ne dépend que de deux paramètres dont l'un, Ro, est directement accessible par la mesure.

Les hauteurs auxquelles apparaissent les échos sur les traînées ionisées suivent des distributions sensiblement gaussiennes centrées sur une altitude h donnée par la relation [23].

$$h = -17 \log F_{(MHz)} + 124$$

A 53 MHz h = 95 km et l'écart type de la distribution est de l'ordre de 3 km. Un météore produit un écho si la première zone de Fresnel de sa trajectoire est totalement ou partiellement dans la zone d'activité définie précédemment. On peut distinguer 3 types de trajectoires (Fig. 11) selon que la première zone de Fresnel est totalement contenue dans la zone d'activité (trajectoire 1) ou que seul le début (trajectoire 2) ou la fin (trajectoire 3) se trouve dans cette zone.

En orientant l'axe Mol dans le sens de la trajectoire, un élément dl de la traînée crée un signal reçu par le radar dE(t) que l'on calcule en prenant compte l'ionisation initiale, la diffusion ambipolaire et la propagation de l'onde.

$$dE(t) = B \exp(-t'/\tau) \exp\left(-j\frac{2\pi l^2}{\lambda Ro}\right) \exp\left(j\omega o\left(t - \frac{Ro}{c}\right)\right) |l| dl \qquad (19)$$

où

$$B = [Gt\,Gr]^{1/2} \frac{\lambda}{2r^2} \exp\left(-4\pi^2 r^2/\lambda^2\right) \frac{re}{\pi} q \qquad (20)$$

- r = distance de l'élément dl au radar r # Ro
- re = rayon effectif de l'électron
- q = densité linéique de la traînée supposée constante pendant la traversée de la zone d'activité.
- $\tau = \frac{\lambda^2}{32\pi^2 D}$ D est la constante de diffusion
- $t' = t - to + l/v$
- l = l'abcisse sur la trajectoire avec origine sur la normale à la trajectoire.

En négligeant les variations de exp (-t'/τ)lors de la traversée de la demi zone de Fresnel et en appelant α la valeur moyenne de ce terme : (21)

$$dE(t) = B\alpha \exp\left(\frac{-j2\pi p^2}{Ro\lambda}\right) |l| dl \exp(-j\omega o\,(t - Ro/2))$$

pour $-\frac{L}{2} < l < o$

$$E(t) \equiv j \frac{B\,Ro\lambda\alpha}{2\pi} \sin\left[\frac{\pi}{2}\left(1 - \frac{l^2}{L^2}\right)\right] \exp\left(j\left[\omega o\left(t - \frac{Ro}{c} - \frac{\pi}{2}\left(1 - \frac{l^2}{L^2}\right)\right)\right]\right) \qquad (22)$$

La phase instantanée vaut alors :

$$\varphi(t) = \omega o\left(t - \frac{Ro}{c}\right) - \frac{\pi}{2}\left(-\frac{l^2}{L^2}\right) \qquad (23)$$

$$\text{d'où } fd = \frac{1}{2\pi}\left[\frac{d\varphi(t)}{dt} - \omega o\right] \frac{lV}{Ro\lambda} \qquad (24)$$

La vitesse apparente vaut donc :

$$Va = \frac{lV}{2Ro} = \frac{V^2(t-to)}{2Ro} \quad \text{soit } Va = \frac{Vr}{2} \qquad (25)$$

Il apparaît donc que la vitesse apparente est comprise entre Vr (trace totalement persistante), et 2Vr (absence de persistance).

III.1.1.c Forme des échos

Compte-tenu des relations (11) (12) (22) (25) la forme des échos peut être déterminée et l'allure générale est celle représentée sur la figure 12.

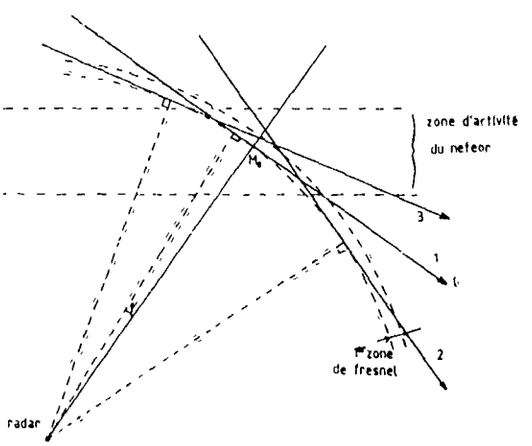


Figure 11

Dans la phase de formation de la traînée, l'amplitude croît (relation 22) et la vitesse apparente mesurée sur la fréquence doppler instantanée décroît linéairement (formule 16).

La vitesse du météore se déduit de la vitesse apparente par la relation

$$V = \left[2\,Ro\frac{dVa}{dt}\right]^{1/2}$$

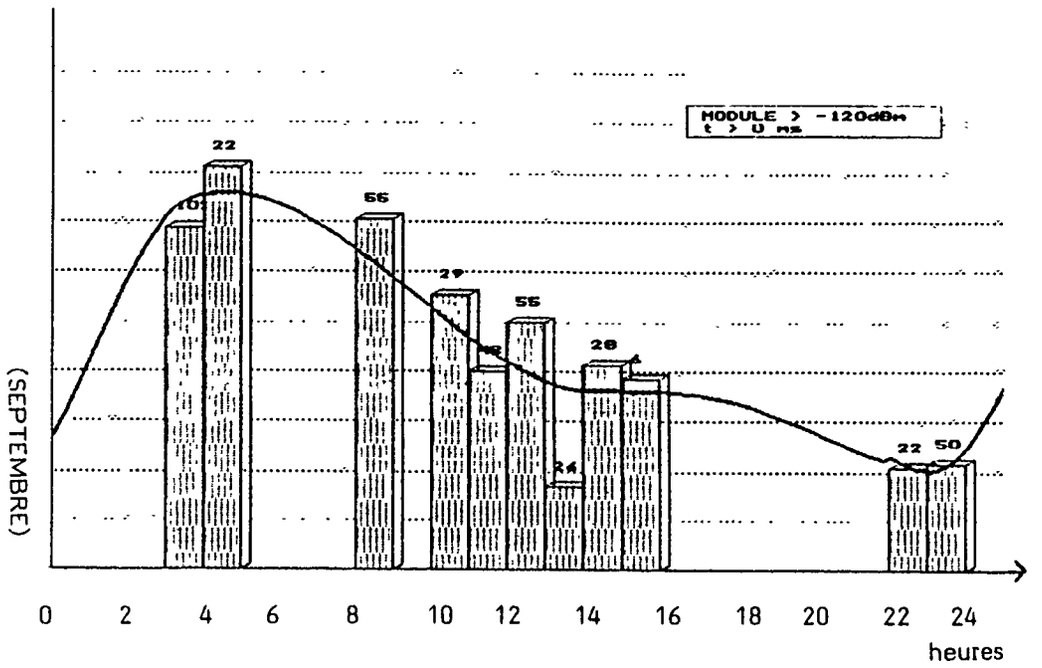


Figure 11(a) Nombre de traînées

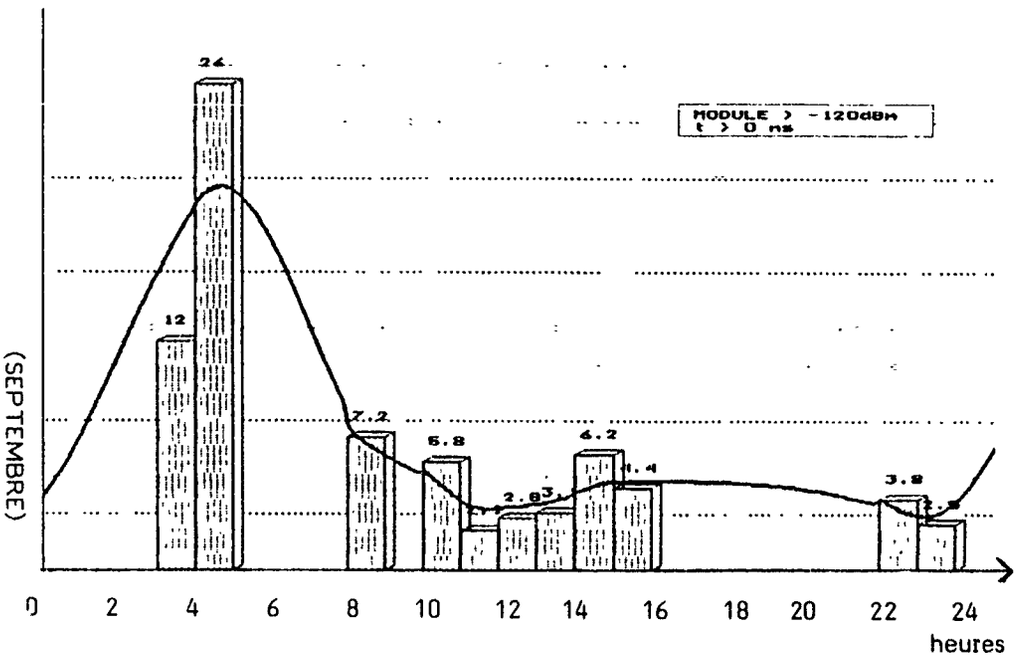


Figure 11(b) Temps d'ouverture ‰

Un calcul simple donne l'angle ψ (fig. 11) qui pemet de déterminer les types de trajectoire 1, 2 ou 3. Si V_{am} est la vitesse moyennee de la phase de formation, alors

$$\psi \# \frac{2 V_{am}}{[Ro \frac{dVa}{dt}]^{1/2}}$$

Dans la phase de persistance de la traîné la vitesse apparente est celle du vent de particules neutres. Lorsque les vents de cisaillement apparaissent, cette trace, comme le montre les résultats expérimentaux fait apparaître des composantes spectrales multiples.

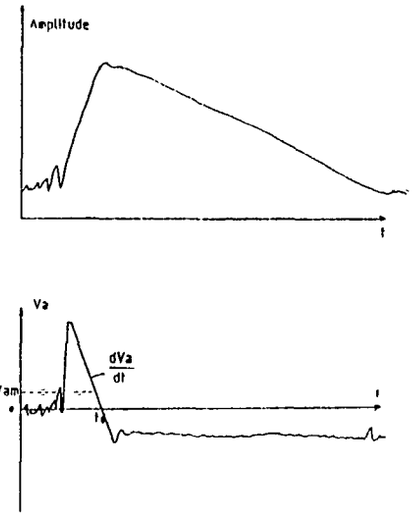


Figure 12

III.1.1.d Résultats expérimentaux

Trois types de résultats sont présentés sur les figures 6, 7 et 8.

La figure 6 donne la signature d'une traînée sous-dense détectée à une distance de 171 km. On peut noter l'allure caractéristique de la variation d'amplitude et de la vitesse apparente en bon accord avec le modèle proposé.Les mesures de $\frac{dV}{dt}$ et de V am donnent V = 45 km/s et ψ = 1,7 mrd, valeurs que consolident les estimations classiques. Dans la phase de persistance, la traînée est déplacée par un vent de particules neutres de vitesse 20 m/s. L'image haute résolution montre l'existence d'une seule raie.

Les figures 7 donnent une représentation semblable pour une traînée sur-dense détectée à une distance de 275 km. Les calculs identiques aux précédents donnent V = 43 km/s , ψ = 6 mrd. La persistance de la traînée montre l'absence de composante radiale de vent. L'analyse spectrale et l'analyse haute résolution montrent l'évolution de l'effet doppler qui met clairement en évidence les phases de formation et de persistance.

Les figures 8 montrent un cas plus complexe d'une signature obtenue en présence de deux composantes de vents neutres, à une distance de détection de 398 km. La phase de formation donne V = 51 km/s et ψ =3 mrd. L'analyse spectrale et l'analyse haute résolution montrent l'évolution rapide du doppler dans la phase de formation et dans la phase de persistance l'exitence de deux raies qui traduisent la présence de deux composantes radiales du vent.

III.2. ANALYSE STATISTIQUE DES TRAINES IONISEES DE METEORES

L'ensemble des expérimentations est récapitulée dans le tableau 4. Elles représentent au total 139 heures de mesure et l'enregistrement de plus de 11000 traînées.

On sait que le nombre N des traînées détectées dépend de la sensibilité S du récepteur ainsi que de la nature de ces traînées.

Pour les traînées sous-denses
$(q < 2.10^{14} m^{-1})$: $N \propto S^{-1/2}$

Pour les traînées sur-denses
$(q > 2.10^{14} m^{-1})$: $N \propto S^{-2}$

où S est le seuil de détection exprimé en watt.

En pratique, le système ne faisant pas la distinction entre les deux types de traînées $N \propto S^{\gamma n/2}$.

γ^n variera entre - 1 et - 4 en fonction de la proportion de traînées sous-denses.

Par extension, on peut exprimer Dc, le taux d'ouverture du circuit (c'est-à-dire le rapport du temps d'utilisation du canal au temps d'observation) :

$$Dc \propto S^{\gamma d/2}$$

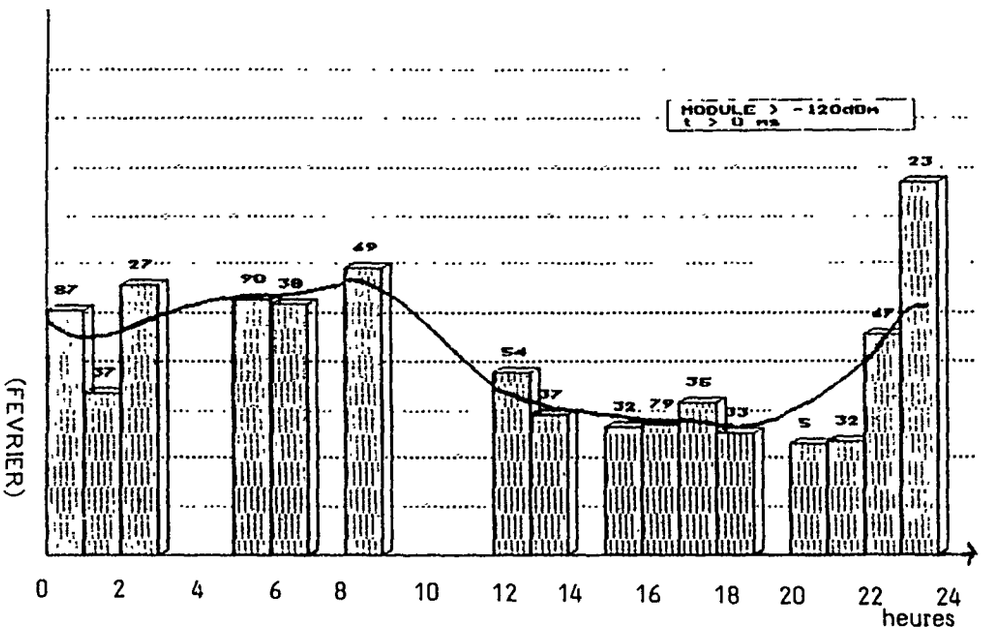


Figure 12(a) Nombre de traînées

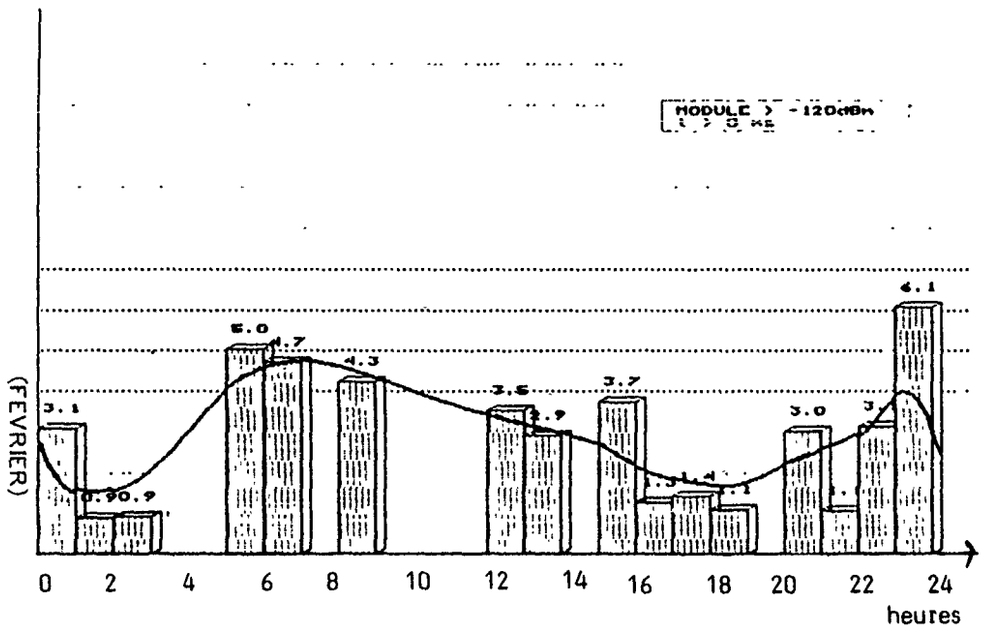


Figure 12(b) Temps d'ouverture ‰

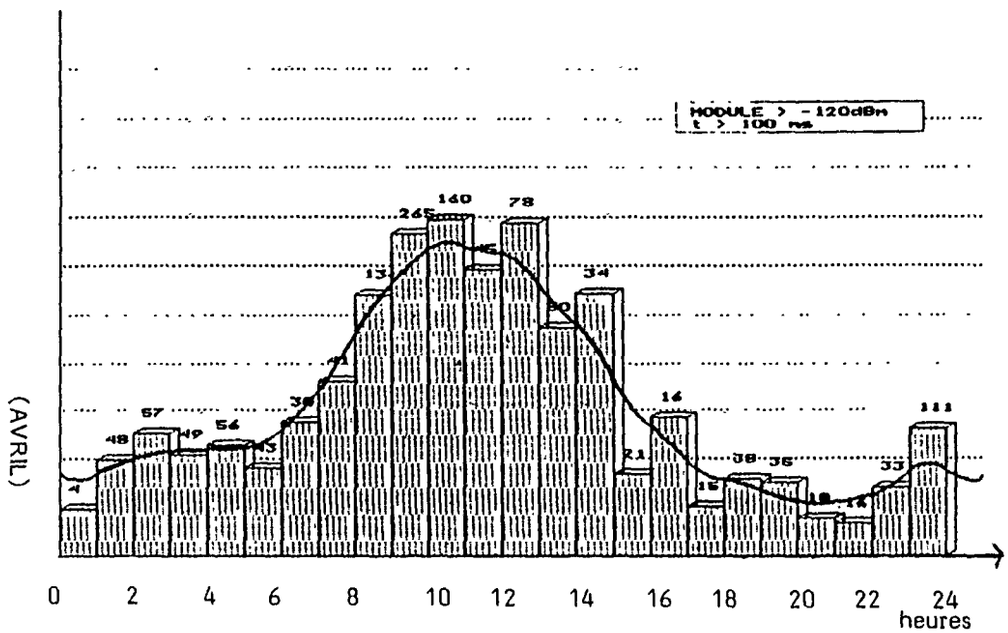


Figure 13(a) Nombre de traînées

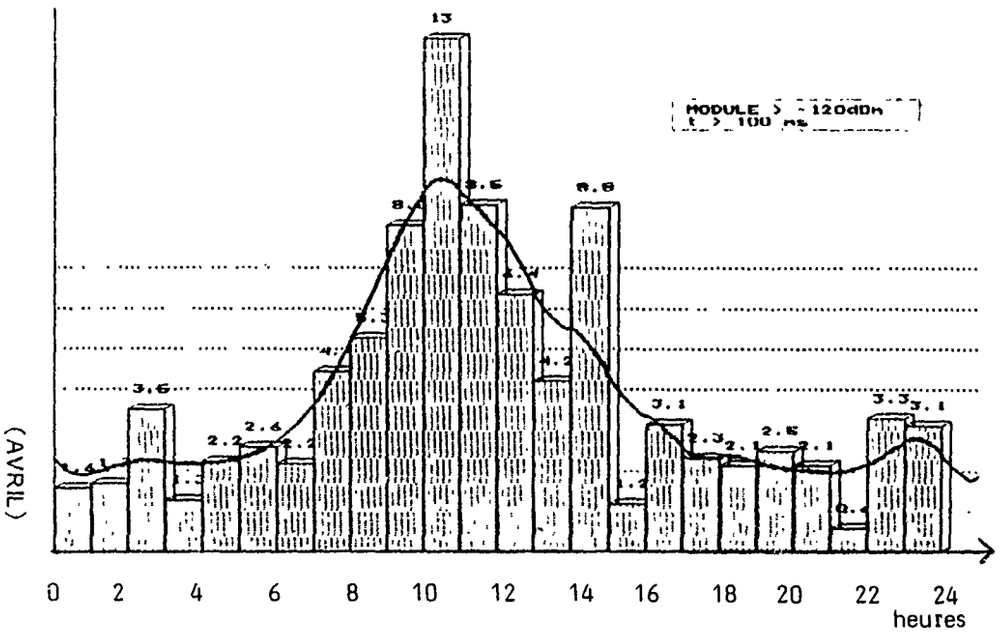


Figure 13(b) Temps d'ouverture ‰

- Les courbes de la figure 15 montrent la variation du temps d'ouverture du canal en fonction des seuils de détection. On remarque très nettement que pour des seuils faibles (-120 dBm à - 105 dBm) la proportion de traînées sous-denses est importante. Le calcul de γd donne des valeurs proches de - 1,7. Par contre, lorsque le seuil est important, γd augmente: pour un seuil S de - 95 dBm, γd vaut - 4. La proportion de traînées sur-denses est alors proche de 100%.

III.2.a Répartition journalière et saisonnière

Les variations saisonnières apparaissent en conformité avec les observations faites par le passé : maximum en Août, minimum en Février.

Les graphismes (figures 10 à 13) donnent les variations journalières pour différentes saisons (respectivement Août et Septembre 1989 ainsi que Février et Avril 1990).

Les cycles journaliers apparaissent, dans leur ensemble, conformes à ceux signalés par divers auteurs (David W. BROWN 1985, Lloyd VANCIL 1984). On distingue un maximum d'activité vers 6 H du matin et un minimum à 18 H (heure du soleil).

Cependant les variations de ces cycles en fonction des saisons montrent des fluctuations:

- Lorsque l'activité moyenne diminue, les variations journalières sont moins importantes (tableau 5).

- Superposé à cette activité moyenne apparaît une forte concentration de traînées très localisées dans le temps (date et heure) correspondant à des pluies de météores, répertoriées dans la littérature. Ainsi au mois d'Août, en pleine période des Perséïdes, le taux d'ouverture du circuit et multiplié par 3,5 à 9 H. Cependant, on peut noter au mois d'Avril une activité importante entre 8 H et 15 H avec un maximum à 10 / 11 H, alors qu'aucune pluie de météores importante n'est connue. Les variations importantes et inattendues du taux d'ouverture (rapport 4) et du nombre de traînées (rapport 3) ne sont donc pas expliquées.

Le tableau 6 représente les variations d'un certain nombre de paramètres significatifs en fonction de l'heure pour différentes séries de mesures dans l'année.(Pour S = - 115 dB et une durée minimale des échos de 100 ms).

N : le nombre de traînées par heure

Dc : le taux d' ouverture du circuit en‰

Pm : la puissance moyene détectée fW

γn et γd, les deux exposants décrits plus haut

A priori, il ne semble pas y avoir une relation simple entre l'activité météorique et l'exposant γd ou γn, c'est-à-dire de la proportion de traînées sous-denses par rapport aux surdenses. On remarque que les perturbations du mois d'Août et d'Avril ne sont pas de même nature: pour des seuils S inférieurs à - 110 dBm γd est très faible au mois d'Août (- 0,95) par contre, il est important au mois d'Avril (- 2). Ainsi, la puissance moyenne reçue au mois d'Avril varie peu (rapport 3), contrairement au mois d'Août (rapport 36 !).

III.2.b Influence de la polarisation

Le tableau 7 (N et Dc sont donnés pour S = -115 dB et une durée des échos > 100 ms) montre que la polarisation horizontale est nettement plus favorable, conformément aux observations effectuées par d'autres auteurs [13].

γd et γn sont plus faibles en polarisation horizontale, le rapport d'efficacité entre les deux polarisations s'accentue donc avec l'augmentation du seuil de détection.

L'accroissement de la surface radar équivalente en polarisation horizontale repousse donc le seuil de détection des traînées sous-denses vers des niveaux supérieurs, ce qui contribue à diminuer γn et γd.

Les résultats des expériences en polarisation croisée montrent le faible effet de dépolarisation.

Les tableaux 8 récapitulent les mesures effectuées au mois d'Avril pour différentes polarisations, différents seuils de détection et différentes durées d'échos:

- Le nombre N de traînées en %
- Le taux d'ouverture Dc en ‰
- Le temps d'attente entre deux météores en minutes.

III.2.c Influence de la fréquence

L'expression des constantes de temps de la loi de décroissance exponentielle des échos (τ et τ'), donne pour les deux fréquences de travail

τ' = 23,5 ms à 40 MHz

τ = 18,5 ms à 53 MHz.

Les graphiques (14) représentent la répartition des traînées de classe 1 en fonction de leur pente $Pe = \frac{\Delta T}{\Delta S}$ (variation de leur durée / variation du seuil de détection).

On constate alors, que la courbe passe par un maximum entre 7 et 8 ms/dB à 53 MHz, et entre 8 et 9 ms/dB à 40 MHz. Si on tient compte du temps de montée des échos (de l'ordre de 3 ms/dB) on retrouve :

Pe' # 5,5 ms/dB soit τ' # 24 ms

Pe # 4,5 ms/dB soit τ # 20 ms.

D'autre part, les tableaux 9 montrent qu'il y a un rapport rm = 3,3 sur le nombre de traînées, et un rapport rd = 3,9 sur le taux d'ouverture du circuit (pour un seuil S = - 115 dBm, et une durée minimale des échos de 100 ms).

Si on exprime Dc en fonction de τ et de N

$$Dc \propto N\tau$$

dans ce cas $rd = \frac{\tau'}{\tau} r_n$

L'expérience nous donne τ' / τ # 1,2 et rd / r_n = 1,18.

Le rapport des énergies reçues est de rm = 4,2 (tableau 9). rm comme rd est de l'ordre de $(\frac{\lambda'}{\lambda})^5 = 3,9$ (λ étant la longueur d'ordre à 53 MHz et λ' à 40 MHz).

Tous ces résultats montrent une très bonne cohérence entre résultats expérimentaux et les relations actuellement admises.

Le calcul de γn et γd (tableau 9) pour des seuils faibles, montre que la proportion de traînées sous-denses est plus importante à 40 MHz,ce qui explique d'une part que rm > rd et d'autre part que l'écart d'efficacité du système s'accroît avec le seuil S (comme le montre les tableaux 10).

III.2.d Influence de l'orientation

Les mesures effectuées à l'ouest dans les mêmes conditions que celles faites à l'est et dans un intervalle de temps assez court n'ont pas montré de variation sensible. Ce résultat vérifié en dehors des pluies de météores est en bon accord avec l'hypothèse d'isotropie des flux de météores isolés.

CONCLUSION

L'étude présentée dans cette communication a été menée avec un système expérimental sensible qui a permis de collectionner plus de 11000 signatures de traînées météoriques et d'effectuer une étude fine grâce à des analyses cohérentes.

Les phases de formation et de pesistance des traînées ont pu être séparées et leur analyse a fourni des informations sur leur vitesse, leur orientation et les mouvements des particules neutres en haute altitude.

Une étude statistique, confirmant les résultats publiés par différents auteurs, a fourni des complémentaires sur la répartition des traînées, les proportions des traînées surdenses et sous-denses et sur les temps d'utilisation du canal en rétrodiffusion.

Ces résultats apportent une meilleure compréhension des phénomènes et constituent une étape dans l'approche de la conception de systèmes opérationnels, l'un des objectifs ultime de ces études.

BIBLIOGRAPHIE

[1] P.E. FORSYTH, E.L VOGAN, D.R. HANSEN, C.O. HINES :
The principes of JANET - a meteor burst communication system.
Proc. IRE Vol. 45 pp 1462-1660 Dec 1957 et 11 articles sur le même thème.

[2] L.L. CAMPELL, C.O. HINES. Bandwidth considerations in a JANET system.
Proc. IRE Dec 1957.

[3] P.J. BARTHOLOME, I.M. VOGT : COMET. A meteor burst system incorporating ARQ and diversity reception.
IEEE Trans Comm. Technol. Vol. COM-16 April 1968.

[4] D.W. BROWN HP WILLIAMS. The performance of meteor burst Communications at different frequencies.
23eme symposium EPP. AGARD 3-7 October 1977 Cambridge Massachusetts U.S.A.

[5] P. SICILIA, D. SORAIS, F. BARRIER - Expérimentation d'un système de transmission de données par canal météorique : Théorème. Symposium EPP. AGARD. 1988.

[6] G.R SUGAR Propagation via meteor trails
Proc. IRE Vol. 52 Feb. 1964.

[7] J.W. KOCH Factors effecting modulation techniques for VHF scatter systems
IRE trans comm. syst. Vol. C3-7 June 1959.

[8] G.H. MILMAN HF scatter from over dense meteor trail.
23eme symposium EPP AGARD 3-7 October 1977 Cambridge Massachusetts U.S.A.

[9] M.D. GROSSI, A. JAVED Time and frequency spread in meteor burst propagation paths.
23eme symposium EPP AGARD 3-7 October 1977 Cambridge Massachusetts U.S.A.

[10] D.W. BROWN Some observations and prediction based on meteor burst communications system modeling.
MILCOM Conf. Proc. Oct. 1985.

[11] D.W. BROWN A physical meteor burst propagation modem and some significant results for communication system design.IEEE J. Selec. Areas Communication SAC-3 1985.

[12] J.A. WEITZEN Predicting the arrival of meteor useful for meteor burst communications Radio Science.
Vol. 21 N6 Nov. Dec. 1986.

[13] P.S. CANNON Polarization rotation in meteor burst communications systems.
Radio Science Vol. 21 N3 May June 1986.

[14] J.A. WEITZEN. Effects of polarisation coupling lois mechanism of design of meteor scatter antennas for short and long range communication.

[15] Y. CHANDRAMOULI,M.F. NEUTS, V. RAMASWAMI. A queing model for meteor burst packet communication systems.
IEEE Trans. Comm. Vol. 37 N10 October 1989.

[16] M.B. PURSLEY - S.D. SANDBERG. Variable rate coding for meteor burst communications.
IEEE Trans. Comm. Vol. 37 N11 November 1989.

[17] J.J. METZNER Improved coding strategies for meteor burst communications.
IEEE Trans. Comm. Vol. 30 N2 February 1990.

[18] M.A. GELLER,s.a. BOWHILL,G.C.HESS. A description of the University of Illinois meteor radar system and some first results.
J. Atmosph. Terr. Phys. N39 1977.

[19] T. ASO, T. TSUDA,S. KATO. Meteor radar observations at Kyoto University.
J. Atmosph. Terr. Phys. N41 1979.

[20] S.K. AVERY,A.C. RIDDLE, B.B. BASLEY. The poker flat. Alaska MST radar as a meteor radar.
Radio Science N18 1983.

[21] R.G. ROP'ER, MWR meteor wind radars Handb. M.A.P. N13 1984.

[22] STUDIO : Electromagnetic founding technique using special and spectral sampling of the reception signal.
C; GOUTELARD - J. CARATORY - A. JOISEL.
Aspects of electromagnetic wave scattering in radio communications. BOSTON October 1977.

[23] D.W.R. Mc KINLEY. Meteor Science and Enginnering.
Mc Graw-Hill Book Co Inc. NY 1961.

Mesures en Rétrodiffusion

Dates	Polarisation	Orientation	Fréquence
10/08 - 13/08/89	Verticale	110° Est	53 MHz
01/09 - 07/09/89	Verticale	110° Est	53 MHz
18/10 - 23/10/89	Verticale	110° Est	53 MHz
25/10/89	Horizontale	110° Est	53 MHz
02/11/89	Verticale	110° Est	53 MHz
07/11 - 09/11/89	Verticale	70° Ouest	53 MHz
01/12 - 11/12/89	Verticale	110° Est	40 MHz
20/02 - 08/03/90	Verticale	134° Ouest	53 MHz
10/04 - 17/04/90	Verticale	46° Est	53 MHz
18/04 - 23/04/90	Croisée	46° Est	53 MHz
26/04 - 02/05/90	Horizontale	46° Est	53 MHz

Tableau 4
Rapport d'activité météorique entre 6h et 18h (heure du soleil)

	Nombre de traînées moyen par heure (N)	Rapport	Taux d' ouverture moyen (Dc)	Rapport
Août *	70	6,6	10	9,5
Septembre	50	4	8	6,3
Mars	20	2,5	2,5	2
Avril*	15	2	2,5	1

Tableau 5
Variation de l'activité météorique en fonction de l'heure

	N	Dc	Pm	γn	γd
Avril					
15 h à 18 h	7,5	0,64	18,5	- 1,5	- 1,5
18 h à 14 h	29,5	2,88	52,5	- 1,5	- 1,7
10 h	40	4,27	62	- 1,5	- 2
Août					
16 h à 24 h	12,5	1,79	34,5	- 1,5	- 1,5
5 h à 8 h	47,5	7,08	184	- 1,5	- 1,5
9 h	66	25,03	1267	- 0,8	- 0,95
Septembre					
10 h à 24 h	14,5	4,28	17	- 1,5	- 1,3
3 h à 5 h	38	7,38	104	- 1,2	- 1

Tableau 6
Comparaison entre un sondage en polarisation verticale et horizontale
(mesures effectuées entre 8h et 12h)

		N	Dc	Pm	γ_n	γ_d
Octobre	Verticale	12,5	1,15	18	- 1,9	- 2
	Horizontale	21	2,52	85	- 1,2	- 1,2
	Rapport	1,67	2,19	4,7		
Avril	Verticale	30,5	2,92	52,5	- 1,5	- 1,8
	Horizontale	54,5	5,33	159	- 1,2	- 1,5
	Rapport	1,8	1,8	3,05		

Tableau 7
Comparaison entre un sondage à 53 MHz et 40 MHz

Temps d'attente ' Temps d'ouverture ‰ Nombre de traînées %

	>-120dBm	>-115dBm	>-110dBm	>-105dBm	>-100dBm	>-95dBm	>-90dBm
t>0ms	0.3' 100 9.68%.	0.6' 54% 3.76%.	1.2' 25% 1.46%.	3.1' 10% 0.58%.	10.9' 3% 0.14%.	35.8' 1% 0.01%.	
t>100ms	0.8' 36% 8.10%.	2.0' 15% 2.92%.	5.2' 6% 1.06%.	13.9' 2% 0.36%.	55.6' 1% 0.09%.	0.00%.	
t>250ms	2.3' 13% 6.19%.	5.7' 5% 2.06%.	17.3' 2% 0.72%.	41.7' 1% 0.25%.	2.1H 0.06%.		
t>500ms	5.0' 6% 4.84%.	15.2' 2% 1.45%.	38.5' 1% 0.53%.	83.4' 0.17%.	0.02%.		
t>1s	10.4' 3% 3.66%.	27.8' 1% 1.12%.	62.6' 0.41%.	0.04%.			
t>2s	25.0' 1% 2.32%.	71.5' 0.69%.	0.07%.				

NOMBRE TOTAL DES ECHOS :1709
DUREE TOTALE DES MESURES : 8 H 20 mn 42 s
POLARISATION VERTICALE

	>-120dBm	>-115dBm	>-110dBm	>-105dBm	>-100dBm	>-95dBm	>-90dBm
t>0ms	1.2' 100 1.85%.	2.7' 46% 0.69%.	9.0' 14% 0.17%.	85.9' 1% 0.01%.			
t>100ms	5.0' 24% 1.47%.	12.5' 10% 0.52%.	40.1' 3% 0.13%.	0.01%.			
t>250ms	13.1' 9% 1.17%.	28.6' 4% 0.41%.	2.0H 1% 0.09%.	0.01%.			
t>500ms	23.1' 5% 0.97%.	50.1' 2% 0.33%.	3.3H 1% 0.07%.				
t>1s	37.5' 3% 0.77%.	2.5H 1% 0.19%.	0.03%.				
t>2s	2.5H 1% 0.35%.	0.07%.					

NOMBRE TOTAL DES ECHOS :510
DUREE TOTALE DES MESURES : 10 H 1 mn 4 s
POLARISATION CROISEE

	>-120dBm	>-115dBm	>-110dBm	>-105dBm	>-100dBm	>-95dBm	>-90dBm
t>0ms	0.2' 100 14.39%.	0.3' 63% 6.72%.	0.6' 33% 3.12%.	1.4' 15% 1.33%.	3.1' 7% 0.51%.	7.4' 3% 0.15%.	42.1' 1% 0.02%.
t>100ms	0.5' 40% 12.23%.	1.1' 19% 5.33%.	2.2' 9% 2.41%.	5.7' 4% 0.98%.	11.5' 2% 0.37%.	31.6' 1% 0.09%.	0.01%.
t>250ms	1.4' 16% 9.29%.	3.0' 7% 3.85%.	6.5' 3% 1.68%.	14.0' 2% 0.71%.	25.2' 1% 0.25%.	0.02%.	
t>500ms	3.4' 6% 6.78%.	7.2' 3% 2.75%.	13.3' 2% 1.24%.	28.0' 1% 0.49%.	0.04%.		
t>1s	6.4' 3% 5.20%.	14.8' 1% 1.91%.	42.1' 1% 0.62%.	2.1H 0.14%.			
t>2s	14.0' 2% 3.27%.	63.1' 0.59%.	0.15%.				

NOMBRE TOTAL DES ECHOS :1217
DUREE TOTALE DES MESURES : 4 H 12 mn 32 s
POLARISATION HORIZONTALE

Tableaux 8

	N	Dc	Pm	γ_n	γ_d
53 MHz	12,6	1,15	18	- 1,9	- 2,0
40 MHz	42 '	4,51	76	- 1,5	- 1,9
Rapport	3,3	3,9			

Tableau 9

Temps d'attente ' Temps d'ouverture % Nombre de traînées %

	>-120dBm	>-115dBm	>-110dBm	>-105dBm	>-100dBm	>-95dBm	>-90dBm
t>0ms	0.3' 100 11.05%.	0.4' 77% 5.60%.	0.8' 37% 2.01%.	2.4' 13% 0.72%.	8.0' 4% 0.22%.	56.0' 1% 0.04%.	
t>100ms	0.7' 45% 9.24%.	1.4' 21% 4.48%.	3.2' 10% 1.57%.	16.0' 2% 0.51%.	37.3' 1% 0.17%.	0.03%.	
t>250ms	2.1' 15% 6.81%.	3.6' 9% 3.38%.	11.2' 3% 1.01%.	28.0' 1% 0.44%.	37.3' 1% 0.17%.		
t>500ms	4.3' 7% 5.50%.	8.0' 4% 2.50%.	18.6' 2% 0.79%.	37.3' 1% 0.40%.			
t>1s	8.6' 4% 4.02%.	16.0' 2% 1.69%.	0.30%.	0.21%.			
t>2s	22.3' 1% 2.26%.	55.9' 1% 0.74%.					

40 MHz

	>-120dBm	>-115dBm	>-110dBm	>-105dBm	>-100dBm	>-95dBm	>-90dBm
t>0ms	0.5' 100 4.84%.	1.3' 41% 1.47%.	3.4' 16% 0.44%.	12.2' 4% 0.14%.	33.0' 2% 0.04%.	3.9M 0.00%.	0.00%.
t>100ms	1.5' 35% 4.01%.	4.7' 11% 1.15%.	14.9' 4% 0.31%.	46.2' 1% 0.11%.	92.5' 1% 0.04%.		
t>250ms	3.6' 15% 3.02%.	11.8' 5% 0.82%.	51.4' 1% 0.17%.	1.9H 0.07%.	0.02%.		
t>500ms	9.0' 6% 2.04%.	27.2' 2% 0.57%.	2.6H 0.10%.	0.04%.			
t>1s	22.0' 2% 1.34%.	1.9H 0.22%.	0.05%.	0.04%.			
t>2s	77.0' 1% 0.59%.						

53 MHz

Tableaux 10

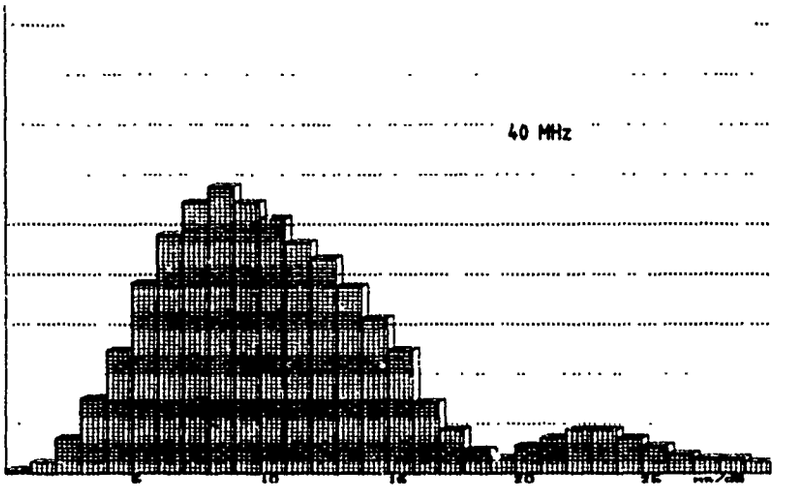


Figure 14(a)

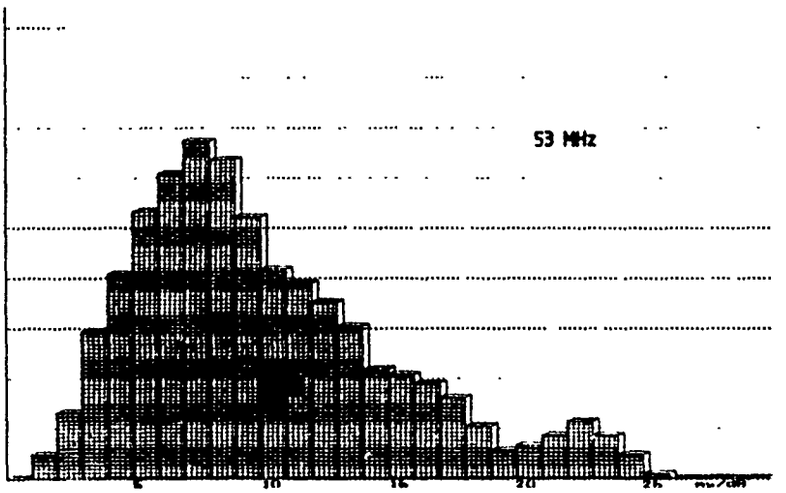
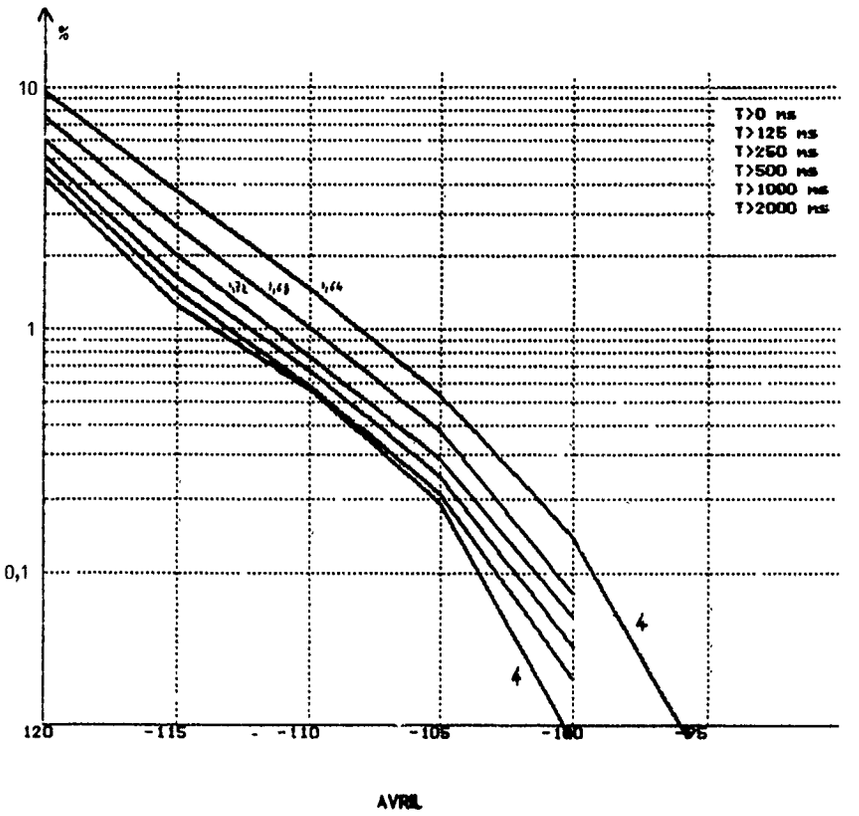


Figure 14(b)

Nombre de traînées en fonction de la constante de temps (ms/dB)



AVRIL

Figure 15 Temps d'ouverture pour chaque seuil en pourcent

EXPERIMENTAL INVESTIGATION OF METEOR BURST FOOTPRINTS

by
M. Rich, P. Heilman, M. Murray, B. Yetso
SRI International
333 Ravenswood Avenue
Menlo Park, California 94025
United States

1.0 **SUMMARY.** Meteor Burst Communications (MBC) provides reliable, beyond-line-of-sight data communications, but some serious throughput limitations are associated with the technology. One limitation is the finite waiting time required to send a message. This waiting time can be reduced by increasing the excess link margin, but this increases the cost and complexity of an MBC system. However, because of the limited footprint of a meteor trail, the waiting time for a system of networked stations can be reduced by spreading the stations over multiple footprints. For example, an array with station spacings of 50 km demonstrated less than one-tenth the waiting time of a single link. SRI International of Menlo Park, California, designed and operated an MBC array in the north-central United States to study meteor burst spatial diversity. This paper describes the effects of spatial diversity on link throughput and describes preliminary observations of signal footprint size and shape.

2.0 **INTRODUCTION.** The throughput and waiting time of an MBC system can be improved by using the spatial diversity of a communications network. Many MBC systems use a single "master" station to communicate with numerous "remote" stations. If these remote stations are emplaced sufficiently far apart to insure that they receive signals from different meteor trails, and if these remote stations are provided with some continuous channel to other remote stations, the network of remote stations will appear to have multiple parallel communication channels to the master station. The number of parallel channels is the diversity of the system and is related to both the size of the array and the size of the meteor burst signal footprint.

To incorporate spatial diversity into the design of an MBC system, the size of the signal footprint must be known. SRI International was contracted by GE Aerospace and Defense to design and operate an experiment to characterize the meteor scatter footprint and the effects of link orientation, polarization, and distance on footprint size and shape. SRI designed and operated a system for five months between December 1988 and April 1989. Numerous test configurations were used throughout the test and over 600 Mbytes of burst data were collected. Following completion of the test program, SRI developed software to analyze the footprint characteristics. This paper presents a preliminary analysis of the data collected on two links with a 625-mile range. The primary goal of this preliminary effort was to develop and test the analysis techniques. The secondary goals were to determine the amount of diurnal variation in the diversity results and the influence of the mixture of underdense and overdense trails on the footprint size.

3.0 **THEORETICAL JUSTIFICATION.** Meteor Burst Communications (MBC) is based on the reflection of signals from an ionized meteor trail to obtain propagation beyond the horizon. The nature of the reflecting surface causes the energy to be preferentially scattered to a particular area. This instantaneous signal "footprint" consists of a two-dimensional distribution of reflected power and is confined to a region where the received power level is sufficient for transferring data. Individual footprints vary in shape but are theoretically bounded by spherical cuts to a conic section, commonly described as asymmetric "boomerangs" [1].

The size and shape of the footprint depend on a number of factors, including carrier frequency, meteor radiant, transmitter power level, trail electron line density, path length, ambient noise level, and signal polarization. Footprint size and shape are statistical in nature because meteor radiants and trail electron line densities are statistical phenomena.

Using more than one receiver in an interconnected receiving system can improve the throughput of an MBC system. The assumption is that any signal received by at least one receiver is then received by all, because of connectivity by means other than meteor propagation. The time to achieve this connectivity is significantly less than the time to receive a message over the meteor communications system. For example, if all receivers are situated in a single footprint, they can be considered one receiver by the meteor system, and no throughput improvement is achieved. If, instead, all receivers are in different footprints, each received message is independent and the throughput is increased by a factor equal to the number of receivers. A typical system will emplace receivers somewhere between these extremes.

Significant difficulties arise in measuring an explicit footprint shape for communications systems planning. Because of the variations in shape and size, measurement of an individual footprint does not provide in enough information to determine an expected or average footprint. Also, since footprints represent large, complex shapes and are seldom aligned with a grid of measuring receivers, detailed measurement requires a large, finely graded array.

One approach to measuring the expected or average footprint is to use a time average of the power received at individual receivers within an array. Analyses have been performed to simulate the scattered signal at various link distances by using statistically distributed meteor radiants. Results included a two-dimensional probability of intercept density function based on the probability of successful message reception at a point some distance from a central receiver. The expected footprint lies within the 50% probability boundary of the distribution. For a transmitter at 800 miles, we expected this region to be 20 to 30 miles wide by 50 to 100 miles long. Additionally, the conditional probability between adjacent footprints was 30 to 40% [1,2].

Given a statistically large number of meteor events that are "well aligned" with the receiver grid, the expected footprint is somewhat easier to measure than the individual footprints. Raw data are analyzed to determine the probability of

accurate and complete signal reception at each site, presupposing that the central receiver (reference point) correctly accepted the message. A fine spacing of receivers is required where the change rate of probability with respect to distance is large (near the center). A more coarse spacing can be used in the fringes of the receiver array, where the derivative is smaller. The array must be at least as large as the required footprint, and preferably twice its size, to determine the conditional dependence of adjacent footprints.

Diversity Calculation. The existence of spatial diversity is due to the finite extent of a meteor-scattered signal on the ground. Spatial diversity for a network of interconnected receivers, then, is given by the ratio of the number of independent signals received by all the array to the number of signals received by a "standard" receiver.

There are two ways to calculate spatial diversity. First, since diversity is caused by a limited footprint area, diversity can be measured by the number of footprints that cover a network of receivers. Adjacent footprints must overlap, because theoretical footprints are infinitely large, having no sharp edges and decreasing with distance from the center. The amount of overlap determines the conditional probability of coincidence between adjacent footprints. To calculate diversity based on the number of footprints necessary to cover a network of receivers, the size of the receiver network, the size of the footprint, and the amount of footprint overlap or conditional probability between adjacent footprints must be known.

A second method of spatial diversity calculation realizes that increased network diversity causes an increase in the number of signals received by the network in a given time interval. A diversity factor can be calculated as the ratio of the number of signals received by the array to the number of signals received by an array element (the "standard" receiver). The number of signals received by an array element could be

- Average number of signals received by all sites in the array
- Number of signals received at a single site
- Number of signals received by a small, compact group of sites.

The average number of signals received by the array is not an accurate reference if the expected number of signals received at each site varies significantly, as occurs if the size of the array is a significant fraction of the distance from the transmitter to the array. The case of a single site as the array element is not generally a good reference, because the diversity calculation is too dependent on the performance or proper operation of this single site. Thus by choosing a small, compact group of sites as the array element of reference, spatial diversity is most accurately determined.

An increase in network diversity decreases the waiting time to inject a message into the network. Once the diversity is known, along with the required confidence factor and adjacent footprint conditional probability, the wait-time reduction factor can be calculated.

The distribution of meteor trails is a statistical counting problem. The average number of trails detected per unit time is determined by the link design and the excess power available beyond that required to close the link. The waiting time is distributed exponentially, because the messages arrive according to a Poisson process. For a given link and a specified amount of time, a message will be received with some confidence probability; for a higher level of confidence, more time is required. From queueing theory, [3] the relationship between wait time at confidence P (WT_p), throughput, and message length is

$$\frac{(WTp)(throughput)}{(message\ length)} = -\ln(1-P) \quad . \qquad (1)$$

With Event 1 (E_1) as the reception of a message by a receiver in Footprint 1, and Event 2 (E_2) as the reception of a message by a receiver in an adjacent footprint, Footprint 2, the probability of reception by either receiver is

$$P(E_1 \text{ or } E_2) = P(E_1) + P(E_2) - P(E_1 \text{ and } E_2) \quad , \qquad (2)$$

where $P(E_1 \text{ or } E_2)$ is the conditional dependence between adjacent footprints, $P(E_1)$ and $P(E_2)$ are the single link probabilities, and $P(E_1 \text{ and } E_2)$ is the confidence for a 2-diverse system having $P(E_1)$ single-link confidence. Since the N-diverse confidence is the reliability design parameter, $P(E_1)$ must be calculated and used for single-link simulation and evaluation. $P(E_1)$ will always be less than the N-diverse confidence when the conditional dependence is less than 1. Extrapolating to an N-diverse system shows

$$P_{X,N} = \frac{P_N}{N - (N-1)P_X} \quad . \qquad (3)$$

where $P_{X,N}$ is the required single-link confidence, P_N is the confidence for a network with diversity N, and P_X is the adjacent footprint conditional dependence.

For P_X less than 1, $P_{X,N}$ will be less than P_N and the wait time calculated in Eq. (1) will be reduced. For an N-diverse network, the confidence P in Eq. (1) is $P_{X,N}$. If $WT_{p,1}$ is the wait time for a single link and $WT_{p,N}$ is the wait time for an N-diverse network, the waiting time improvement factor, f(N), for an N-diverse network with constant throughput can be defined as

$$f(N) = \frac{(WT_{p,N})}{(WT_{p,1})} \quad . \qquad (4)$$

Combing Eq. (1) and Eq. (4) to determine the waiting time reduction factor given the single link and network confidence factors results in

$$f(N) = \frac{\ln(1 - P_{X,N})}{\ln(1 - P_N)} \quad , \qquad (5)$$

with $P_{X,N}$ and P_N related by Eq. (3). A plot of waiting time improvement factor as a function of level of diversity is shown in Figure 1.

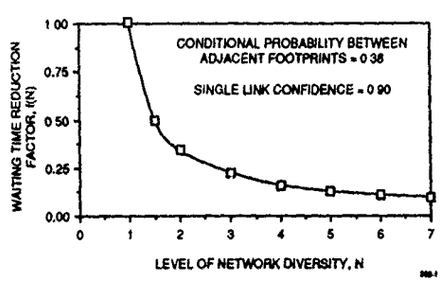


FIGURE 1 EFFECT OF SPATIAL DIVERSITY ON WAITING TIME

4.0 **EXPERIMENT DESCRIPTION.** To determine the spatial diversity of a networked array, SRI International installed a meteor burst receiver array centered in Lewistown, Montana. The array consisted of 33 receivers located in concentric hexagons about the central site: 4 receivers were in a 5-mile radius from the center, 6 receivers at 15 miles, 6 receivers at 30 miles, 12 receivers at 60 miles, and 4 receivers at 100 miles from the array central receiver. Figure 2 is a geographical approximation of the site locations. (Rings are numbered sequentially, not all sites are shown)

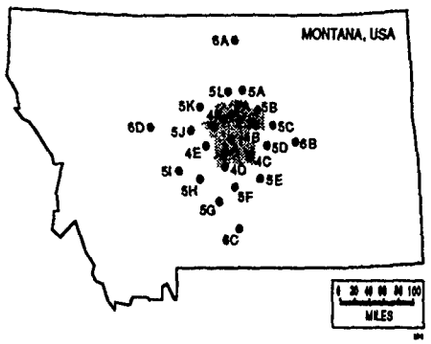


FIGURE 2 SPATIAL DIVERSITY RECEIVER ARRAY (Central Montana)

Installed at each array site was a two-channel horizontal receiver and a two-channel vertical receiver. Frequency A (48.858 MHz) and Frequency C (46.959 MHz) corresponded to the vertical receiver's Channels 1 and 2, respectively. Frequency B (46.759 MHz) and Frequency D (48.758 MHz) corresponded to the horizontal receiver's Channels 1 and 2, respectively. The receiver bandwidth was 9 kHz, and the receiver system normally used horizontally and vertically polarized antennas with gains of 5.3 dBi and 4.5 dBi, respectively.

The receivers had 256 kbytes of on-board memory, providing three to ten days of data logging depending on the system configuration and meteor activity. They had an operational battery life of approximately ten days. ASCII data were downloaded every three to seven days from each of the 33 receiver sites and stored on 3.5-in. floppy diskettes. The data were later transferred to 5.25-in. diskettes, 9-track tape, and 8-mm tape cartridges for analysis using an MS-DOS personal computer. A total of 62 days of valid data were collected throughout the experiment, which was performed from 4 December 1988 through 7 April 1989.

The experiment consisted of four major test configurations. A total of four 500-W transmitters were used: one pair moving west-to-east from the center of the array, and the second pair moving north-to-south from the center of the array. Each pair of transmitters operated with two Yagi antennas (11 dBi of gain), one vertically polarized and one horizontally polarized. The antennas were mounted on 20-ft carbon masts for the first three tests and a 45-ft steel foldover tower for the last test. For the first test, the two transmitter pairs were colocated in Tekoa, Washington, 400 miles to the west of the central receiver site. After this test, the transmitters were configured for simultaneous north-to-south and east-to-west test operation, from sites 175, 400, and 625 miles from the center of the array. A fifth transmitter was located on Judith Peak, near Lewistown, Montana. This transmitter ensured the accuracy of the receiver time-keeping equipment by broadcasting a 2-s burst from the center of the array twice per hour.

Each 100-bit transmitted message comprised a 60-bit header, 32-bit transmit time code indicating time of transmission, and an 8-bit cyclical redundancy checksum (CRC). The correlation receiver allowed message detection at a 5.2-dB signal-to-noise ratio, with a message false-alarm probability of 10^{-7}. The 100-bit, FSK-modulated transmit signal was broadcast every 40 ms using a baud rate of 2.5 kbps. The system logged the signal power and correlation level every 40 ms during the burst and logged calibration data (ambient noise level, temperature, and battery voltage) twice per hour.

5.0 **INDIVIDUAL SITE ANALYSIS.** The data collected at each receiver site provided a chronology of the meteor events and receiver noise measurements since the previous data collection. Data processing began on the field data by filtering out erroneous data caused by an improper CRC, abnormal background noise, or data collection problems. Regression analysis was performed on the data to correct transmitter timebase problems, to correct for drift in the receiver clocks, and to determine the arrival time of bursts without message numbers. This initial phase of processing created daily logs of meteor arrival times and lengths.

The intent of this experiment was to study the meteor trail, a physical process, rather than the communications resulting from the trail occurrence, an artificial process influenced by the equipment and experimental design parameters. Two induced problems were readily apparent. First, the data collected by the receiver had a duration based on a signal power level measured in the receiver. The threshold level was inconsistent throughout the array, thus, the burst duration was corrected by analyzing the received signal-to-noise ratio. Background noise was measured by the calibration system every half hour. The duration of a burst was defined as the signal power sufficient to provide fewer than 9 bit errors in the 56-bit header of each message. Second, typical meteor burst communication systems exhibit multipath fading of the received signal. Fading is caused by a variety of natural phenomena perturbing the meteor trail. Data analysis was needed to join signal events that seemed to be caused by individual trails because such fades would indicate more and shorter trails than actually occurred. Thus, the next step in the data processing sequence was to identify and eliminate multipath fading effects from the meteor logs.

To analyze the signal fades, a histogram of the fade durations was created. We hypothesized that two processes could cause a gap in the signal history: signal fading, which we were attempting to remove, and the expected random time between successive meteor trails. As these two processes have greatly different means, they will appear as two different slopes in the histogram. The sample histogram illustrated in Figure 3 clearly shows two different distributions. To separate multipath fading gaps from true meteor arrival gaps, gap statistics from one whole day were gathered and the first 50 points were fitted to a power curve ($y=Ax^{-b}$). From the slope of the curve, the cumulative 90%

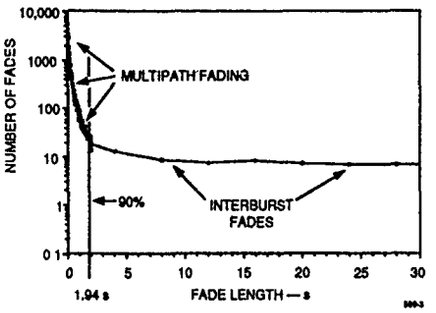


FIGURE 3 HISTOGRAM OF FADE LENGTHS (DAY 081, FREQUENCY C)

point was determined; i.e., the gap size that 90% of the gaps were less than. The 90% values ranged from 0.704 s to 4.216 s. Figure 3 shows a 90% point of 1.94 s. We averaged the 90% values from six days, representing the various test configurations, to obtain a value of 1.6 s. Gaps of 1.6 s or less in the meteor log files were considered to be multipath fading from a single trail and were joined. Gaps greater than 1.6 s were considered separate meteor trails This gap size was chosen as a good tradeoff between joining too many independent trails and breaking too many trails into multiple events.

Once the field data were corrected for time errors, duration errors, and signal fades, the meteor log files were assembled and processed in 4-hr blocks synchronous with Greenwich Mean Time (GMT). This process involved counting the number of trails of length greater than or equal to 40, 120, and 200 ms. Site data were recorded for all four receiver channels. An average number of trails detected by an individual site was calculated, ignoring any sites that were not operating.

Event Coincidence. The next step in the data processing sequence was to correlate the meteor events between the various sites within the array. A large linked list of site data was formed and searched for independent trails. Independent trails were identified by finding gaps in time, which marked the boundary from one meteor trail to the next. For each 4-hr period, the number of independent trails seen by the array, excluding the 100-mile sites, was tabulated. (The 100-mile sites were only used for footprint evaluation, and not for diversity calculations, because of their sparse positioning.) The diversity of the array was calculated by dividing the number of independent trails seen by the entire array by the average number of trails detected by individual receivers within the array. Diversities of 8 to 12 were typical.

Additional processing was performed to refine the diversity measurement. Meteor reflections have historically been categorized in two ways, as reflections from underdense trails or as reflections from overdense trails. Each type of reflection has different scattering characteristics and therefore different footprint sizes. The data set was divided into overdense and underdense categories to allow separate analysis of these two trail types.

Underdense trails are defined as trails whose ionization density is low enough to allow an incident signal to pass through the trail without major perturbation [2]. Scattering occurs from the individual electrons, and the scattered signal is very coherent. The good coherency allows the signal to be evenly spread in space, creating a moderate-sized footprint without major fades.

Overdense trails have a sufficiently high density of electrons to reflect the incident signal from a large scattering surface. Signals reflecting from different parts of the surface can combine destructively at some locations. The large reflection surface and greater reflected power cause the signal to be scattered over a large area, with fades in some regions of the footprint. The large ionization density also causes the trail to last longer, as the trail duration depends on diffusion of the ionized particles. Upper atmosphere winds cause the trail to break up and drift across the receiver array, creating dynamic footprint effects. Even though the signal from an overdense trail is only correlated over a small area and many fade cells exist, there are many areas of correlation and these areas move throughout the trail lifetime. Thus, an overdense trail becomes a large-area effect, having a much larger footprint over its lifetime than does an underdense trail.

An accurate model of a networked MBC system would have to take the two types of scattering into account because the footprints from underdense and overdense trails differ greatly . Similarly, design of networked MBC systems will have to consider the mix of underdense and overdense trails, depending on the link design. To determine a mix of these two trail types in our test links, a procedure to separate signals from the two types of trails was developed.

Two methods of dividing the data into overdense and underdense sets were investigated. The more obvious approach was to divide the trails based on trail duration. However, our plots of trail duration contained no clues as to how to differentiate overdense and underdense trails. Many factors can influence the duration of the trail besides the ionization density; e.g., trail orientation, fade parameters, trail location with respect to the receiving array, and the receiving equipment. An alternate approach was to divide the data based on the number of sites that detected a given trail. We called this the visibility of the trail. We considered the visibility of a trail to be a better measure of the trail ionization density than trail duration because the signal footprint is determined by the scattering process. However, if the signal footprint could be used to divide the signal records into underdense and overdense data sets, the resulting histogram of trail durations could be used as a verification of the accuracy of the set division.

The visibility of a trail was analyzed in three ways. The first and simplest was to look at the number of sites that received signal energy at each 40-ms time sample without regard to trail boundaries. This *individual visibility* measurement heavily weighted the longer trails, because the longer trails have more samples. However, this technique was not a good measure of the coverage area of individual trails. The second method, *instantaneous visibility*, counted the maximum number of sites that together detected signal energy at the same time during a trail's duration. This was a better measure of the coverage of an individual trail, but only at a single point in time, and it ignored any trail drift effects. The third technique, *total visibility*, counted the number of sites detecting trail energy throughout a trail's duration, providing a measurement of the total, rather than instantaneous, coverage of the trail. Histograms of the total visibility of the trails showed greater dispersion of the overdense trails and enabled a more accurate division of the underdense and overdense data sets.

Histogram data used to determine the division between underdense and overdense events were based on the central site, the 30-mile sites, and the 60-mile sites only. These sites represented a regular hexagonal array without the bias of the additional short-range sites. The histograms showed two slopes that converged between four- and five-site detection of the meteor trail. Figure 4 shows histograms of total visibility for two 4-hr periods. The two slopes are related to the variance in site coverage probability. The underdense trails, having a smaller footprint than the overdense trails, can be seen in the histogram region showing detection by one to four sites. Overdense trails produce the slope for detection by five and more sites.

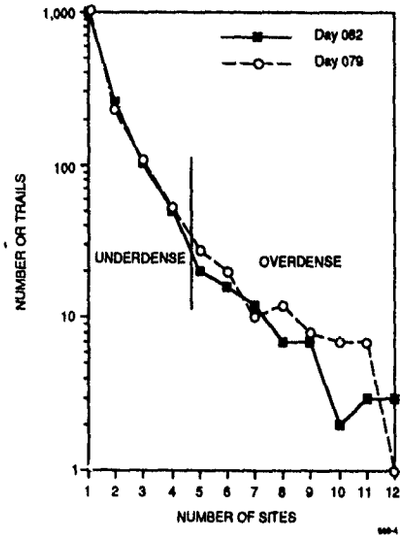


FIGURE 4 TOTAL VISIBILITY HISTOGRAM (FREQUENCY C, 1600-2000 GMT)

The data files were reassembled and the linked list was pruned of overdense trails based on total visibility by more than four sites. Underdense diversities, visibility histograms, and trail duration histograms were recalculated for each 4-hr data block. This process then was repeated appropriately to prune underdense trails from the linked list. As expected, the burst duration histograms indicated that most of the underdense activity was confined to the shorter trails. Most of the longer trails seen in the underdense category were trails that illuminated only the edges of the test array. Typically, less than 5% of the trails in the underdense category exceeded 3 s in duration, and less than 10% of the trails in the overdense category were less than 1 s in duration.

Footprint Sizing. Another calculation made during the event coincidence processing was used to evaluate the actual size and shape of the average footprint. We have previously defined the footprint as a probability distribution surface; for example, the probability of detecting a signal at some point away from a signal-detecting "reference receiver." We used the five central receivers as our reference, and considered a reference signal to be a signal detected at any of these sites. For each site in the array (including the 100-mile sites), we estimated a conditional signal detection probability, B|A. The B|A probability at each site was the ratio of the number of signals that both the site in question and the reference site detected to the total number of signals detected at the reference site.

The extensive footprint analysis presented herein is based on data from Julian Days 077 through 083, when the receiver array was performing most efficiently. One transmitter pair was located 625 miles south of the array, and the second transmitter pair was located 625 miles east of the array. First, the receiver array was divided into six 60° segments, with each receiver site allocated to one of the six segments. Each segment contained between four and six sites. The B|A probabilities for each site were averaged over the six days. A logarithmic curve was fit to a plot of the resulting data points as a function of distance from the center of the array. Analysis involved determining the m and b parameters of the exponential curve of regression of the form y=b*m^x for each sector. Five distances corresponding to probabilities of 10%, 25%, 50%, 75% and 90% were interpolated from each curve. From these parameters, five footprints were determined for three conditions: underdense meteor trails, overdense meteor trails, and unseparated data containing both underdense and overdense trails. The footprint surface was plotted by curve fitting the five probability points in each sector.

The data for the horizontally polarized south-to-north link was used to evaluate the footprint shapes. Figure 5 depicts footprints for composite overdense and underdense meteor events. The plot uses a linear scale, with two concentric circles indicating 10- and 30-mile distances from the center of the array. The 90% probability footprint is closest to the center of the array, while the 10% probability is farthest from the center of the array. True north is along the 0° axis. Figure 6 depicts footprints for overdense meteor trails only. The plot uses a linear scale, with two concentric circles indicating 40- and 80-mile distances from the center of the array. Figure 7 shows footprints for underdense meteor trails only. The logarithmic-compressed scale for the underdense footprints corresponds to 100-, 10-, 1-, and 0.1-mile distances.

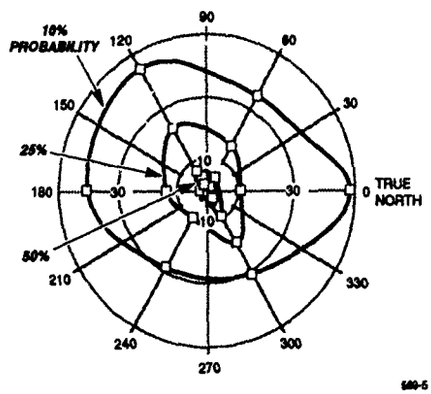


FIGURE 5 FOOTPRINTS FOR COMPOSITE OVERDENSE AND UNDERDENSE METEOR TRAILS (F4, DAYS 077-082)

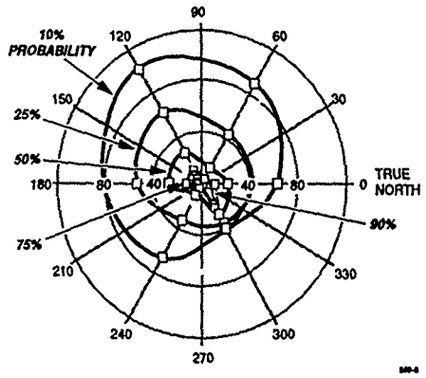


FIGURE 6 FOOTPRINTS FOR OVERDENSE METEOR TRAILS (F4, DAYS 077-082)

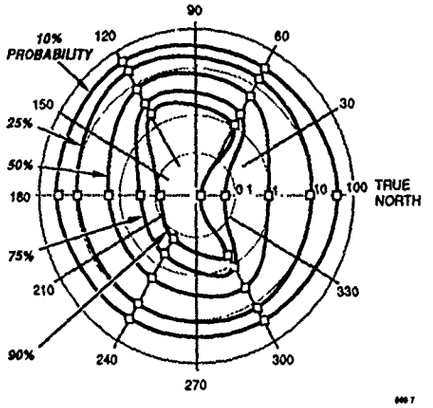


FIGURE 7 FOOTPRINTS FOR UNDERDENSE METEOR TRAILS (F4, DAYS 073-082)

6.0 **OBSERVATIONS.** The data analysis for this paper is based on six days near the end of the test, when the transmitters were located 625 miles from center of the receiver array. Data analysis concentrated on the south-to-north link because of equipment problems with the east-to-west link discovered during data processing. Diversity was analyzed for all frequencies and all time blocks in this interval. An average footprint analysis was performed on all frequencies during this interval, but the footprint surface was only determined for the south-to-north horizontally polarized link. Selected individual trails were also reviewed. Our intent was to develop data analysis techniques that could be used to analyze all the data. These six days represented some of the most consistent data collected and the most common link configuration. Further processing, when time and resources permit, will complete the analysis for the remaining data.

Data for each 4-hr time block were averaged over the six days. Table 1 summarizes overdense and underdense event diversity, listing, for vertically and horizontally polarized links, the average number of trails detected at each site, the number of independent trails detected at each site, and the diversity of the array. Time plots of the south-to-north array diversity are shown in Figure 8. The 4-hr diversity ranged from 10.3 to 11.5 for the vertically polarized link, and from 11.3 to 12.9 for the horizontally polarized link. The vertically polarized link showed 1.6 times the individual site activity of the horizontal link, but this was determined to be predominantly a frequency effect, as the vertical link was operating 1.8 MHz below the horizontal link. (The east-to-west link, with the frequencies reversed, showed the opposite effect.) Horizontal link diversity was typically 10% higher than that of the vertical link. This difference held for both link directions and did not seem to be a frequency effect.

Table 1
UNDERDENSE AND OVERDENSE BURST DIVERSITY

Time Block (GMT)	Average Trails Detected / Site	Independent Trails Detected by the Array	Array Diversity
VERTICALLY POLARIZED LINKS			
0 to 4	31.3	355	11.5
4 to 8	75.3	866	11.5
8 to 12	155.2	1605	10.4
12 to 16	175.8	1836	10.5
16 to 20	142.9	1631	11.4
20 to 24	73.5	841	11.5
Average			11.1
HORIZONTALLY POLARIZED LINKS			
0 to 4	23.1	264	11.4
4 to 8	45.7	549	12.0
8 to 12	92.9	1088	11.7
12 to 16	105.9	1243	11.7
16 to 20	84.0	1063	12.6
20 to 24	48.4	608	12.5
Average			12.0

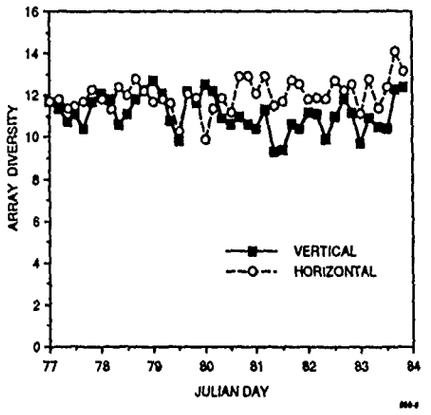


FIGURE 8 AVERAGE SPATIAL DIVERSITY

The receiving array was designed to accurately determine the extent of the footprint without an excessive number of sites. If we sized the array correctly, the diversity should be approximately half the number of receivers in the array. This ratio would provide an average of two receivers in each signal footprint. Of the 33 receivers in the array, 19 were in a regular hexagonal pattern. Although this number of receivers is sufficient for determining the diversity, a slightly smaller and more regular interstitial spacing would have yielded more accurate results.

As expected, a great deal of diurnal variation occurred in the number of bursts at a site: Peak interval activity was about 5.5 times the quiet interval activity. Weak diurnal activity was detected in the diversity calculation. Figure 9 shows a diurnal peak for both site activity and array diversity. The diversity spectrum also shows a harmonic of the diurnal peak. Diversity and diurnal variations were analyzed further by dividing the trails into underdense and overdense sets. The diurnal variation in the ratio of underdense to overdense trails would imply that the two types of trails had different formation processes, suggesting a diurnal variation to the array diversity.

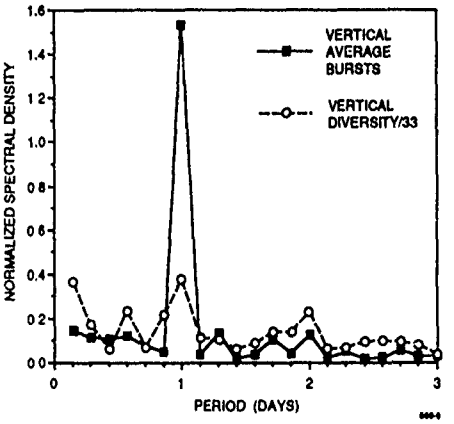


FIGURE 9 COMPARISON OF AVERAGE NUMBER OF BURSTS AND DIVERSITY

Tables 2 and 3 summarize diversity for underdense and overdense events, respectively, listing, for vertically and horizontally polarized links, the average number of trails detected at each site, the number of independent trails detected by the array, and the diversity of the array. Array diversity is strongly influenced by the underdense meteor trails. For the south-to-north link, there were 6.6 times as many underdense trails as overdense trails for the vertically polarized link, and 5.3 times as many for the horizontally polarized link. However, because of the greater duration of the overdense trails, the ratio of total underdense trail duration to total overdense trail duration was less: 3.1 on the vertically polarized link and 4.4 on the horizontally polarized link. The diurnal variation of the ratio of number of trails was imperceptible, discounting the existence of separate formation processes, while the ratio of total trail duration showed significant variation. The ratio of total underdense trail duration to total overdense trail duration was greatest in the afternoon, when trail activity is at a minimum. The variation in total trail duration is probably inflated by the small number of overdense trails identified during these intervals. The ratio of underdense to overdense trail diversity was 6.8 for the vertically polarized link and 7.2 for the horizontally polarized link. The ratio of independent events is nearly the same for the horizontally and the vertically polarized links: 37.9 and 39.3, respectively. While the lack of diurnal variation in these data yields little insight into the formation of trails, it greatly simplifies the data processing, enabling extraction of effects based on frequency, link orientation, and signal polarization and improving the overall quality of the results. Part of the reason that a large volume of data was collected was because diurnal variations were considered a possibility. In particular, the limited diurnal variation in the diversity enabled us to combine the probability data, thereby improving our estimate of the footprint size.

Table 2
UNDERDENSE EVENT DIVERSITY

Time Block (GMT)	Average Trails Detected / Site	Independent Trails Detected by the Array	Array Diversity
VERTICALLY POLARIZED LINKS			
0 to 4	26.7	346	13.1
4 to 8	66.5	847	12.8
8 to 12	131.6	1555	11.9
12 to 16	146.0	1780	12.3
16 to 20	116.3	1583	13.7
20 to 24	61.8	815	13.3
Average			12.8
HORIZONTALLY POLARIZED LINKS			
0 to 4	18.8	255	13.5
4 to 8	38.4	533	13.8
8 to 12	79.0	1060	13.4
12 to 16	86.5	1170	13.5
16 to 20	68.8	1033	15.0
20 to 24	40.3	592	14.6
Average			14.0

Table 3
OVERDENSE EVENT DIVERSITY

Time Block (GMT)	Average Trails Detected / Site	Independent Trails Detected by the Array	Array Diversity
VERTICALLY POLARIZED LINKS			
0 to 4	5.0	10	1.7
4 to 8	9.8	19	2.0
8 to 12	26.0	50	1.9
12 to 16	30.3	56	1.8
16 to 20	27.1	49	1.8
20 to 24	11.8	26	2.2
Average			1.9
HORIZONTALLY POLARIZED LINKS			
0 to 4	4.3	9	1.9
4 to 8	7.4	16	2.1
8 to 12	14.2	28	2.0
12 to 16	19.6	37	1.9
16 to 20	15.3	31	2.0
20 to 24	8.2	17	2.0
Average			2.0

The diversity evaluations generated a single numerical representation of the footprint size. This representation is a weak function of the performance at any individual site. The footprint area analysis depends strongly on site locations and individual site operations, because the footprint area evaluation required the calculation of 30 plot points. Prior to plotting the footprints, the underlying data were analyzed to find trends that supported the footprint size and shape hypothesis. The data with the best fit to the logarithmic approximation curve were in the sectors with the greatest dispersion of receiver sites. As anticipated,

the south-to-north link generally had more dispersion of signal energy in the north and south sectors than did the east-to-west link. The corollary was true for the east-to-west link. This energy distribution is not as well demonstrated in the logarithmic polar plots used for the underdense trail footprint. The log radial scale was used to provide more definition to the central probability rings. Few sites had conditional probabilities in excess of 0.5; therefore, the probability curves greater than 0.5 have less confidence than those below 0.5. The use of a master site consisting of multiple central sites reduced the sensitivity of the calculation to a single site yet increased the dispersion of the probability surface. Data previously processed with a single master site showed both greater conditional probabilities near the center of the array and smaller conditional probabilities near the array edges.

The area of the regular array was 10,221 square miles. The average diversity of the south-to-north horizontally polarized link was 12. The average footprint area, based on the diversity, was the ratio of these two, or 852 square miles. In defining the footprint size on the basis of the footprint plots, the basic problem is defining the appropriate conditional probability. On the logarithmic plots, the conditional probability rings are roughly circular. If we expect a footprint size of 852 square miles, a circular footprint radius would be 16.5 miles, corresponding to 20% conditional probability on the underdense and combined footprints, and to 65% probability on the overdense footprints. These values imply that the footprint surfaces are slightly smaller than the diversity calculations would indicate. There are two reasons for this. First, the use of multiple sites as a master site greatly increased the number of trail detections attributed to the master site, since a large portion of the trails detected by the array are seen by only one site. This reduced the conditional probabilities for all sites, including those sites used as part of the master site group, and therefore reduced the size of the footprint. Second, calculation of conditional probabilities was based on instantaneous coincidence rather than trail coincidence. Therefore, sites that detected a trail non-coincidentally with the master site were not included in the conditional set, thereby reducing the conditional probabilities and the footprint size. Together, these two factors increased the dispersion of the probabilities and probably made the footprints appear smaller and more circular than expected.

The goal of the data analysis to date has been to study the dominant influences on meteor burst spatial diversity so as to describe a first-order model of the meteor burst signal footprint. The data will allow study of numerous secondary effects when time and resources permit. Once all the data are processed consistently, we plan to decorrelate the effects of signal frequency, polarization, and link orientation with the magnetic field. The effects of signal frequency on diversity may be too weak to identify because of the closely spaced frequencies used in this test. The present data seem to indicate a 10% greater diversity for the horizontally polarized signals, but the influence of link orientation is too weak an effect to identify with the data analysis we have completed to date. The data collected at other link distances will be processed to determine the influence of link range on footprint size. And while we have identified little diurnal variation of footprint size or trail density makeup at the 625-mile link range, we will continue this investigation with the remaining data.

Further data analysis will be performed to study dynamic influences on diversity. Two dynamic processes seem to be present in the data but have not been quantified: the effects of trail formation and trail drift. By timing the leading edge of the detected bursts at the various array sites, we have identified a leading edge velocity of approximately 200 km/s, or roughly twice the expected trail formation speed. During longer bursts, we detected a much slower and wider area propagation at approximately the speed of the neutral winds (25 m/s). The array temporal sampling rate was 40 ms/sample, and the spatial sampling ranged from 4.2 to 334 km, with typical spacing of 60 km. Typical overdense trails crossed a portion of the array in 10 to 20 time samples, a sufficient amount to characterize the direction and speed of the trail. Future work will include development of data analysis techniques to characterize the velocities of these dynamic effects as well as to quantify the effects of these transient phenomena on array diversity.

ACKNOWLEDGMENTS

The authors of this paper would like to thank the people from Lewistown, Montana, who helped out in planning and operating the test: Bud Zormeir, J. C. Gilpatrick, Jeffrey Finn, Gerry Moline, Dale and John Ployhar, and Deana Taylor. We would also like to thank Paul Boose of GE Services and Barbara Camph and others within SRI International for their support of the program.

REFERENCES

1. Embry, U.R., Stanford Electronics Laboratories, Stanford, California, "The Probability of Intercepting Radio Signals Scattered by Meteor Trails, " March 1962, Technical Report No. 773-1.
2. Davies, K., "Ionospheric Radio Propagation, " National Bureau of Standards, April 1965, p. 365.
3 Kleinrock, L., "Queueing Systems, Volumes I and II," New York, John Wiley and Sons, 1975, 1976.

EIGENVECTOR WEIGHTING AS AN ADAPTIVE ARRAY INTERFERENCE CANCELLATION TECHNIQUE

R.W. Jenkins and K.W. Moreland
Communications Research Centre
Box 11490, Station H
Ottawa, Canada K2H 8S2

SUMMARY

Adaptive antenna techniques normally make use of known properties of the desired signal, such as its direction or an embedded code, to distinguish it from interference. When this is not possible, algorithms such as power ratio inversion or Gram-Schmidt are used. These techniques tend to invert the relative powers of signals arriving at the array, and thus are effective when the interference is substantially stronger than the desired signal. However, when signal levels are close, such techniques fail. This paper presents analytic and modelling studies of the eigenvector weighting technique. It is shown that this technique performs substantially better than power ratio inversion techniques. Like the power ratio inversion methods, eigenvector weighting is most effective when there is a large separation in signal powers. However, even when the signals are close in power, satisfactory cancellation can be achieved with eigenvector weighting for a large fraction of cases. The actual performance depends on array geometry and number of elements.

INTRODUCTION

In a crowded, possibly hostile, radio environment, communication is threatened by interference due to other users or enemy jammers. Many-user nets require a means of discriminating between concurrent users. Adaptive antenna arrays provide spatial discrimination against interference, and a means of separating the signals from multiple users in a net. These arrays normally employ techniques which use known features of a desired signal, such as a predetermined embedded code or a known direction, in order to distinguish it from other signals.

Predetermined identification techniques are not appropriate for some communications signals. Strategies which may be used in such cases include the power ratio inversion algorithm [1] and the Gram-Schmidt orthogonalization approach [2]. These techniques tend to invert the relative powers of signals incident on the array, yielding a desired signal at the output which is as much above the co-channel interference level as its input counterpart is below the interference.

Eigenvector weighting has been proposed as a means of separating independent incident signals [3]. Using eigenvectors of the covariance matrix of the input signals as weights to transform the array inputs yields a set of uncorrelated output signals. This lack of correlation is a property also of the independent incident signals, and it has been suggested that the output signals produced by eigenvector transformation consist mainly of the independent incident signals. There is a fallacy in this reasoning. There exist linear combinations of the incident signals which are uncorrelated There are also other transformations such as Gram-Schmidt orthogonalization which cause the outputs to be uncorrelated. Nonetheless, modelling and simulation studies [3,4] indicate good performance for the eigenvector technique when the signals of interest are well separated in power.

This paper presents analytic and modelling studies of the eigenvector technique. Its performance relative to power ratio inversion is considered.

SIGNAL FORMULATION

A complex baseband formulation is used in this analysis. Column vectors are indicated by $\tilde{}$, matrices, by bold type, and the complex conjugate, transpose, and conjugate transpose operations, by *, ^{T}, and ^{H}, respectively.

Consider m independent signals with baseband time dependence $s_j(t)$, $j=1,m$, incident on an n-element array. Let g_{ij} be the complex gain of the ith element of the array to the jth signal. The resultant signal at element i is given by

$$x_i = \sum_{j=1}^{m} g_{ij} s_j + n_i \qquad (1)$$

where n_i is the noise present on the ith channel. The array response vector $\tilde{g}_j$ to the jth signal is defined as

$$\tilde{g}_j = [\, g_{1j},\ g_{2j},\ \cdots,\ g_{nj} \,]^T \qquad (2)$$

and the n x m array response matrix **G** to the collection of signals, is likewise defined as

$$\mathbf{G} = \begin{bmatrix} g11 \cdots g1j \cdots g1m \\ \cdots \cdots \cdots \cdots \cdots \\ gi1 \cdots gij \cdots gim \\ \cdots \cdots \cdots \cdots \cdots \\ gn1 \cdots gnj \cdots gnm \end{bmatrix} = [\tilde{g}_1, \tilde{g}_2, \cdots, \tilde{g}_m] \quad (3)$$

Using the array response matrix, (1) can be expressed in vector form as

$$\tilde{x} = \mathbf{G}\,\tilde{s} + \tilde{n}\,, \text{ where} \quad (4)$$

$\tilde{x} = [x_1, x_2, \cdots, x_n]^T$, $\tilde{s} = [s_1, s_2, \cdots, s_m]^T$, and $\tilde{n} = [n_1, n_2, \cdots, n_n]^T$ are the array input, signal, and noise vectors respectively.

The covariance matrix **R** of the input signals is given by

$$\mathbf{R} = \begin{bmatrix} \langle x_1 x_1^* \rangle & \langle x_1 x_2^* \rangle & \cdots & \langle x_1 x_n^* \rangle \\ \langle x_2 x_1^* \rangle & \langle x_2 x_2^* \rangle & \cdots & \langle x_2 x_n^* \rangle \\ \cdots & \cdots & \cdots & \cdots \\ \langle x_n x_1^* \rangle & \langle x_n x_2^* \rangle & \cdots & \langle x_n x_n^* \rangle \end{bmatrix} = \langle \tilde{x}\,\tilde{x}^H \rangle \quad (5)$$

where $\langle\ \rangle$ represents the statistical average, which is approximated in practice by a time average. The input signals s_j, $j=1,m$ and channel noise signals n_i, $i=1,n$ are assumed to be independent and thus non-correlating. Using (4) and (5), the covariance matrix can be written as

$$\mathbf{R} = \mathbf{G}\mathbf{P}\mathbf{G}^H + \mathbf{N} \quad (6)$$

where $\mathbf{P} = \begin{bmatrix} P_1 & 0 & \cdots & 0 \\ 0 & P_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & P_m \end{bmatrix}$ is the matrix of the signal powers $P_j = \langle s_j s_j^* \rangle$, and $\mathbf{N} = \begin{bmatrix} N_1 & 0 & \cdots & 0 \\ 0 & N_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & N_n \end{bmatrix}$,

the matrix of the channel noise powers $N_i = \langle n_i n_i^* \rangle$. In the subsequent analysis, the noise power is assumed to be independent of channel number, so that **N** may be written

$$\mathbf{N} = \sigma^2 \mathbf{I} \quad (7)$$

where **I** is the n x n identity matrix, and σ^2 is the noise power per channel. Although the noise level may not be constant between channels in practice, this introduces little error provided the noise levels are much lower than the signal levels of interest.

EIGENVECTOR SEPARATION TECHNIQUE

In the eigenvector separation technique the eigenvectors of the covariance matrix **R** are used as weight vectors to transform the array input signals vector $\tilde{x}$ into an output vector $\tilde{y}$. If the technique is successful the components of $\tilde{y}$ should consist primarily of the separated independent signals s_j. The interference cancellation problem is then reduced to selecting the output channel which contains the desired signal.

Since **R** is a covariance matrix, it is Hermitian, its eigenvalues are real and non-negative, and its eigenvectors orthogonal [5]. Define $\mathbf{E} = [\tilde{e}_1, \tilde{e}_2, \cdots, \tilde{e}_n]$ as the matrix whose column vectors are the normalized eigenvectors $\tilde{e}_i$ of **R**, ordered according to the size of their corresponding eigenvalues λ_i, from largest to smallest. As the eigenvectors $\tilde{e}_i$ are normalized and orthogonal, **E** is a unitary matrix.

The eigenvector equation can be written

$$\mathbf{R}\mathbf{E} = \mathbf{E}\Lambda \quad (8)$$

where Λ is the diagonal matrix whose ith diagonal element is the eigenvalue λ_i corresponding to the eigenvector $\tilde{e}_i$. Since $\mathbf{E}^H = \mathbf{E}^{-1}$ for unitary matrices, premultiplying both sides of (8) by $\mathbf{E}^H$ yields

$$\mathbf{E}^H\mathbf{R}\mathbf{E} = \Lambda\,. \quad (9)$$

The eigenvectors are used as weights to transform the array input signals vector $\tilde{x}$,

$$y_i = \tilde{e}_i^H\,\tilde{x}\,, \text{ or } \tilde{y} = \mathbf{E}^H\,\tilde{x} \quad (10)$$

where $\tilde{y} = [y_1, y_2, \cdots, y_n]^T$ is the vector of output signals. From equation 9 the covariance matrix **R'** of the output signals is the diagonal matrix of the eigenvalues,

$$\mathbf{R}' = \langle \tilde{y}\,\tilde{y}^H \rangle = \mathbf{E}^H \langle \tilde{x}\,\tilde{x}^H \rangle \mathbf{E} = \mathbf{E}^H\mathbf{R}\mathbf{E} = \Lambda \quad (11)$$

Equation 11 implies that the use of eigenvectors as weights yields a set of output signals y_i which are non-correlating.

In equation 6 the covariance matrix **R** was written as the sum of two matrices, one due to the signals, and the other due to noise:

$$\mathbf{R} = \mathbf{S} + \mathbf{N}, \text{ where } \mathbf{S} = \mathbf{G}\mathbf{P}\mathbf{G}^H \quad (12)$$

If the array response vectors $\tilde{g}_j$ to the signals s_j, $j=1,m$, are independent, then **S** is of rank m and has m non-zero eigenvalues. **S** is Hermitian and its eigenvectors are orthogonal. Let $\tilde{u}_i$, $i=1,m$, be the normalized eigenvectors of **S** and γ_i be their corresponding eigenvalues. Since **N** is of the form $\sigma^2 I$ (equation 7),

$$\mathbf{R}\tilde{u}_j = (\gamma_i + \sigma^2)\tilde{u}_j \ . \qquad (13)$$

It is evident from (13) that the eigenvectors $\tilde{u}_j$ of **S** are also eigenvectors of **R**, and the eigenvalues of **R** are given by

$$\lambda_i = \gamma_i + \sigma^2 \qquad (14)$$

Therefore $\tilde{u}_j$ can be replaced by $\tilde{e}_i$, $i=1,m$, and expressed as columns of the matrix $E_S = [\tilde{e}_1, \tilde{e}_2, \cdots, \tilde{e}_m]$, ordered as before from largest to smallest. The eigenvectors E_S are determined by the signals and span the signal subspace of **R**.

In addition to these orthonormal vectors, an arbitrary set of n-m orthonormal vectors $\tilde{e}_j$, $j=m+1,n$, can be chosen from the remaining orthogonal subspace of **R**. These vectors are transformed by **S** into the null vector, and are eigenvectors of **R** corresponding to the common eigenvalue σ^2 :

$$\mathbf{R}\tilde{e}_j = (\mathbf{S} + \mathbf{N})\,\tilde{e}_j = \mathbf{N}\tilde{e}_j = \sigma^2\tilde{e}_j \qquad (15)$$

The matrix of these remaining eigenvectors is denoted by

$$E_N = [\tilde{e}_{m+1}, \tilde{e}_{m+2}, \cdots, \tilde{e}_n] \qquad (16)$$

The total set of eigenvectors is then represented by the columns of the matrix

$$\mathbf{E} = [E_S, E_N] \qquad (17)$$

where E_S spans the signal subspace and E_N spans the noise subspace of **R**.

From equations 11, 14, and 15 the output correlation matrix $\langle \tilde{y}\,\tilde{y}^H \rangle$ the output signals are noncorrelating, with powers given by the diagonal elements of $\langle \tilde{y}\,\tilde{y}^H \rangle$ according to

$$P_i(\text{out}) = \langle y_i\, y_i^* \rangle = \begin{cases} \gamma_i + \sigma^2 & \text{for } i=1 \text{ to } m \\ \sigma^2 & \text{for } i=m+1 \text{ to } n \end{cases} \qquad (18)$$

Consequently, under eigenvector weighting, the first m output channels contain the signals plus noise while the last n-m contain noise only.

CONDITIONS FOR PERFECT SIGNAL SEPARATION

One is tempted to suggest that since the original independent signals are non-correlating, the eigenvector transformation has separated the array input signals into the original signals. While this is not true in general, it is enlightening to explore the conditions under which perfect signal separation is achieved.

It is now shown that perfect separation of signals by the eigenvector technique occurs only if the array signal-response vectors are orthogonal

From equations 4 and 10 the output signal $\tilde{y}$ can be expressed in terms of the input signals $\tilde{s}$ and noise $\tilde{n}$ by

$$\tilde{y} = \mathbf{E}^H\tilde{x} = \mathbf{E}^H\mathbf{G}\,\tilde{s} + \mathbf{E}^H\tilde{n} = \mathbf{A}\,\tilde{s} + \mathbf{E}^H\tilde{n} \qquad (19)$$

where $\mathbf{A} = \mathbf{E}^H\mathbf{G}$ is an n x m matrix having ijth component a_{ij}. The signal for the ith output channel is given by

$$y_i = \sum_{k=1}^{m} a_{ik}s_k + \text{noise term} \qquad (20)$$

Perfect signal separation means that each independent signal s_k is found in one and only one output channel. The signal ordering is arbitrary and can be chosen so that s_k appears in the kth channel, for $k=1,m$. This means, using (20), that the components of **A** can be written as $a_{ik} = a_i\,\delta_{ik}$, where δ_{ik} is the Kroenecker delta. Since $\mathbf{E}^H\mathbf{G} = \mathbf{A}$, and **E** is unitary, the array response matrix can be expressed as

$$\mathbf{G} = \mathbf{EA} = [\tilde{e}_1, \tilde{e}_2, \cdots, \tilde{e}_n] \begin{bmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_m \\ 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

or,

$$\mathbf{G} = [\tilde{g}_1, \tilde{g}_2, \cdots, \tilde{g}_m]$$

$$= [a_1\tilde{e}_1, a_2\tilde{e}_2, \cdots, a_m\tilde{e}_m] \qquad (21)$$

Since the eigenvectors $\tilde{e}_j$ are orthonormal, the array response vectors $\tilde{g}_j$ are orthogonal.

The converse is also true (orthogonal response vectors implies that eigenvector weighting separates the signals perfectly), provided the signal-related eigenvalues are distinct. This is shown next.

If the array response vectors $\tilde{g}_j$ are orthogonal, they can be shown to be eigenvectors of **R** by postmultiplying (6) by $\tilde{g}_j$,

$$\mathbf{R}\tilde{g}_j = \mathbf{G}\mathbf{P}\mathbf{G}^H \tilde{g}_j + \mathbf{N}\tilde{g}_j = |\tilde{g}_j|^2 P_j \tilde{g}_j + \sigma^2 \tilde{g}_j = \lambda_j \tilde{g}_j \tag{22}$$

where $\lambda_j = |\tilde{g}_j|^2 P_j + \sigma^2$ is the corresponding eigenvalue of **R**. Since the signal eigenvalues λ_j are distinct, the corresponding normalized eigenvectors consist of the unique signal eigenvectors given by

$$\mathbf{E}_S = [\tilde{g}_1/|\tilde{g}_1|, \tilde{g}_2/|\tilde{g}_2|, \cdots, \tilde{g}_m/|\tilde{g}_m|] \tag{23}$$

plus the orthogonal noise eigenvectors $\mathbf{E}_N = [\tilde{e}_{m+1}, \tilde{e}_{m+2}, \cdots, \tilde{e}_n]$. From (23), (17) and the orthogonality of **G** and $\mathbf{E}_N$, **A** can be written as

$$\mathbf{A} = \mathbf{E}^H\mathbf{G} = \begin{bmatrix} \mathbf{E}_S^H \\ \mathbf{E}_N^H \end{bmatrix} \mathbf{G} = \begin{bmatrix} |g_1| & 0 & \cdots & 0 \\ 0 & |g_2| & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & |g_m| \\ 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \tag{24}$$

Thus **A** has a quasi-diagonal form, which from (20) is equivalent to perfect signal separation at the output.

TWO-SIGNAL CASE

When only two signals (e.g., a communications signal and a single interfering signal) are incident on an array using eigenvector weighting, only the first two output channels will contain signal contributions (equation 18). The signals in these two channels are uncorrelated (equation 11). As will now be shown, this lack of correlation has interesting implications.

The output signals y_1, y_2 can be written in terms of the initial independent signals s_1, s_2 and channel noise n_k, k=1,n, as

$$y_i = a_{i1}s_1 + a_{i2}s_2 + \sum_{k=1}^{n} w_{ki}^* n_k, \quad i = 1,2 \tag{25}$$

where w_{ki}, k=1,n, are the weights $\tilde{w}_i$ used for the ith output channel. The non-correlation of outputs y_1 and y_2 can be expressed using (25) and the fact that the signals s_l and noise n_k do not correlate, as

$$\begin{aligned} 0 &= \langle y_1 y_2^* \rangle \\ &= a_{11}a_{21}^* P_1 + a_{12}a_{22}^* P_2 + \tilde{w}_1^H \tilde{w}_2 \sigma^2 \end{aligned} \tag{26}$$

P_1,P_2, and σ^2 are the signal and noise powers respectively. The eigenvector weights are orthogonal, making the last term of (26) zero. Equation 26 becomes

$$\frac{P_1}{P_2} = -\frac{a_{12}\,a_{22}^*}{a_{11}\,a_{21}^*} \tag{27}$$

From (25), the power due to the jth signal s_j in the ith output channel is given by

$$\beta_{j(i)} = a_{ij}\,a_{ij}^* P_j \tag{28}$$

and the ratios of the powers of the signals s_1 and s_2 in the ith output channel, by

$$\frac{\beta_{1(i)}}{\beta_{2(i)}} = \frac{a_{i1}\,a_{i1}^*}{a_{i2}\,a_{i2}^*}\,\frac{P_1}{P_2} \tag{29}$$

By manipulating (27) and (29), it is seen that

$$\frac{\beta_{1(1)}}{\beta_{2(1)}} = \frac{\beta_{2(2)}}{\beta_{1(2)}} \tag{30}$$

Interpreting one of the signals as a desired signal and the other as interference, equation 30 states that the signal-to-interference ratio is inverted between the two output signal channels.

PERFORMANCE RELATIVE TO POWER RATIO INVERSION

The Gram-Schmidt weighting technique [2], like the eigenvector weighting technique, produces outputs signals which do not correlate. Equation 30 can be shown to apply to this technique as well, for low noise levels [6]. Since for the Gram-Schmidt technique, the first output channel is the same as the first input channel, the second output channel may be interpreted as inverting the relative powers of the two signals as seen by the first array element. The constrained matrix inversion technique can be shown to have a performance similar to that of the Gram-Schmidt method [2,6]. These two techniques thus have a performance similar to that of the power ratio inversion algorithm [1].

This section compares the performance of eigenvector weighting to that of the above-mentioned power-ratio-inverting techniques in separating two incoming signals. It is shown that for arrays with similar element antenna patterns, eigenvector weighting performs better than power ratio inversion for all signal directions, with power ratio inversion providing a greatest lower bound to the eigenvector performance.

It is noted that the array input $\tilde{x}$ of a number of signals s_i, i=1,m incident on an n-element array with common element pattern $g(\theta,\phi)$, where θ and ϕ represent signal arrival direction azimuth and elevation respectively, is the same as the array input of a number of signals $s_i' = g(\theta_i,\phi_i)s_i$, i=1,m, incident on an array of identical geometry and isotropic element patterns. Therefore the condition of identical element patterns may be replaced by the easier-to-handle condition of isotropic patterns.

From equation 4, the signal in the first eigenvector output channel is given by

$$y_1 = \tilde{e}_1^H \tilde{x} = \tilde{e}_1^H \mathbf{G} \tilde{s} + \tilde{e}_1^H \tilde{n}$$

$$= (\tilde{e}_1^H \tilde{g}_1)s_1 + (\tilde{e}_1^H \tilde{g}_2)s_2 + \tilde{e}_1^H \tilde{n} \qquad (31)$$

The first term on the right side of (31) is due to signal 1, the second, to signal 2, and the third, to noise. Therefore the ratio of signal powers in channel 1 is given for the eigenvector technique by

$$\left.\frac{\beta_{1(1)}}{\beta_{2(1)}}\right|_{eiv} = \frac{|\tilde{e}_1^H \tilde{g}_1|^2}{|\tilde{e}_1^H \tilde{g}_2|^2} \frac{P_1}{P_2} \qquad (32)$$

According to the ordering adopted in this paper, $P_1 \geq P_2$. This ordering, together with equations 30 and 32, implies that the eigenvector technique provides better signal separation than power ratio inversion when $|\tilde{e}_1^H \tilde{g}_1| > |\tilde{e}_1^H \tilde{g}_2|$, and when $|\tilde{e}_1^H \tilde{g}_1| = |\tilde{e}_1^H \tilde{g}_2|$, the signal separation provided by eigenvector weighting matches that of power ratio inversion. Therefore the problem, of proving that the performance of eigenvector weighting is everywhere better than that of power ratio inversion with power ratio inversion providing a greatest lower bound to eigenvector performance, is reduced to proving that

$$|\tilde{e}_1^H \tilde{g}_1| \geq |\tilde{e}_1^H \tilde{g}_2| \qquad (33)$$

everywhere with the equality being a limiting case.

The proof of (33) is of some length and is given in the Appendix.

COMPUTER-MODELLING

The eigenvector weighting technique was modelled on a PC using the MATLAB numeric computation package, for the case of a communications signal and a single jamming signal.

Except as noted, array element patterns were modelled as isotropic. The array response matrix $\mathbf{G} = [\tilde{g}_1, \tilde{g}_2]$ (equations 1-3) was calculated from a given array geometry and signal directions. The ideal covariance matrix $\mathbf{R}$ was then calculated from response matrix $\mathbf{G}$, together with given input signal powers P_1, P_2, and channel-independent noise power σ^2, using equations 6 and 7. From the ideal covariance matrix thus obtained, the normalized eigenvectors $\tilde{e}_k$, k=1,n, were computed, and ordered according to the sizes of their eigenvalues, strongest first.

The output powers in the kth channel, k=1,2, due to the jth signal, j=1,2, were computed from

$$\beta_{j(k)} = P_j |\tilde{e}_k^H \tilde{g}_j|^2 \qquad (34)$$

The corresponding noise power in each of the output channels was given by σ^2.

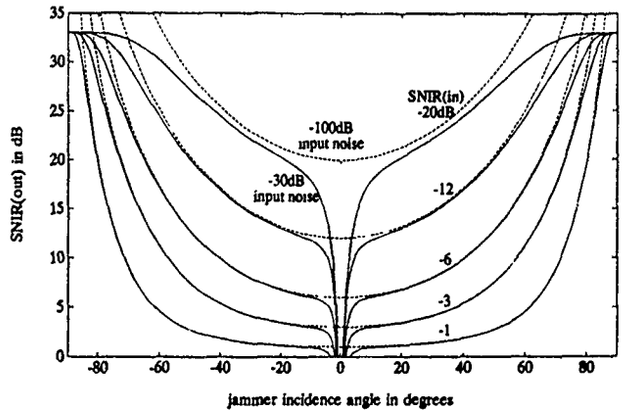
The output channel favouring the communications signal was selected for evaluation purposes This was usually the second channel, as the communications power was normally less than the jamming power.

Figure 1a shows the computed ideal performance of the eigenvector separation technique, for the case of a two-element half-wavelength-spaced array and a normally incident communications signal The output signal to noise-plus-interference ratio SNIR(out) is plotted as a function of jamming incidence angle for various values of SNIR(in). Two sets of curves are shown: one for negligible noise power (-100dB relative to communications) and the other, for a plausible noise power (-30dB)

With negligible input noise power, when the jamming incidence angle is close to that of the communications (0°), the performance equals that of power ratio inversion. (A singular point for which SNIR(out) is undefined occurs when the angles are exactly equal; this is not shown on the figures.) When the jamming angle approaches either + or -90°, SNIR(out) becomes infinite, suggesting perfect separation. The response vector of the array to the jamming is orthogonal to the communications response vector for these directions. At intermediate angles, performance lies between that of perfect separation and power ratio inversion It is interesting to note, that even with near-equal power levels (SNIR(in) = -1 and -3dB in Figure 1a), acceptable performance (taken as SNIR(out) > 10dB) is achieved for a significant fraction of jammer incidence angles.

The addition of input noise, as illustrated in Figure 1a, causes the performance to drop at nearly-

Figure 1a. SNIR(out) as a function of jamming incidence angle, for a normally incident communications signal and various values of SNIR(in) on a λ/2-spaced 2-element array using eigenvector weighting. Curves are shown for negligible input noise power (-100dB relative to communications) and plausible input noise power (-30dB).



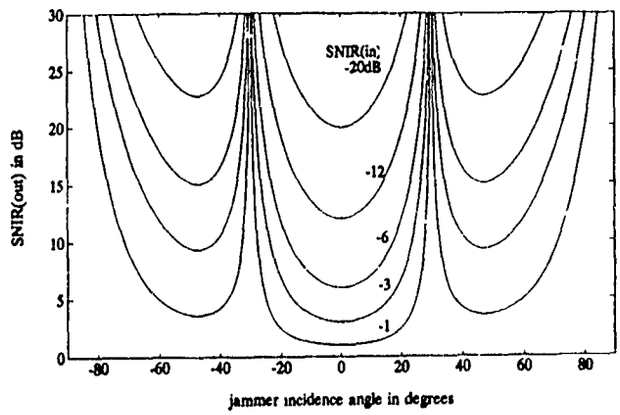
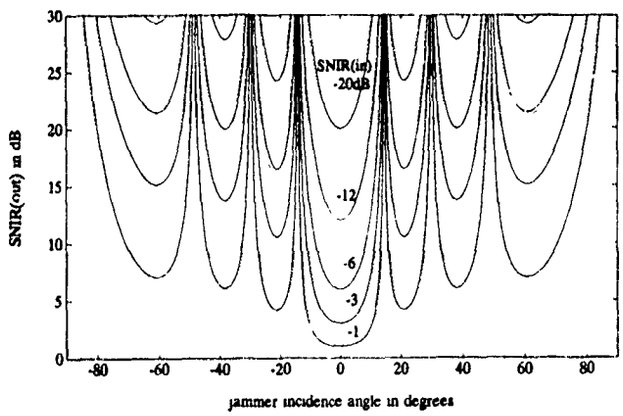


Figure 1b. SNIR(out) as a function of jamming incidence angle, for a normally incident communications signal and various values of SNIR(in) on a λ/2-spaced 4-element linear array using eigenvector weighting. Input noise power was negligible (-100dB relative to communications).

Figure 1c. SNIR(out) as a function of jamming incidence angle, for a normally incident communications signal and various values of SNIR(in) on a λ/2-spaced 8-element linear array using eigenvector weighting. Input noise power was negligible (-100dB).



equal signal directions. This happens because the output communications power approaches zero when its direction approaches that of the stronger jamming signal, while the output noise power remains finite. This is a feature of most interference-cancelling techniques, not only eigenvector weighting. The addition of noise also limits the maximum value of SNIR(out) which is achieved at directions where the input signals are orthogonal, to the input signal plus noise ratio SNR(in) (30dB) plus the array gain (3dB). Apart from these limitations, the observations noted

for the noise-absent case are true for the noise-present case.

In the remainder of the modelling results discussed, a negligible noise power (-100dB) has been used. This is done in order to simplify the results so that other features are clearer. The addition of noise does not alter the conclusions thus obtained.

Figures 1b and 1c show the corresponding performances for a four and eight-element linear half-wavelength spaced array, respectively. Like the two-element case, the performance drops to that of power ratio inversion when the jamming angle approaches that of the communications, and increases to that of perfect signal separation, at jamming angles corresponding to orthogonal jamming and communications response vectors. The performances for near-equal input power levels (SNIR(in) = -1, -3dB) are seen to be acceptable (SNIR(out) > 10dB) over a wider range of jamming angles, as the number of elements increases.

Figures 2a, b, and c show corresponding three-dimensional plots for the case of SNIR(in) = -3dB, for the two, four, and eight-element half-wavelength spaced array respectively. SNIR(out) is plotted as a function of the jamming and communications incidence angle, in this case. The plot extends over all jamming and communications incidence angles, -90 to +90°, relative to the array. Provided that the angles lie near the horizon, the incidence angles represent equally likely azimuth angles. In order to provide a measure of the fraction of cases for which communications is acceptable, the plots are truncated at the SNIR(out) = 10dB level. The fraction of angles for which SNIR(out) > 10dB is thus illustrated. As previously suggested by figures 1a, b, and c, acceptable communications is achieved for near-equal input powers over a wider fraction of angles, as the number of elements increases.

The modelling represented by Figures 2a, b, and c was repeated for other values of SNIR(in), to

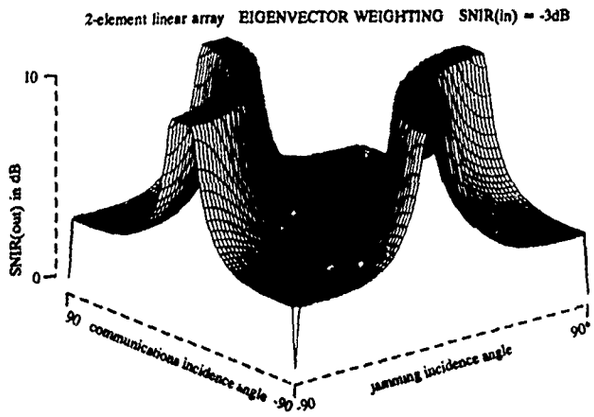


Figure 2a. Three-dimensional plot of SNIR(out) as a function on communications and jamming incidence angle, for input communications, jamming and noise signals of 0, 3, and -100dB respectively, on a λ/2-spaced 2-element array using eigenvector weighting. The plot is limited above at 10dB.

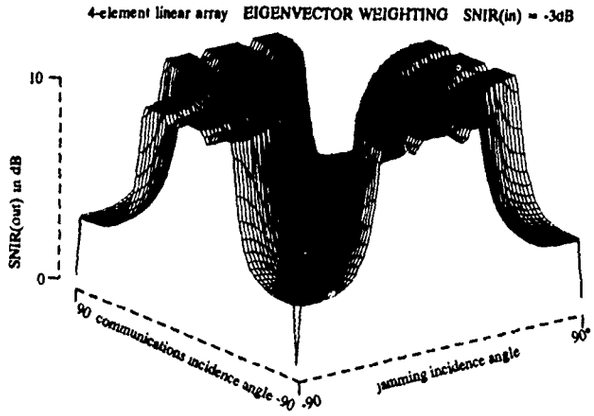


Figure 2b. Three-dimensional plot of SNIR(out) as a function on communications and jamming incidence angle, for input communications, jamming and noise signals of 0, 3, and -100dB respectively, on a λ/2-spaced 4-element linear array using eigenvector weighting. The plot is limited above at 10dB.

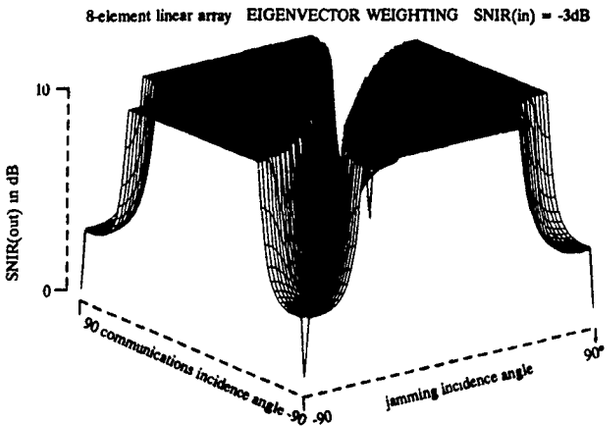


Figure 2c. Three-dimensional plot of SNIR(out) as a function on communications and jamming incidence angle, for input communications, jamming and noise signals of 0, 3, and -100dB respectively, on a λ/2-spaced 8-element linear array using eigenvector weighting. The plot is limited above at 10dB.

determine the extent to which acceptable communications can be achieved with eigenvector weighting, at near-equal signal levels. The fraction of angles, for which acceptable communications (SNIR(out) > 10dB) is achieved, is plotted for the two, four and eight-element linear arrays as a function of SNIR(in) in Figure 3. This fraction is seen to be symmetric about SNIR(in) = 0°. This result arises from the performance symmetry in the two-signal case, described in equation 30 in terms of the relative output powers being inverted between the two signal output channels. The improvement in performance with increasing number of elements is also evident. To put these results in perspective, it should be noted that the power-inversion techniques would not result in acceptable communications at these values of SNIR(in), for any angles. While not a completely satisfactory solution for the case of

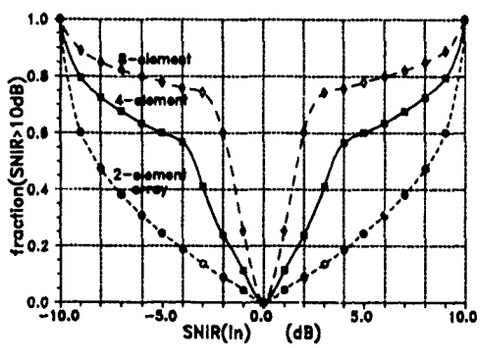


Figure 3. Fraction of communications and jamming directions for which SNIR(out) exceeds 10dB, as a function of SNIR(in), for 2, 4, and 8-element l/2-spaced linear arrays using eigenvector weighting.

near-equal powers, Figure 3 shows that eigenvector separation can provide acceptable performance in a large fraction of cases, and thus is substantially better than the power-inverting alternatives.

The effect of gain differences in the input channels is illustrated in Figures 4a and b. These figures demonstrate the effects of reducing the gain in one channel, relative to the others, for the two and four-element linear half-wavelength arrays The communications signal was normally incident and SNIR(in) was -3dB, with a negligible noise level. SNIR(out) was plotted as a function of jamming incidence angle, for various channel gain mismatches It can be seen that the array performance for those directions where eigenvector separation performs best deteriorates rapidly with channel gain differences. This effect is worst for the two element case; for very large gain differences, the 2-element performance is everywhere reduced to that of power ratio inversion. For the four-element case the behavior is more complex, and depends on which and how many elements have reduced gain. The example shown in Figure 4b is for a single attenuated channel. When three of the four channels are severely attenuated, the performance was also found to be reduced to that of power ratio inversion. These results indicate that efforts should be made to provide equal input channel gains for systems where the eigenvector separation technique is used.

The superior performance of the eigenvector technique over power ratio inversion methods was shown analytically for the case of two signals and identical element patterns. On an intuitive basis, the increased degrees of freedom used in determining the weights for the first and second output channels suggest that the eigenvector technique is likely to be

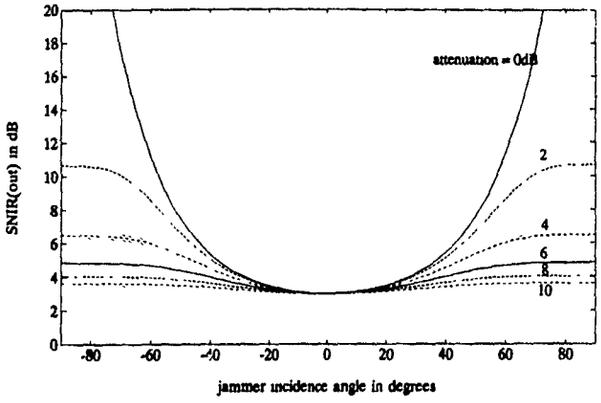


Figure 4a. Effect of channel gain differences on eigenvector weighting performance of a λ/2-spaced 2-element array, for a normally incident communications signal and communications, jamming, and noise powers of 0, 3, and -100dB, and various values of attenuation applied to one of the channels.

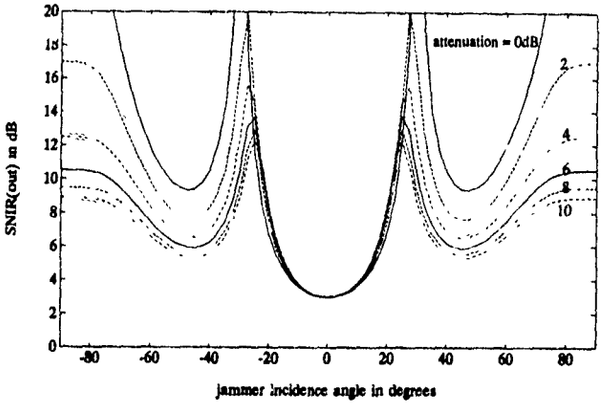


Figure 4b. Effect of channel gain differences on eigenvector weighting performance of a λ/2-spaced 2-element linear array, for a normally incident communications signal and communications, jamming, and noise powers of 0, 3, and -100dB, and various values of attenuation applied to one of the channels.

superior to power ratio inversion (as demonstrated by Gram-Schmidt orthogonalization which fixes the first weight vector and permits only one degree of freedom in the second weight vector) much more often than not.

A plausible example of non-identical element patterns is provided in the four-element square array with differently-directed offset-circular element patterns, illustrated in Figure 5. The performance for this array using a front-to-back ratio of 10dB in the element patterns, for SNIR(in) = -3dB, is illustrated in Figure 6. In this figure, SNIR(out) is plotted versus jamming and communications incidence angles in a 3-dimensional format. As can be seen in this figure, the array performance has a minimum of SNIR(out) = 3dB, which occurs when the jamming incidence angle approaches that of the communications. Substantially better performances are achieved for a large fraction of possible

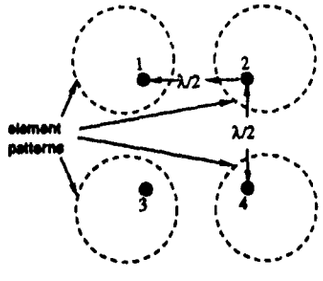


Figure 5. Modelled square array with differently-directed offset-circle element patterns.

directions. Thus the eigenvector performance with this array is better than that of a similar array with non-directive elements using power-ratio inverting techniques.

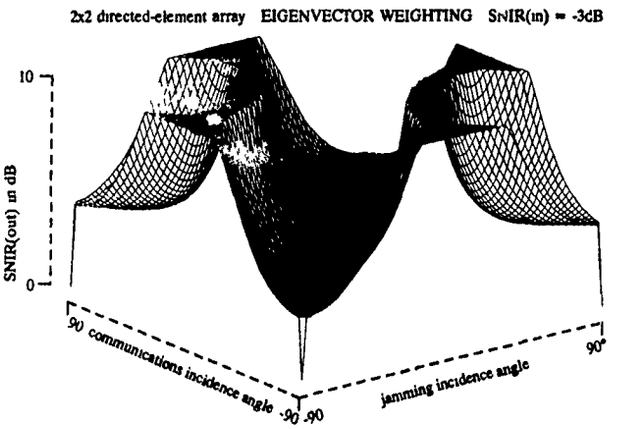


Figure 6. Three-dimensional plot of SNIR(out) obtained with eigenvector weighting, as a function of communications and jamming incidence angles, for input communications, jamming, and noise signals of 0, 3, and -100dB respectively, for the array of Figure 5 with a forward to back ratio of 10dB in the element patterns. The plot is limited above at 10 dB.

Further, since the signals which are weighted and combined in any adaptive array technique are those seen by the element antennas, the effect of directive elements on power-ratio inversion methods is expected to be merely a shift in the value of SNIR(in) corresponding to a certain value of SNIR(out). For any given set of signal directions, the poor values of SNIR(out) corresponding to near-equal input powers (as seen by the antennas) would still occur under power ratio inversion.

From the above, it can be inferred that in situations where a single algorithm is to be used to cover a range of possible signal powers and directions, the eigenvector technique is preferred whether the array elements are differently directed or not. However, if the implementation is such that different algorithms can be tried on a case-by-case basis, then including a power ratio inversion technique along with eigenvector weighting will at times improve performance, if the array elements are differently directed.

SIMULATION

The ideal covariance matrix **R** assumed in the analysis and used to obtain the modelling results reported herein differs from the time-average estimate found in practice, in that the estimation errors due to the finite averaging period and random signal variations have not contributed to the covariance matrix elements. These errors give rise to small apparent correlations between the two independent signals and the noise signals. As a check on the practical applicability of the analysis and the reliability of the modelling studies, a more detailed simulation was carried out for the two and four-element arrays.

The simulations included a binary-phase-coded communications signal plus a single jamming signal, incident on half-wavelength spaced linear arrays The jamming signal was either pure-tone or white noise. The input signals were constructed on a sample-by sample basis, using a sample rate of 5 samples/bit. Noise values and white-noise jamming values were assumed to be independent between samples. From each block of 64 input samples, a block covariance matrix was formed. A weighted average was taken between the block covariance matrix and the previous covariance matrix estimate, in order to form a new covariance matrix estimate. The time constant of the weighted averaging was 2.54 blocks. The simulation was run for 32 blocks, and from the weights thus obtained for each block, SNIR(out) found, on a block-by-block basis.

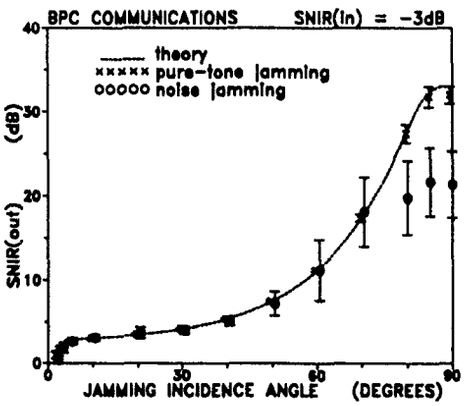


Figure 7a. Theoretical and simulated performance of a λ/2-spaced 2-element array with eigenvector weighting, in the presence of a normally-incident 0dB communications signal, a 3dB jamming signal, and -30dB input noise Mean values and rms deviations are shown for pure-tone and noise jamming simulations.

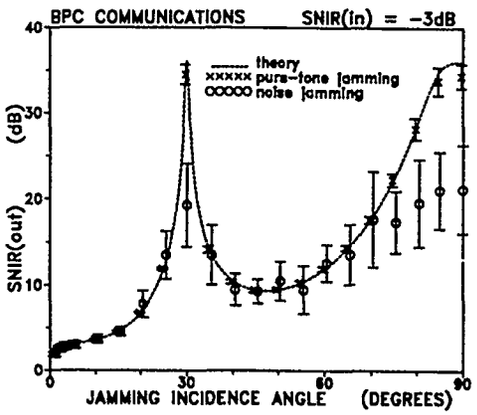


Figure 7b. Theoretical and simulated performance of a λ/2-spaced 4-element linear array with eigenvector weighting, in the presence of a normally-incident 0dB communications signal, a 3dB jamming signal, and -30dB input noise. Mean values and rms deviations are shown for pure-tone and noise jamming simulations.

Figures 7a and b show the simulation performance, for 3dB input jamming power and -30dB noise power, relative to the input communications power, for the two and four-element arrays respectively. SNIR(out) is plotted as a function of the jamming incidence angle, for a normally incident communications signal. Average values and rms deviations (on the block-by-block results) are plotted for the cases of pure-tone and white-noise jamming. Also shown are the theoretical statistics-independent curves, computed from the previous modelling program, for these input power levels.

The pure-tone jamming results agree well with the theoretical curves for both the two and four-element arrays, implying that the random binary-phase coding of the communications does not substantially affect performance. The noise jamming results are in agreement with the theoretical curves, at values of SNIR(out) up to 15dB. At higher valuesof SNIR(out), the random nature of the noise jamming causes the average performance to be reduced below that of the theoretical curve. Also, the block-to-block rms variation in performance is greater for the noise jamming, and for higher values of SNIR(out). It should be noted that the reduction in performance below the theoretically expected value, found for the case of noise jamming, does not extend down to the minimum values of SNIR (approx. 10dB) needed for acceptable communications.

The simulation results support the validity of the conclusions drawn from the analysis and the ideal modeliing studies, for practical applications.

REFERENCES

[1] R.T. Compton, Jr., "The Power-Inversion Adaptive Array: Concept and Performance", IEEE Trans. Aerospace Electron. Syst., Vol. AES-15, pp. 803-814, 1979

[2] K. Gerlach and F. Kretschmer, Jr., "Convergence Properties of Gram-Schmidt and SMI Adaptive Algorithms", IEEE Trans. Aerospace Electron. Syst., Vol. AES-26, pp. 44-56, 1990

[3] C.M. Hackett, Jr., " Adaptive Arrays Can Be Used to Separate Communication Signals", IEEE Trans. Aerospace Electron. Syst., Vol. AES-17, pp. 234-247, 1981.

[4] R. Spohn, V. Frost, R. Balasubramaniam, R. Moats, "Development of an Eigenvector Sorting Algorithm for HF Antenna Arrays", RADC-TR-88-68 Final Technical Report, University of Kansas, 1988

[5] S. Haykin,"Adaptive Filter Theory ", Ch. 2, Prentice-Hall, 1986

[6] R.W. Jenkins and K.W. Moreland, "A Comparison of the Eigenvector Separation and Gram-Schmidt Adaptive Antenna Techniques", submitted for publication, 1990

APPENDIX

It is proven here that, given isotropic array elements,

$$|\tilde{e}_1^H \tilde{g}_1| \geq |\tilde{e}_1^H \tilde{g}_2| \qquad (33)$$

In equation 12, the covariance matrix $\mathbf{R}$ is written as $\mathbf{R} = \mathbf{S} + \mathbf{N}$, where $\mathbf{S} = \mathbf{GPG}^H$ and $\mathbf{N} = \sigma^2\mathbf{I}$ are the signal and noise contributions respectively. The eigenvectors $\tilde{e}_1$, $\tilde{e}_2$ of $\mathbf{R}$ were shown to be also eigenvectors of the n x n matrix $\mathbf{S}$. To show (33) is true, it is helpful to find $\tilde{e}_1$ by solving the eigenvector equation of $\mathbf{S}$. Using the factorization

$$\mathbf{S} = (\mathbf{G}\sqrt{\mathbf{P}})(\mathbf{G}\sqrt{\mathbf{P}})^H, \qquad (A1)$$

where $\sqrt{\mathbf{P}} = \begin{bmatrix} \sqrt{P_1} & 0 \\ 0 & \sqrt{P_2} \end{bmatrix}$ and $\mathbf{G} = [\tilde{g}_1, \tilde{g}_2]$,

the eigenvector equation of $\mathbf{S}$ can be expressed as

$$(\mathbf{G}\sqrt{\mathbf{P}})(\mathbf{G}\sqrt{\mathbf{P}})^H \tilde{e}_k = \gamma_k \tilde{e}_k \qquad (A2)$$

Premultiplying (A2) by $(\mathbf{G}\sqrt{\mathbf{P}})^H$ yields the simpler 2-dimensional eigenvector equation

$$\mathbf{T}\tilde{v}_k = \gamma_k \tilde{v}_k , \qquad \text{(A3)}$$

where

$$\tilde{v}_k = (\mathbf{G}\sqrt{\mathbf{P}})^H \tilde{e}_k \qquad \text{(A4)}$$

and

$$\mathbf{T} = (\mathbf{G}\sqrt{\mathbf{P}})^H(\mathbf{G}\sqrt{\mathbf{P}}) \qquad \text{(A5)}$$

T is a 2 x 2 matrix whose eigenvalues are the eigenvalues γ_k of **S**, and whose eigenvectors are given by $\tilde{v}_k$. The eigenvector $\tilde{e}_k$ of **S** can be found in terms of $\tilde{v}_k$ by using (A4) to replace the factor $(\mathbf{G}\sqrt{\mathbf{P}})^H\tilde{e}_k$ on the right-hand side of (A2) by $\tilde{v}_k$, thus obtaining

$$\tilde{e}_k = \frac{1}{\gamma_k}(\mathbf{G}\sqrt{\mathbf{P}})\tilde{v}_k.$$

The eigenvector scale factor $\frac{1}{\gamma_k}$ is arbitrary and can be dropped, so that

$$\tilde{e}_k = (\mathbf{G}\sqrt{\mathbf{P}})\,\tilde{v}_k \qquad \text{(A6)}$$

Proceeding now to find the eigenvectors $\tilde{v}_k$ of **T**, **T** is expanded in terms of its elements using (A1) and (A5) to obtain:

$$\mathbf{T} = \begin{bmatrix} P_1|\tilde{g}_1|^2 & \sqrt{P_1P_2}\,\tilde{g}_1^H\tilde{g}_2 \\ \sqrt{P_1P_2}\,\tilde{g}_2^H\tilde{g}_1 & P_2|\tilde{g}_2|^2 \end{bmatrix} \qquad \text{(A7)}$$

The quadratic equation $\det(\mathbf{T}-\gamma\mathbf{I}) = 0$ is solved for the eigenvalues of **T**, obtaining

$$\gamma_{1,2} = \frac{P_1|\tilde{g}_1|^2+P_2|\tilde{g}_2|^2}{2} \pm \sqrt{P_1P_2}\,|\tilde{g}_1^H\tilde{g}_2|\sqrt{1+\chi^2} \qquad \text{(A8)}$$

$$\text{where } \chi = \frac{P_2|\tilde{g}_2|^2-P_1|\tilde{g}_1|^2}{2\sqrt{P_1P_2}\,|\tilde{g}_1^H\tilde{g}_2|} \qquad \text{(A9)}$$

γ_1 corresponds to the plus sign in (A8) since it is the larger eigenvalue by the ordering convention used. Setting

$$e^{j\beta} = \frac{\tilde{g}_1^H\tilde{g}_2}{|\tilde{g}_1^H\tilde{g}_2|} \quad \text{and} \quad \mu = \chi + \sqrt{1+\chi^2}, \qquad \text{(A10)}$$

$\tilde{v}_1$ is given by

$$\tilde{v}_1 = \begin{bmatrix} 1 \\ \mu e^{-j\beta} \end{bmatrix} \qquad \text{(A11)}$$

Substituting (A10) and (A11) in (A6), $\tilde{e}_1$ is found to be

$$\tilde{e}_1 = (\mathbf{G}\sqrt{\mathbf{P}})\,\tilde{v}_1 = \sqrt{P_1}\,\tilde{g}_1 + \sqrt{P_2}\,\mu e^{-j\beta}\,\tilde{g}_2 \qquad \text{(A12)}$$

and

$$|\tilde{e}_1^H\tilde{g}_1| = |\tilde{g}_1|^2(\sqrt{P_1} + \xi\mu\sqrt{P_2}), \qquad \text{(A13)}$$

$$\text{where } \xi = \frac{|\tilde{g}_1^H\tilde{g}_2|}{|\tilde{g}_1|^2}. \qquad \text{(A14)}$$

Similarly,

$$|\tilde{e}_1^H\tilde{g}_2| = |\tilde{g}_1|^2(\xi\sqrt{P_1} + \mu\sqrt{P_2}) \qquad \text{(A15)}$$

The array elements are assumed to have isotropic gains. Therefore, $\tilde{g}_k$ is of the form

$$[\rho_1 e^{j\delta 1k}, \rho_2 e^{j\delta 2k}, \cdots, \rho_n e^{j\delta nk}]^T,$$

implying that

$$|\tilde{g}_1| = |\tilde{g}_2| = (\textstyle\sum_{i=1}^{n} \rho_i^2)^{1/2} \qquad \text{(A16)}$$

From (A15), (A16) and the Cauchy-Schwartz inequality,

$$\xi = \frac{|\tilde{g}_1^H\tilde{g}_2|}{|\tilde{g}_1|^2} \le \frac{|\tilde{g}_1||\tilde{g}_2|}{|\tilde{g}_1|^2} = 1, \text{ which implies that}$$

$$0 \le \xi \le 1 \qquad \text{(A17)}$$

From (A9), the input power ordering ($P_1 \ge P_2$), and isotropy ($|\tilde{g}_1| = |\tilde{g}_2|$), χ is non-positive. Applying this to (A10), μ is seen to lie in the range

$$0 \le \mu \le 1 \qquad \text{(A18)}$$

From (A15), (A16), (A17) and (A18), together with the input power ordering, it can be shown that

$$|\tilde{e}_1^H\tilde{g}_1| - |\tilde{e}_1^H\tilde{g}_2| = |\tilde{g}_1|^2(1-\xi)(\sqrt{P_1} - \mu\sqrt{P_2}) \ge 0 \qquad \text{(A19)}$$

Therefore, $|\tilde{e}_1^H\tilde{g}_1| \ge |\tilde{e}_1^H\tilde{g}_2|$, which proves that the eigenvector technique provides better signal separation than the power ratio inversion techniques under the condition of isotropic element patterns. When the signals are close in direction, then ξ approaches unity (A14), thus causing $|\tilde{e}_1^H\tilde{g}_1|$ to approach $|\tilde{e}_1^H\tilde{g}_2|$ (A19). Therefore, power ratio inversion provides a greatest lower bound to the eigenvector separation performance, given the isotropy condition.

DISCUSSION

George H. HAGN

What is the impact of coherent multipath on the method you have presented ?

Have you gone beyond the theorefical stage which you presented to us to build a practical system ?

AUTHOR'S REPLY

If the multipath is completely coherent, then the resultant signal from that source will have a certain time-independent phase relationship between antenna elements. The technique will deal with it as it would any simple-mode signal, and performance would not be reduced. If coherence is partial however, a deterioration in performance might result.

We have a test bed for HF and VHF applications, part of which is a programmable real-time processing system. The eigen vector method will be implemented on that test bed, in 2-element form, in the near future, and, in 4-element form, about a year from now.

TECHNIQUES FOR ESTIMATING THE EFFECTS OF MAN-MADE RADIO NOISE ON DISTRIBUTED MILITARY SYSTEMS

by
D.B.Sailors
Ocean and Atmospheric Sciences Division
Naval Ocean Systems Center
San Diego, CA 92152-5000

SUMMARY

This paper reviews techniques available for estimating the effects of man-made radio noise on distributed military systems using empirical man-made noise models. The models given in CCIR Report 258 are reviewed along with the empirical data base upon which they are based. Results of measurements of man-made noise are presented for six Pacific Ocean sites and for three Atlantic Ocean/Europe sites. Accumulative probability distribution models of increasing complexity are reviewed. Tests of fit of these distributions are presented for select samples of measured man-made noise data.

1. INTRODUCTION

Increased utilization of electrical and electronic devices for man's well being and security has also increased the amount of undesired electromagnetic energy in telecommunication systems. This man-made radio frequency interference is characteristically impulsive in form and random in occurrence and originates from such common sources as automobile ignition systems, high voltage transmission lines, electrical power generating stations, and electrical applicances and machinery. Man-made noise predominates over noise of natural origin[1] at many locations especially during the daytime.

In the solution of telecommunication problems, it is highly desirable to be able to estimate the radio noise at any location, frequency, or time of day caused by these different sources. One approach to solving the problem has been the development of empirical models of man-made noise. The empirical models available today are based on an empirical data base acquired by the Institute for Telecommunication Sciences(ITS)[2]. Data exist on man-made radio noise available power levels in the U.S. and on the time and location variabilities of those levels for specific environmental categories: rural,residential and business. These models have been adopted by the CCIR in Report 258-5[3]. These levels are presented as F_a, the effective noise figure, in dB(kT_o)[3] or dB($kT_o b$)[1]; however, these are essentially equivalent units[4]. The distribution of F_a is needed to calculate the distribution of SNR(defined as the ratio of average signal power to average noise power) and hence the probability of successful communication or the loss of circuit reliability due to interference.

This paper presents empirical models of man-made noise that can be used in the solution of telecommunication problems for distributed military systems. First, the empirical data for which the CCIR models are based is discussed. Ther the CCIR Report 258 models are described. Four models for the probability distribution of the short-term(1 min.) mean values of man-made radio noise available power levels for the specific environmental categories are given(these models are not amplitude probability distributions(APD)). The extension to frequencies above 250 MHz, based on data measured in Canada, is described[5-7]. Results of recent measurements of man-made noise made on a circular disposed antenna array(CDAA) by the Naval Electronics Engineering Activity, Pacific are discussed[8]. Results of these measurements for Guam are compared to measurements made near the CDAA antenna in Guam during 1974 on a short vertical rod[9].

2. EMPIRICAL DATA BASE

In 1974 the Institute for Telecommunication Sciences(ITS) reported the use of a measurement system with a rms detector to obtain data in the band 250 kHz to 250 MHz with a short vertical antenna near ground at various sites in the U.S.[2]. Over 300 hours of data were obtained simultaneously on ten frequencies over a period from 1966 through 1971 in six states(Colorado, Maryland, Texas, Virginia, Washington and Wyoming) and in the District of Coluumbia. Data were obtained for three environmental categories: rural, residential, and business. Data were measured in 31 rural areas, 38 residential areas, and 23 businsss areas. At each location the area sampled varied from a few square blocks in the business areas to few square miles in the rural areas. In the noise measurement method used, 10-sec samples of the running average(time constant about 50 sec.) of F_a were recorded. Thus, 360 samples of F_a were obtained each hour for a given measurement location and frequency. These results were analysed statistically at ITS. The least-suares fit for F_{am}, the median value of F_a, given in [2-3] for each environmental category is reproduced as Figure 1 and are given here as Table 1 for each of the frequencies. The slope with frequency was found to be -27.7 dB/decade for each environmental category.

These man-made noise data are daytime values. At night these 20-50 MHz levels can drop 5-10 dB to a minimum around 0400 hours local time, and between 100 MHz and 250 MHz they can drop 3-5 dB. At the lower frequencies in the HF band, the night levels are frequently controlled by atmospheric noise from lightning; and the man-made levels cannot be observed. The diurnal variation decreases for the MF band and is again only 3-5 dB at 0.25 MHz, with values at night being slightly higher than during the day.

An indication of the variation encountered from location to location within each environmental category is given by Spaulding and Disney[2]. An example

distribution of local median values of man-made noise at 20 MHz in residential areas is given as Figure 2. The values σ_T is the standard deviation of all measured medians for all frequencies combined about the regression line of Figure 1. The authors quote values of 7.00 dB, 5.00 dB, and 6.45 dB for business, residential, and rural areas, respectively. A better estimate of the location variability of Fam for these environmental categories for a specific measured frequency may be obtained from the standard deviations σ_{NL} given in Table 1.

Once the best estimate of the location variability has been found, the variation of F_a at that location with time should be considered. Figure 3 gives the distribution of F_a values obtained on 20 MHz during an hour (0839-0939 hours local time) in a residential area in Boulder, CO. The median and the upper and lower deciles are indicated. The time variability for the different environmental categories has been estimated by ITS for each of ten measurement frequencies in terms of the upper and lower deciles, D_u and D_l(in dB, relative to the median value). These values, summarized in Table 1, are the root-mean-squares of all the location values for each frequency and environmental category.

Table 1. Summary of selected measured noise parameters for business, residential and rural environmental categories

Environmental Category	Freq. (MHz)	F_{am} (dB(kT_o))	D_u (dB)	D_l (dB)	σ_{NL} (dB)
Business	0.25	93.5	8.1	6.1	6.1
	0.50	85.1	12.6	8.0	8.2
	1.00	76.8	9.8	4.0	2.3
	2.50	65.8	11.9	9.5	9.1
	5.00	57.4	11.0	6.2	6.1
	10.00	49.1	10.9	4.2	4.2
	20.00	40.8	10.5	7.6	4.9
	48.00	30.2	13.1	8.1	7.1
	102.00	21.2	11.9	5.7	8.8
	250.00	10.4	6.7	3.2	3.8
Residential	0.25	89.2	9.3	5.0	3.5
	0.50	80.8	12.3	4.9	4.3
	1.00	72.5	10.0	4.4	2.5
	2.50	61.5	10.1	6.2	8.1
	5.00	53.1	10.0	5.7	5.5
	10.00	44.8	8.4	5.0	2.9
	20.00	36.5	10.6	6.5	4.7
	48.00	25.9	12.3	7.1	4.0
	102.00	16.9	12.5	4.8	2.7
	250.00	6.1	6.9	1.8	2.9
Rural	0.25	83.9	10.6	2.8	3.9
	0.50	75.5	12.5	4.0	4.4
	1.00	67.2	9.2	6.6	7.1
	2.50	56.2	10.1	5.1	8.0
	5.00	47.8	5.9	7.5	7.7
	10.00	39.5	9.0	4.0	4.0
	20.00	31.2	7.8	5.5	4.5
	48.00	20.6	5.3	1.8	3.2
	102.00	11.6	10.5	3.1	3.8
	250.00	0.8	3.5	0.8	2.3

3. CCIR REPORT 258 MODELS

The environmental categories for which predictions are available in CCIR Report 258-5 include: business, residential, rural, and quiet rural[3]. Business areas are defined as any area where the predominant usage throughout the area is for any type of business(e.g. stores and offices, industrial parks, large shopping centers, main streets or highways lined with various business enterprises, etc.). Residential areas(urban or suburban) are defined as any area used predominately for single or multiple family dwellings with a density of at least two single family units per acre and no large or busy highways. Rural areas are defined as locations where land usage is primarily for agricultural or similar pursuits, and dwellings are no more than one every five acres. Quiet rural areas are defined as locations chosen to ensure a minimum of man-made noise.

In all cases results are consistent with a linear variation of the median value of F_a, F_{am}, with frequency of the form:

$$F_{am} = c - d \log f, \qquad (1)$$

where f is the frequency expressed in MHz. The constants c and d are given in Table 2. As these results are based on the work of Spaulding and Disney[2], equation (1) is valid only in the range 0.25 to 250 MHz for all the environmental categories except quiet rural and galactic noise. The formula for galactic noise from radio stars which is incident on the ionosphere is included here for comparison only. Note that Table 2 also contains data from Spaulding and Disney for parks and university campuses and for inter-state highways.

Table 2. Values of the Man-made Noise Constants c and d[2]

Environmental category	c	d
Business	76.8	27.7
Inter-state highways	73.0	27.7
Residential	72.5	27.7
Parks and university campuses	69.3	27.7
Rural	67.2	27.7
Quiet rural	53.6	28.6
Galactic noise	52.0	23.0

Skomal has reviewed man-made noise data collected in a range of countries at various distances from metropolitan areas[10-11]. Results cover the frequency range 500 kHz to 1 GHz. In the frequency interval 100 to 800 MHz, he showed that the frequency decrement moderates from that given in Table 1 to -10 to -15 dB/decade. This is consistent with the presence of a localized maximum in the UHF-band emission spectrum of vehicular ignition interference.

At four sites in downtown Ottawa, Canada, measurements of the VHF-UHF radio environment were carried out over a 17-day period in November 1976[5]. A linear regression equation of the frequencies 200 to 500 MHz was given as $F_{am} = -15.8 \log f + 48.4$(f in MHz). In the frequency range from 200 to 300 MHz, the results using this equation compare favorably with those using the business area equation of [2].

Man-made noise level measurements were made of the UHF radio environment over a four-month period(1982-1983) in and around

Ottawa, Canada by Lauber and Bettrand[6]. The antenna noise temperature T_a was measured at sites typical of business, residential, and rural areas at frequencies from 600 to 950 MHz. For each of these areas, a minimum of 1800 one-second measurements were combined to form a cumulative probability distribution, which shows the time and location variability of the noise. The highly skewed shape of the distribution, especially those from the business areas, showed that the UHF radio environment is composed of a background noise level upon which is superimposed a highly variable man-made noise level from vehicle ignition noise.

4. NOISE LEVEL DISTRIBUTION MODELS

Hagn and Sailors[12] have presented four statistical models of increasing complexity(simple Gaussian, composite Gaussian, chi-suare, Gaussian from chi-square) which utilize the time and location variability of F_a to predict the accumulative probabilities of man-made radio noise available power levels for short, vertically-polarized antennas located near the ground. These models, which have now been included in CCIR Report 258, are useful in predicting the probability that the short-term signal-to-noise ratio for a given communication system equals or exceeds a value required for successful communication.

For the simple Gaussian model, the mean is approximated by F_{am} from Table 1, and the standard deviation σ_N is given by

$$\sigma_N = (\sigma_{NL}^2 + \sigma_{NT}^2)^{\frac{1}{2}} \qquad (2)$$

assuming that the temporal and spatial variabilities are uncorrelated. The parameter σ_{NT} is the standard deviation of the temporal variability and values are obtained from D_u and D_l of Table 1 using

$$\sigma_{NT} = \frac{1}{1.28}\left[\frac{(D_u^2 + D_l^2)^{\frac{1}{2}}}{2}\right] \qquad (3)$$

The location variability σ_{NL} is given in Table 1.

The composite Gaussian model is the simplest model which takes into account skewness in the distribution. For this model the standard deviations for the upper and lower halves of the distribution are given by σ_{Nu} and σ_{Nl}, respectively, and the mean is given by F_{am}. The corresponding upper and lower deciles for these distributions are obtained from σ_{Nu} and σ_{Nl}, respectively, using the equations

$$D_{Nu} = 1.28\ \sigma_{Nu} \qquad (4)$$

$$D_{Nl} = 1.28\ \sigma_{Nl} \qquad (5)$$

The standard deviations themselves are computed

$$\sigma_{Nu} = \left[\left(\frac{D_u}{1.28}\right)^2 + \sigma_{NL}^2\right]^{\frac{1}{2}} \qquad (6)$$

$$\sigma_{Nl} = \left[\left(\frac{D_l}{1.28}\right)^2 + \sigma_{NL}^2\right]^{\frac{1}{2}} \qquad (7)$$

The chi-suare model also takes into account skewness. In the chi-square model the parameter ν is the number of degrees of freedom. The chi-square model satisfies the relationship $F_a = a + b\,\chi^2$. The mean noise is $\langle F_a \rangle = F_{a\chi}{}^2 = a + b\,\nu$, and its standard deviation is $\sigma_{N\chi}{}^2 = b(2\,\nu)^{\frac{1}{2}}$. The parameters $a, b, \nu, F_{a\chi}{}^2$, and $\sigma_{N\chi}{}^2$ are given in Table III of CCIR Report 258-5[3] for the ten discrete frequencies between 0.25 and 250 MHz. These parameters were found using a method developed by Zacharisen and Crow[13].

A second Gaussian model using parameters can be estimated from the chi-square approximations. For this model the mean is given by $F_{a\chi}{}^2$ and the standard deviation $\sigma_{N\chi}{}^2$.

Figure 4 shows the predictions for 20 MHz in a residential environment. All four models predict very similar values between the deciles. In the tails of the distribution(beyound the deciles), the simple Gaussian model predicts the lowest levels; and the chi-square model predicts the highest levels.

Hagn and Sailors[12] also made a preliminary comparison of the models with measured data. Figure 5 is a plot of the model predictions of Figure 4 superimposed on the data of Figure 3. Since there was no location variability data for this sample, σ_{NL} was set to zero for the comparison. By inspection, the best fit over the interval 1% to 80% is given by the Gaussian model with parameters estimated from the chi-square. The simple Gaussian model predictions are consistently too low; and, in the more interesting upper half of the distribution, the chi-square model predictions are good over the interval 10% to 80% but too high for the upper tail of the distribution. The shape of the measured distribution in the lower tail is most closely approximated by the chi-square.

Hagn and Sailors applied a Kolmogorov-Smirnov(K-S) goodness-of-fit test[14-16] to examine more rigorously how well distributions of sample man-made noise data conform to the hypothesized distributions. The results for data and model predictions of Figure 5 show that the chi-square model is accepted for a significance level of 5% for values of probability of exceeding the ordinate up to 90 percent, whereas, the simple Gaussian model does not fit well throughout the entire distribution. The K-S test was also applied to 20 MHz data obtained for 31 minutes in a Boulder, CO, business area. This particular distribution sample was the only one available which was not used to develop the model. The distribution was slightly skewed as can be seen in Figure 6. As the mean for this sample was 49.7 dB(kT_o) compared to 40.8 dB(kT_o) from Table 1; clearly, all the model predictions would be too low. In this case the authors set σ_{NL}= 0 and adjusted the median of the simple Gaussian model to 49.2 dB(kT_o) for the comparison. For a significance level 5%, the Gaussian model with parameters estimated using the chi-square was overwhelmingly rejected. At this significance level, it was found that it would be incorrect to reject the simple Gaussian, composite Gaussian, and chi-square models. The final conclusions were

that the models were most useful in the probability interval 0.1 to 0.9, but further checks against measured data are needed to determine the limits of their applicability.

5. OTHER EMPIRICAL MODELS

There are other useful empirical models that either supplement the CCIR Report 258 models or in one case replace it. This includes: (1) A model developed for use at frequencies in the VHF and UHF bands; (2)A model known as the quasi-minimum atmospheric noise model; and (3)An airborne man-made radio noise model.

The first of these is a model that was developed for use at VHF and UHF by Hagn et al.[7]. For frequencies between 1 MHz and 200 MHz, the model given in CCIR Report 258-8 is used. However, at frequencies above 200 MHz, the following equations for F_{am} in dB(kT_o) are used:

$$F_{am} = 49.4 - 15.8 \log f \quad (8)$$

$$F_{am} = 45.2 - 15.8 \log f \quad (9)$$

$$F_{am} = 39.2 - 15.8 \log f \quad (10)$$

for business, residential, and rural environments, respectively, where the frequency f is in MHz. At 200 MHz, this model and CCIR Report 258-5 produce identical results. Equation (8) differs only 1 dB from that of Lauber and Bertrand[5] at 200 MHz for business areas. The slope of these curves is identical to that of Lauber and Bertrand at these higher frequencies. The above equations extend the noise models upward in frequency until the man-made and/or galactic noise drops into the internal noise of the receiving system. Lauber and Bertrand[6] have compared this model along with one due to Skomal[11]. The results are presented in Figure 7. The Hagn et al.[7] results are within $\pm$3 dB with the exception of the higher frequencies for the rural areas.

The Hagn et al. model also provides approximations for the standard deviations for the simple Gaussian model of Hagn and Sailors[12]. For frequencies between 1 MHz and 100 MHz, these standard deviations are given as 10.5, 8.5, 6.5, 4.5, and 1 dB for business, residential, rural, quiet rural, and galactic environments, respectively. For frequencies above 100 MHz, the values are based on the more limited data of Lauber and Bertrand[5]. The standard deviation σ_N, in dB, is given by:

$$\sigma_N = 10.5 - 9 \log(\tfrac{f}{100}), \text{ or 2 dB} \quad (11)$$

$$\sigma_N = 8.5 - 9 \log(\tfrac{f}{100}), \text{ or 2 dB} \quad (12)$$

$$\sigma_N = 6.5 - 9 \log(\tfrac{f}{100}), \text{ or 2 dB,} \quad (13)$$

respectively, for business, residential, and rural man-made noise environments. These equations are useful for $\sigma_N \pm 0$ dB. The standard deviation σ_N is not allowed to go negative at higher frequencies but merely drops to 2 dB and stays constant for higher frequencies until F_{am} decreases to 0 dB(kT_o).

Another noise model often used as a man-made noise model in some applications is one known as the quasi-minimum atmospheric noise model[17]. It is based on a comprehensive examination of expected noise at many locations and for all seasons using data from the National Bureau of Standards noise measurement program, and shipboard measurements made at sea in the San Diego,CA area. Table 3 is a list of the values for this model. A fit to these data is given by

$$F_{am} = 60.0 - 27.5 \log f \quad (14)$$

where f is in MHz. Equation (14) is also given in Table 3. Note that the median rural noise equation is nearly equal to the quasi-minimum atmospheric noise equation. Results for rural noise equation are also presented in Table 3.

Measurements indicate that airborne man-made radio noise from a distant metropolitan area can be detected once an aircraft rises above the local optical horizon. Above 10,000 feet measurements show a broad noise signature representative of an entire metropolitan area. Roy has reported an airborne man-made radio noise model developed to evaluate the effect of man-made radio noise on the operation of meteor burst communication systems[18-19]. Equations developed by Skomal[11] were used to construct the model. Two parametric equations were used to model the height gain of man-made radio noise as a function of distance, 0 to 150 miles, from the source. Coefficients for these equations were calculated from data measured over Seattle[20]. Roy used two hundred of the nation's largest cities and 62 of the largest counties and military installations as sources of radio noise. Day and nighttime contours were produced in the 25 to 75 MHz range for altitudes between 30 and 70 thousand feet. These maps show that very little of the continental United States is free of airborne man-made noise. Minimum noise levels are found during the night at low altitudes for distances greater than 100 miles from most metropolitan areas. As an example Figure 8 shows daytime 45 MHz contours for an altitude of 5 thousand feet. Contours of constant radio noise power in dB above kTb are plotted for values of 15,20, and 25 dB. Shaded areas in the continental United States represent areas containing noise power 3 dB or less above galactic noise.

Table 3. Quasi-minimum atmospheric noise levels in dB above kT_o

Frequency (MHz)	Level Chase & Tirrell[17]	Equ.(14) (Quasi-Minimum)	Equ. (1) (Rural)
2	52	51.72	58.86
4	44	43.44	50.52
6	39	38.60	45.65
8	36	35.17	42.18
10	36	32.50	39.50
12	31	30.32	37.31
15	28	27.66	34.62
20	25	24.22	31.16
25	22	21.56	28.48
30	20	19.38	26.28

6. RECENT MEASUREMENTS

When one uses the models of CCIR Report 258, it is necessary to decide which environmental category a receive site belongs. This author had an opportunity to determine the environmental categories for the receive sites of an HF communication system. For this purpose he acquired man-made noise data measured on CDAA antennas at the receive sites of the communication stations at Adak,AL; Diego Garcia; Guam; Honolulu,HI; San Diego, CA; and Stockton,CA.

These measurements were made by the Naval Electronics Engineering Activity, Pacific(NEEAP) on the Automated AN/FRM-19(v) Test System. This system includes an Automated Noise Measurement System(ANMS). The ANMS allows the site to measure its baseline/strong signal level on a periodic basis. The ANMS acquires data every two hours at seven frequencies over a seven day period. This data is then averaged over this period, and a baseline noise/strong signal level is found. The ANMS acquires data data from a 0.5 MHz window centered on the seven frequencies(usually 2.5,5.0,7.5,10.0,12.5,20.0,30.0 MHz). The ANMS searches for the lowest power level or "hole" in this 0.5 MHz window. This 0.5 MHz window at a frequency of interest is divided into 83 channels of 3.6 kHz. The ANMS then samples the noise waveform in each channel by taking 50 consecutive points. By taking 50 samples, a window is set up in the time domain that will capture time related RFI noise, such as powerline noise. After the ANMS scans the 0.5 MHz window, it returns to the "hole" to measure the system noise level. The average power(true rms voltage) and V_d(voltage deviation) are then calculated from the digitized waveform. After the ANMS has collected the noise data for seven days, an operator can then proceed to plot out the data using different plotting modules to analyze the noise environment.

Baseline noise data was provided by NEEAP for the forenamed communications sites measured by the ANMS for the local noon period. Local noon was chosen as the most likely time of day that man-made noise would most likely be present. The data for Guam, measured from September 17, 1987 through September 24, 1987, was examined first because that was the only site for which there was other measured results available for comparison[9]. Figure 9 shows this data measured by the Stanford Research Institute(SRI) along with the data measured by NEEAP on a CDAA antenna. The SRI data was measured on a short vertical rod. For comparison purposes both quiet rural and galactic noise are included. Clearly there is a large difference between these two measurements part of which is due to the measurements by NEEAP on a CDAA type antenna rather than the standard short vertical rod.

George Hagn of SRI International provided this author with some conversion factors which he obtained while making noise measurements in Iceland[21]. Mr Hagn made noise measurements at the edge of the ground screen of the CDAA antenna with a calibrated nine foot vertical rod and then made noise measurements through the CDAA antenna to compare to the other measurements. He developed a rough set of corrections that could be applied to noise measurements made on the omni-beam output of the CDAA antenna. These are given in Table 4. The way they are used is that first the noise measurements in dBm are converted to dB above 1 micro-volt by adding 107(-107 dBm=0 micro-volts across a 50 ohm resistance). Then the antenna correction factor is added. Finally, the noise field strength is converted to F_a by using the relationship between F_a and noise field strength found in CCIR Report 322-3[1] taking into account the receiver bandwidth, which in this case was 3.6 kHz.

Table 4. CDAA to 9 Foot Rod Conversion Factors

Frequency (MHz)	Antenna Factor dB/1 Meter
2.0 - 2.5	-3,± 6 dB
2.5 - 8.0	-23,± 8 dB
8.0 -13.0	-15,± 8 dB
13.0-30.0	-14,± 8 dB

Figure 10 shows the results of appling these antenna factors to the NEEAP measurements. The curve marked Guam-NEEAP is the result of appling the antenna factor disregarding the error term. The curve marked Guam-NEEAP Adjusted has had the upper limit of the error term added to the antenna factor for 5, 20 and 30 MHz. Although the Guam-NEEAP curve is an improvement over that show in Figure 9 for this case, the Guam-NEEAP Adjusted is an even better improvement. Based on this comparison, the measured noise at the other five sites was adjusted using Table 4; for 5,7.5, 20, and 30 MHz, the upper error, 8 dB, was added to the antenna factor. Figure 11 shows the results of the NEEAP measurements with these antenna factors applied. Curves for rural and quiet rural and galactic noise are show for comparison.

In addition to providing man-made noise measurement data for Guam, Shepard et al.[9] also show the results of measurements made at Keflavik, Iceland; Rota, Spain; and Bremerhaven, Germany. These results are repeated here as Figure 12.

7. DISCUSSION

Engineers and operational analysts who use these CCIR man-made models are faced with the problem of determining whether any environment is most like a U.S. business area, a U.S. residential area, or a U.S. rural area over a decade ago. The only other alternative is to assume the environment is similar to a quiet rural area near one of the quiet stations where the atmospheric noise data were gathered over three decades ago. Hagn has provided additional insight into this problem[22].

One problem already encountered in this paper is the effect of the receive site antenna on the measured or estimated noise values. In fact in the case encountered here additional data is needed comparing CDAA antenna and short vertical rod noise measurements. Hagn and Shepherd[23] have provided some insights into this problem of the effect of different antennas on noise estimation and

measurement; however, the problem is not solved.

In the discussion of the simple Gaussian model, it was assumed that the median value of man-made noise Fam was an efficient estimator of the mean. Sailors[24] has reviewed techniques for estimating the mean and standard deviation of a parameter from its quantiles. Results using these techniques were compared to that obtained from estimating the sample means and standard deviations from data samples of man-made noise. For man-made noise modeling, it was determined that the mean could be estimated to an accuracy of 0.1 dB using the expression $F_{am} + 0.237(D_u - D_l)$. In the example given in Figures 4 and 5, the median noise was 36.46 dB. The value of the mean using this expression is 37.43 dB. This compares to 37.5 dB given by the simple Gaussian obtained from the chi-square.

8. REFERENCES

1. CCIR,"Characteristics and Applications of Atmospheric Radio Noise Data," Report 322-3, International Radio Consultative Committee, International Telecommunication Union, Geneva, Switzerland, 1986.

2. Spaulding, A.D. and R.T. Disney, "Man-Made Radio Noise I: Estimates for Business, Residential, and Rural Areas," OT Report 74-38, Office of Telecommunications, U.S. Department of Commerce, Boulder, CO, June 1974.

3. CCIR, "Man-Made Noise," Report 258-5, International Radio Consultative Committee, International Telecommunication Union, Geneva, Switzerland, 1990.

4. Spaulding, A.D., "Man-Made Noise: The Problem and Recommended Steps Toward Solution," OT Report 76-85, Office of Telecommunications, U.S. Department of Commerce, Boulder, CO, April 1976.

5. Lauber, W.R. and J.M. Bertrand, "Preliminary Urban VHF/UHF Radio Noise Intensity Measurements in Ottawa, Canada," Second Symposium and Technical Exhibition on Electromagnetic Compatibility, Montreux, Switzerland, 28-30 June 1977.

6. Lauber,W.R. and J.M. Bertrand, "Man-made Noise Level Measurements of the UHF Radio Environment," Symposium Record, 1984 IEEE National Symposium on EMC, San Antonio, TX, 14-26 April 1984.

7. Hagn, G.H., B.M. Sifford and R.A. Shepard, "The SRICOM Probabilistic Model of Communication System Performance--A User's Manual for Engineers, Applications Programmers and Systems Programmers," Final Report, SRI Project 3603, Contract NT-81-RC-16011, SRI International, Arlington, VA, May 1982(with errata).

8. "Automated AN/FRM-19(V) Test System Operation and Maintenance Manual," Version 3, Naval Electronics Engineering Activity, Pacific, Pearl Harbor, HI, October 1987.

9. Shepherd, R.A., J.B. Lomax, V.D. Cone, D.L. Nielson, and G.C. Edwards, "MF and HF Man-Made Radio-Noise and Interference Survey--Gaum," Final Report, SRI Project 3328, Contract N00039-74-C-0292, Stanford Research Institute, Menlo Park, CA, June 1974.

10. Skomal, E.N., "An Analysis of Metropolitan Incidental Radio Noise Data," IEEE Trans. Electromagn. Compat., Vol. EMC-15, No. 2, pp. 45-57, May 1973.

11. Skomal, E.N., Man-Made Noise, Van Nostrand Reinhold Co.,New York, NY, 1978.

12. Hagn, G.H. and D.B. Sailors, "Empirical Models for Probability Distributions of Short-term Mean Environmental Man-made Radio Noise Levels," Electromagnetic Compatibility 1979, Proceedings of 3rd Symposium and Technical Exhibition on Electromagnetic Compatibility, Rotterdam, pp. 355-360, May 1-3, 1979.

13. Zacharisen, D.H. and E.L. Crow, "Fitting Distributions of Telecommunication Variables with Chi-square Distributions," Radio Science, Vol. 5, pp. 1307-1315, November 1970.

14. Massey, Jr., F.J.,"The Kolmogorov-Smirnov Test for Goodness of Fit," J. Amer. Statist. Assn., Vol. 46, pp. 68-78, March 1951.

15. Smirnov, N., "Table for Estimating the Goodness of Fit of Empirical Distributions," Annals of Maths. Stat., Vol. 19, pp. 279-281, 1948.

16. IBM Corporation, "Nonparametric Statistics," System 360 Scientific Subroutine Package, Version III, IBM, White Plains, NY, 63-64, August 1970.

17. Chase, W.M. and C.W. Tirrell, "TRED HF Communication Systems Analysis," Naval Ocean Electronics Laboratory Center Tech. Rep. 1786, San Diego, CA, 1971.

18. Roy, T.N., "Airborne Man-made Radio Noise Assessment," NOSC TR 677, Naval Oceans Systems Center, San Diego, CA, 15 April 1981.

19. Roy, T.N., "Airborne Man-made Radio Noise Assessment," 1982 IEEE International Symposium on Electromagnetic Compatibility, Santa Clara, CA, 8-10 September 1982.

20. Buehler, W.E. and C.D. Lunden, "Signature of Man-made High-Frequency Radio Noise," IEEE Trans. Electromagn. Compat., Vol. EMC-8,pp. 143-152,1966.

21. Hagn, G.H., "Calibration of Nine-Foot Rod Antenna for MF and HF EMC Measurements and Comparison with AN/FRD-10 Antenna: Volume I: Groundwave," Final Report, Contract N00039-77-C-0349, SRI Project 6792, Stanford Research Institute, Menlo Park, CA, September 1978.

22. Hagn, G.H., "Man-made Radio Noise and Interference," AGARD Conference Proceedings 420, Effects of Electromagnetic Noise and Interference on Performance of Military Radio Communication Systems, Lisbon, Portugal, 26-30 October 1987.

23. Hagn, G.H. and R.A. Shepherd, "Selected Radio Noise Topics," Final Report, SRI Project 5002, Contract No. NT 83 RAC-36001, SRI International, Arlington, VA, June 1984.

24. Sailors, D.B., "Estimation of the Mean and Standard Deviation from Quantiles in Interference Modeling," 1980 IEEE International Symposium on Electromagnetic Compatibility, Baltimore, MA, 7-9 October 1980.

9. ACKNOWLEDGEMENT

The author would like thank Brian Kutara of the Naval Electronics Engineering Activity, Pacific, Pearl Harbor, HI for providing the noise data for the six Pacific communication stations. He would also like thank Mr. George Hagn of SRI International, Arlington, VA for providing him the information on converting noise measuroments for the CDAA antenna to a 9 foot rod.

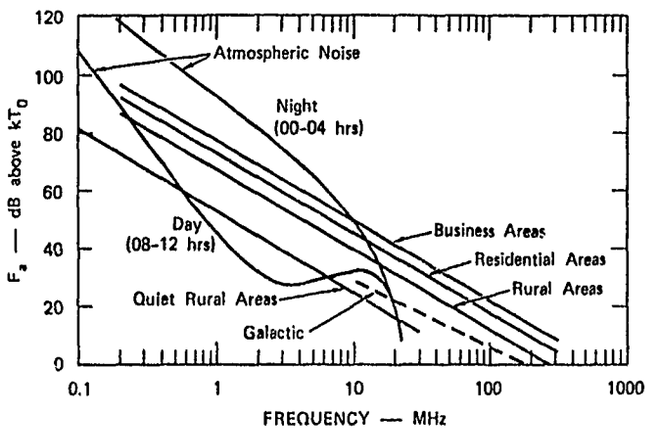


Figure 1. Estimates of median values of man-made, atmospheric, and galactic noise expected near Washington, D.C. during summer[1-2].

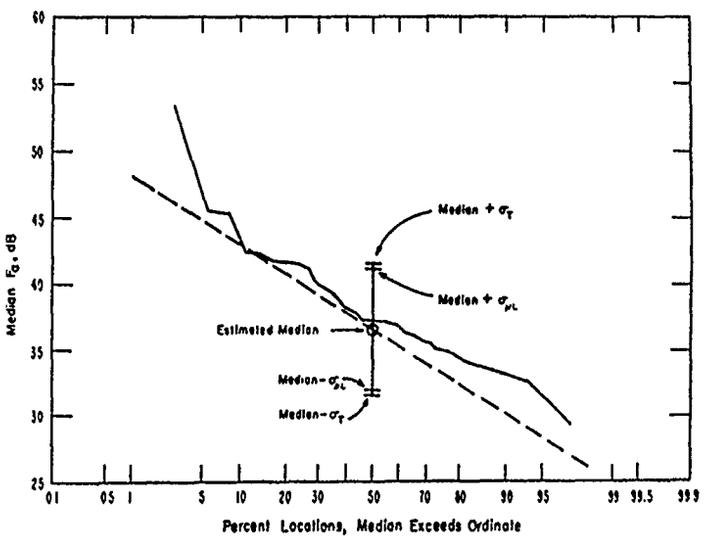


Figure 2. Distribution of location median of F_a values for man-made noise in residential areas at 20 MHz[2].

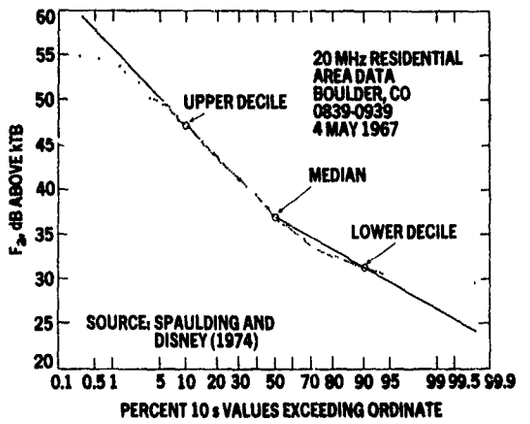


Figure 3. Example distribution of F_a at one location during one hour with short-term median and deciles indicated[2].

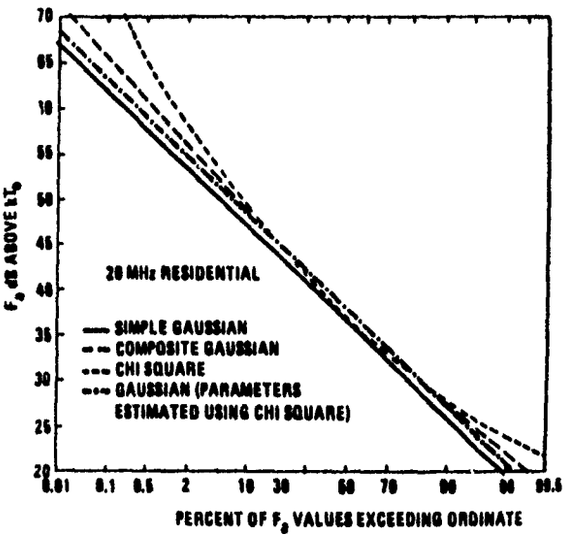


Figure 4. Comparison of model predictions for 20 MHz in a residential area[12].

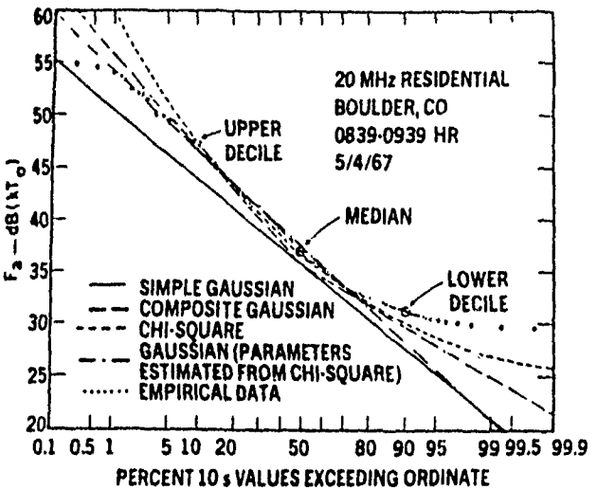


Figure 5. Example comparison of models and data[12].

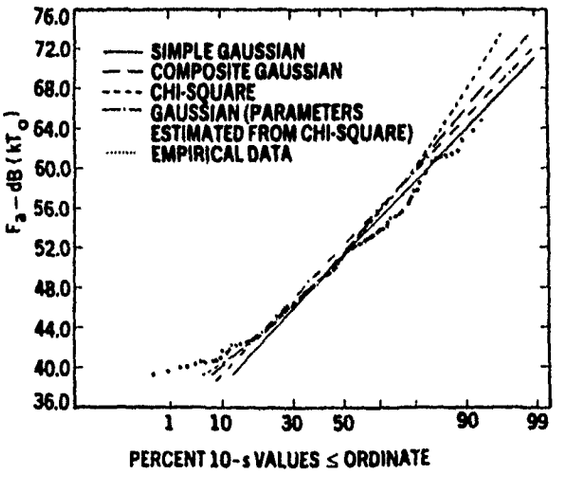


Figure 6. Comparison of models and data: business[12].

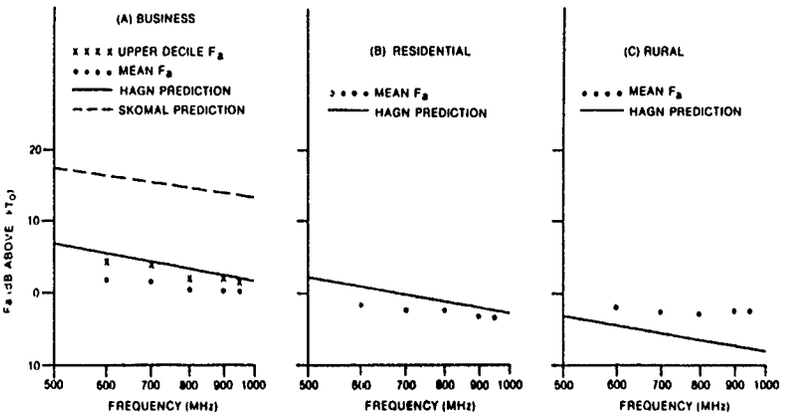


Figure 7. Comparisons of measurements with recent prediction models[6].

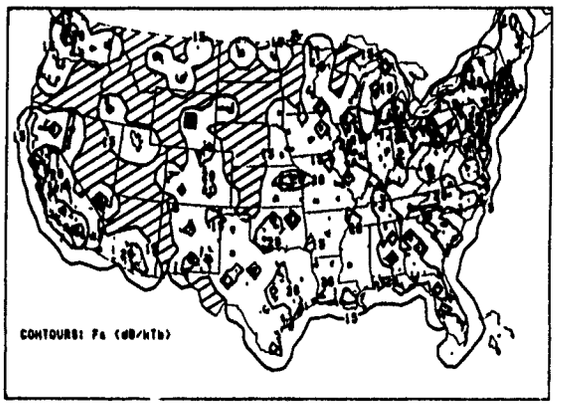


Figure 8. Daytime 45 MHz airborne man-made radio noise map of the continental United States for an altitude of 5 thousand feet[18-19]. Shaded areas represent noise power 3 dB or less above galactic noise.

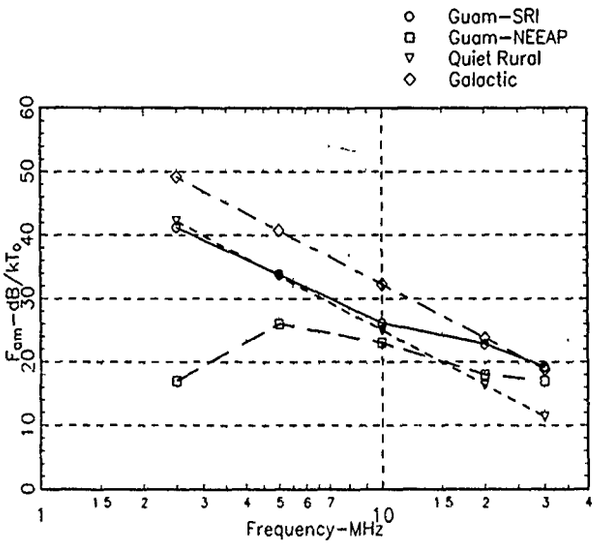


Figure 9. Comparison of man-made noise measurements at Guam made by Stanford Research Institute(SRI) on a short vertical rod with data measured by the Naval Electronics Engineering Activity, Pacific(EEAP) on a CDAA antenna.

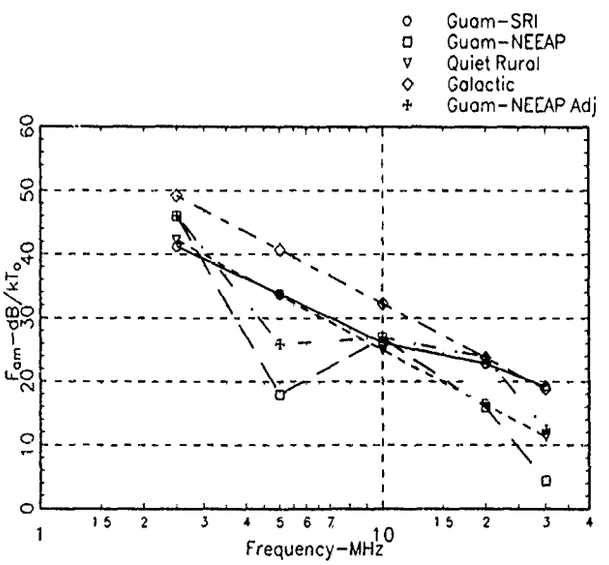


Figure 10. Comparison of man-made noise measurements at Guam made by Stanford Research Institute(SRI) with data measured by the Naval Electronics Activity, Pacific(NEEAP).

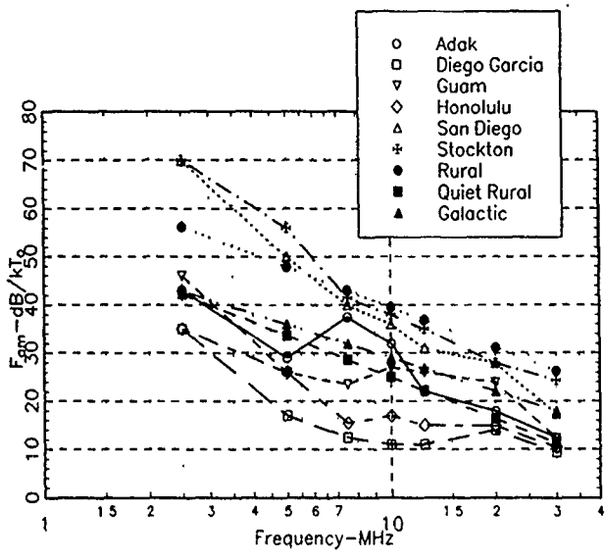


Figure 11. Man-made noise environment at sites in the Pacific Ocean.

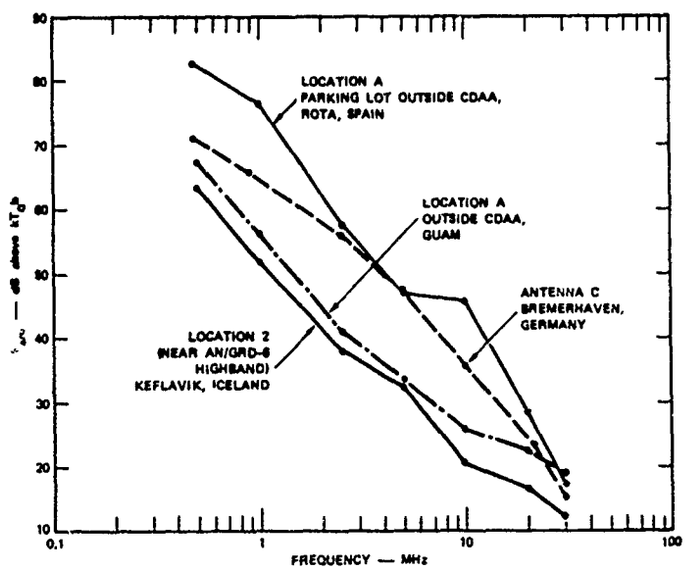


Figure 12. A comparison of F_{am} at quiet locations[9].

DISCUSSION

John S. BELROSE

I wonder what sort of "ground-plane" was used with the 9-foot rod that you used to make the radio noise measurements ? Stimulated by a CCIR interest to measure HF field strenghts imploying a short vertical whip, we (at the Communications Research Centre, Ottawa) have made a study (which is still ongoing) to determine how the antenna factor depends on ground conductivity and on the type of "grounding". We have made measurements using a short active whip fed against a support rod. The support rod was : 1) grounded by a ground stake ; 2) fed against a radial ground screen, having radials of various numbers and lengths, and elevated at different heights ; 3) isolated from ground (a dipole) ; and 4) employing an Uffer ground, a ground rod enclosed in concrete. The concrete remains damp and provides a more stable ground, than a ground rod. The antenna factor depends on all of the above parameters.

AUTHOR'S REPLY

One attempts to have a short antenna over perfect ground. In the case of the 9 foot vertical rod referred to in the paper, Shephard et al. [9] used a nine foot rod on a 2-ft diameter aluminum base ; this itself was then mounted on top of a van. SRI made a considerable effort to verify that the antenna was in fact a short antenna over perfect ground.

ADDED COMMENT - George H. HAGN

The 9-ft rod on with a 2-ft aluminum base mounted on a vehicle roof was calibrated for 2 vehicles in 1972 in Iceland. The groundwave antenna factor was measured to be -2,8dB(m^{-1}) for a small van and - 1,6dB(m^{-1}) for a small Volkswagen for the entire HF band. One should always calibrate the antenna as used, wherever possible to reduce this source of error ?

R.M. HARRIS, UK

You mentioned a report which studied man-made noise at altitude. Do you know if they measured conventional atmospheric noise as a function of altitude ? I do not know of any such work having seen reported, do you know of any ?

AUTHOR'S REPLY

In the frequency range from 25 to 70MHz, covered in the work of Roy [18-19], atmospheric noise is not as important a source as is galactic noise and man-made noise. John Herman may have reported some studies of atmospheric noise at satelite heights and a person who I think is named Art Mille of Convair Corporation may have reported the measurement of noise measured by an airplane as it approached Akron, Ohio. In the latter study the noise measurements were peak noise rather than r.m.s. noise.

ADDED COMMENT - George H. HAGN

Atmospheric noise from lightning in the frequency range 25-70MHz propagates primarily by line-of-sight (LOS) from active storm cells. Its variation with altitude should be small once the LOS geometry is achieved.

E. ANAGNOSTAKIS

Mr Sailors, could you, please, elaborate a little more on the choice of the number of degrees of freedom in the case in which you employ the chi-square technique for determining the gaussian parameters of your interpretation ?

AUTHOR'S REPLY

In the case in which the chi-square technique is employed to obtain the Gaussian parameters, the degrees of freedom are determined exactly as would be used to find the degrees of freedom for the chi-square model. Reference [13] given in the paper gives further details on how to obtain the degrees of freedom for the chi-square given the median and upper and lower decile. An algorithm to do this is contained in the HFMUFES IV HF predict' n program.

TOWARD A GLOBAL MODEL OF HF OTHER-USER INTERFERENCE

G.H. Hagn
Information & Telecommunication Sciences Center
SRI International
1611 North Kent Street
Arlington, VA 22209

ABSTRACT

Information on other-user interference (sometimes called co-channel interference) is important in predicting the performance of HF systems. For many systems, especially frequency-hopping communications systems, other-user interference can be a greater problem than man-made noise from incidental radiators (e.g., powerlines) or atmospheric noise from lightning. We currently have a set of local environment models for man-made noise (CCIR Report 258) for business, residential, rural and quiet rural areas and a worldwide model for atmospheric noise from lightning (CCIR Report 322). However, there are no worldwide or regional models for HF other-user interference. This paper will describe a method for development of a regional model for HF other-user interference in terms of congestion for allocated bands (e.g., fixed, mobile, amateur, etc.). Congestion is defined as the fraction of the channels in an allocated band of contiguous channels with detected energy exceeding a specified threshold. A global congestion model might eventually be achieved as a combination of regional models.

Occupancy of individual channels and congestion of allocated HF bands vary with type of service, frequency, time of day, season, angle of arrival, type of receiving antenna, receiver bandwidth, threshold, geographic location and sunspot number. Researchers at UMIST, Manchester, UK, have noted the large variability of the occupancy of individual 1-kHz channels; however, by aggregating contiguous 1-kHz measurements across individual allocated bands, they have also noted the stability of the statistics of HF bands and congestion for a given time of day and season (e.g., winter, noon). UMIST also has successfully modeled band congestion at one location (Gott, et al., AGARD-CP-420, Lisbon, 1987). A method of extrapolating the empirically-based UMIST statistical model of band congestion from a point (e.g., one receive site in the UK) to a region (e.g., Europe) will be described. The model coefficients would be derived from measurements at locations in NATO countries separated by about 500 km (or a distance determined by a correlation test). The coefficients for the regional model for a given arbitrarily chosen location, would be derived by interpolation. The applicability of the regional modeling approach to the development of a global model for HF band and congestion (analogous to the CCIR Report 322 worldwide model for atmospheric noise from lightning) will be discussed.

DISCUSSION

ICA FUERXER

Why we do not define a model based on the description of a transmitter, the propagation and the activity to find the head lines of the problem to solve. It could be the opportunity to specify the measuring system specifications and the way to validate the model in Europe.

AUTHOR'S REPLY

I believe that a propagation model can be used to gain insights by running it for known transmitter locations and characteristics. I do not think that the International Telecommunication Union has an accurate or complete enough data base of HF assignments and equipments to generate the basic model. An empirical data base is needed to define the model that is the subject of this paper.

C. GOUTELARD, FR

Il s'agit d'un commentaire qui s'inscrit dans la discussion présente. Tout d'abord il me semble que les prévisions des interférences doivent être à court terme, car aux variations ionosphériques s'ajoutent les variations dues aux utilisateurs et dont la variabilité est bien plus grande.

Ensuite, il me semble que l'on a de bonnes raisons d'être optimiste sur la stabilité du spectre. En effet si l'on doublait, ou si l'on divisait par 2 le nombre d'émetteurs existant sur terre on n'obtiendrait qu'une variation de 3dB du niveau moyen du spectre HF.

Enfin, je pense que les études à mener ne doivent pas être menées en s'interrogeant sur le coût d'un tel système. Il faut le faire de toute façon.

AUTHOR'S REPLY

The cost must be considered when determining if a regional (or global) model is practically achievable. The distance correlation data available now indicate that sampling can be done at a spacing greater than 500 km. Therefore, a regional model for Europe seems practical from an economic stand point.

MODELISATION SPATIO-TEMPORELLE DES INTERFERENCES ELECTROMAGNETIQUES H.F. EN EUROPE OCCIDENTALE

T. CANAT, J. CARATORI, C. GOUTELARD

LETTI - Université Paris-Sud - Orsay - France

- oOo -

I. - INTRODUCTION -

L'encombrement spectral de la gamme des ondes décamétriques prouve son exploitation permanente par de nombreux utilisateurs. Dans ce spectre, qui s'étend de 3 à 30 MHz, l'importance des interférences est très grande à cause des puissances utilisées - certains émetteurs ont des puissances d'émission de 1MW - mais aussi à cause des conditions de propagation qui permettent de grandes portées avec de faibles pertes de puissance.

La qualité d'une liaison est caractérisée par son rapport signal/(bruit + interférences). On s'est beaucoup penché, durant les années passées, sur le calcul et la prévision de la puissance du signal. Les caractéristiques variables de l'ionosphère, liées à l'indice d'activité solaire dont on possède des valeurs suffisamment précises depuis 1730, ont incité les chercheurs à définir des méthodes de prévision. Ces méthodes dont l'efficacité est connue permettent, avec des temps de préavis variables, de donner satisfaction aux utilisateurs.

Cependant, la qualité d'une liaison dépend tout autant du bruit et des interférences, même si ce terme apparaît au dénominateur d'un rapport. Le bruit et plus encore les interférences dépendent des conditions ionosphériques et sont donc fonction du temps et de l'espace. D'autres causes de variations existent, liées notamment à la densité des émetteurs dans la zone de plusieurs milliers de kilomètres qui entoure le lieu de réception. Ainsi, l'encombrement spectral est beaucoup plus important en Europe qu'en Amérique, qu'au coeur de l'océan pacifique...

Une vision plus large des prévisions ionosphériques consiste à prévoir, non plus seulement la puissance du signal reçu, mais la valeur complète du rapport signal/(interférences + bruit), qui, comme il vient d'être dit, ne dépend plus uniquement que de l'activité solaire, mais également de l'activité des autres utilisateurs.

La prévision de l'activité des interférences constitue une partie de cette présentation, mais cette prévision doit être complétée, pour les systèmes modernes, par une évaluation des variations spatiales des interférences.

L'utilisation d'antennes directives et le traitement du signal par les antennes permettent de réduire l'influence des interférences par la discrimination de leur direction d'arrivée. Ainsi, l'étude de la répartition des interférences en fonction de l'azimut devient une connaissance à priori utile pour la conception des systèmes.

La variation spatiale des interférences concerne à la fois leur prévision et l'élaboration de systèmes à grande base utilisant la diversité d'espace. En effet, un système de prévision des interférences doit prendre en compte à la fois les variations temporelles et les variations spatiales, de façon à établir le temps de préavis optimum et le maillage spatial le moins coûteux limité par un coefficient de corrélation minimum entre les points d'une maille. Ces considérations sont examinées dans une seconde partie de cette présentation.

Les résultats présentés s'appuient sur un ensemble de mesures d'environ 60 millions de points. Ces résultats ont été obtenus avec des systèmes informatisés, seuls capables de permettre un traitement complet. Deux systèmes ont été conçus. Le premier, relativement lourd, a permis d'effectuer des analyses très fines, le second plus léger est constitué par des systèmes identiques transportables utilisés pour l'étude des corrélations spatiales.
Ces systèmes sont décrits dans la première partie de cette communication.

Les résultats d'une telle étude sont exploités dans un projet de réception à grande fiabilité utilisant une base de réception étendue à diversité.

II. - SYSTEME DE MESURE -

L'étude du spectre H.F. montre que la différence entre le niveau de l'émetteur le plus puissant et le niveau du bruit de fond, atteint couramment 100dB, voire davantage. Une telle variabilité impose des contraintes sévères à l'appareillage de mesure. Celui-ci devra posséder à la fois : une grande sensibilité, de manière à pouvoir mesurer effectivement le bruit plancher, ainsi qu'un point d'interception d'ordre 3 élevé, de façon à permettre la mesure du niveau du signal le plus intense, sans engendrer d'intermodulation. Sa dynamique doit donc être aussi grande que possible.

Le spectre H.F. évolue aussi en fonction de l'heure et de la saison, aussi l'appareillage devra-t-il permettre de relever et d'enregistrer, de façon automatique, un grand nombre d'échantillons couvrant l'ensemble des situations possibles. Il devra donc être programmable via un ordinateur.

Dans le but de faciliter l'étude expérimentale du spectre H.F., le LETTI a développé deux systèmes de mesure répondant chacun à des critères spécifiques.

Le premier d'entre eux, dont la figure 1 représente le synoptique, a été conçu dans le but d'effectuer des mesures de grande qualité, à poste fixe. Il se compose :

- D'un analyseur de spectre synthétisé et programmable Marconi 2382.

- D'un ordinateur HP 9000 type A330 équipé d'un disque dur de 40Moctets.

- D'un "tiroir mesure" programmable et étalonnable, destiné à améliorer les performances globales du système.

Le tiroir mesure contient essentiellement :

- Des commutateurs permettant de sélectionner le signal à analyser : charge 50 Ω ou générateur en mode étalonnage, antenne active ou réseau à balayage électronique en mode mesure.

- Une batterie de filtres passifs passe-bande de largeur un demi-octave, destinés à éliminer les produits d'intermodulation d'ordre 2.

- Un ensemble d'amplificateurs large-bande, fort-niveau, faible-bruit, et d'atténuateurs, dont le rôle est de fixer le facteur de bruit, donc la sensibilité de la chaîne.

Ce système possède une sensibilité maximale d'environ -140dBm pour une largeur de bande de 1KHz (facteur de bruit F ≈ 4dB), une dynamique maximale de 85dB et une dynamique totale de 140dB, ce qui assure, dans tous les cas, d'excellentes mesures.

Le second système a été conçu dans le but d'étudier la corrélation spatiale des interférences. La nécessité de transporter fréquemment le matériel a orienté le choix vers un système compact et léger. La figure 2 montre que la chaîne de mesure se compose :

- D'une antenne active AFT2.

- D'un récepteur de trafic synthétisé et programmable NARDEUX T-264, capable de mesurer le niveau du signal, et de la transmettre à un calculateur via un bus IEEE-488.

- D'un calculateur associé à un lecteur de disque souple.

Ce système possède une sensibilité un peu moins bonne que celle du système fixe, et sa dynamique est beaucoup plus faible. Cependant, l'utilisation du circuit de contrôle automatique du gain (C.A.G.) permet de bénéficier d'une dynamique totale supérieure à 100dB, ce qui est amplement suffisant pour l'étude comparative des interférences seules.

Dans les deux systèmes, un programme de mesure permet à l'ordinateur de :

- Fixer les paramètres.

- Lancer une mesure à l'instant prévu.

- Acquérir les résutlats et les convertir en ASCII (ce qui réduit considérablement le volume de données à stocker, et augmente l'autonomie).

- Stocker les résultats sur disquette, sur disque dur ou sur bande selon le cas.

Il faut noter que, du fait de l'évolution de l'ionosphère, et plus particulièrement de l'absorption, le niveau des signaux reçus évolue au cours de la journée. En conséquence, le programme adapte automatiquement les caractéristiques de la chaîne de mesure à celles des signaux analysés, assurant ainsi en permanence des performances optimales.

Enfin, selon la fréquence des mesures et la nature du support utilisé pour le stockage des données, l'autonomie se situe entre quelques jours et quelques semaines.

III. - ETUDE SPATIALE DES INTERFERENCES -

Les études présentées s'inscrivent dans une étude plus vaste qui a porté sur le bruit et sur les interférences [1].

Les mesures à azimut variable, effectuées pour étudier la distribution spatiale du bruit et des interférences, ont été réalisées à la station STUDIO (Système de Traitement Universel et de Diagnostic IOnosphérique), située à 40 km au sud de Paris. Cette station dispose d'un réseau plan d'antennes biconiques à balayage électronique, couvrant environ 1 hectare, dont le lobe de rayonnement peut être orienté dans une direction quelconque (figure 3).

Les spectres observés dans la bande 3-30MHz sont bien connus et les variations temporelles et spatiales sont illustrées par les enregistrements de la figure 4.

III.1. **Bruit et interférences**

La distinction entre le bruit et les interférences a été faite à partir de la fonction de densité de probabilité d'amplitude du signal reçu dans une bande de fréquence définie.

La méthode de mesure utilisée est illustrée par les courbes de la figure 5, résumant les traitements effectués sur des analyses spectrales de 50KHz obtenues avec des résolutions de 300Hz. La figure 5.a, qui représente la densité de probabilité de bruit $p_a(x)$ du système de mesure fait ressortir une distribution de Rayleigh parfaitement conforme à la théorie :

$$p_a(x) = \frac{x}{q_a^2} e^{-x^2/2q_a^2} \qquad (1)$$

où q_a est lié à la valeur quadratique moyenne σ_a^2 du bruit par la relation :

$$\sigma_a^2 = q_a^2 \qquad (2)$$

Lorsque x tend vers zéro, il vient

$$\log p(x) = 0,1\, x_{dB} - 0,1\, \sigma_{a_{dB}} - 0,94 \qquad (3)$$

Cette droite asymptotique Δ a permet d'évaluer le bruit d'appareillage.
Les mesures donnent les densités de probabilités $p_a(x)$ dont les figures 5 b et 5.c donnent deux exemples mesurés en France, au sud de Paris, avec le système d'aériens de la station STUDIO, dans les directions Est et Ouest. On peut noter la différence importante qui existe entre ces enregistrements, qui révèlent des interférences beaucoup plus importantes dans la direction Est que dans la direction Ouest.

La distinction entre le bruit d'appareillage, le bruit atmosphérique et les interférences a été effectuée en remarquant que les interférences ont des puissances supérieures au bruit et apparaissent donc pour les fortes valeurs de l'amplitude. Pour les faibles valeurs, seul le bruit et les interférences faibles subsistent et on retrouve une loi de distribution de Rayleigh asymptotique à la droite

$$\log p(x) = 0,1\, x_{dB} - 0,1\, \sigma_{dB} - 0,94 \qquad (4)$$

La droite Δ est translatée par rapport à la droite Δa d'une quantité

$$\Delta_0 = (\sigma - \sigma_a)\ dB \qquad (5)$$

Il devient donc possible de corriger l'effet d'appareillage et les enregistrements des figures 5.b et 5.c montrent :

- Que le bruit composite (bruit + interférences faibles), est plus important dans la direction Est que dans la direction Ouest.

- Qu'au-delà du maximum de la courbe de densité de probabilité, la caractéristique de Rayleigh disparaît à cause des interférences de fort niveau.

- Que la présence d'interférences de fort niveau est plus importante dans la direction Est que dans la direction Ouest.

III 2. **Répartition azimutale des interférences**

Une étude de la répartition azimutale des interférences a été effectuée à l'aide du réseau de la station STUDIO orientable tout azimut par commande électronique.

Le niveau élevé des interférences montre qu'il est possible de les évaluer par une simple mesure de la puissance reçue dans une bande d'analyse, la puissance du bruit étant négligeable.

La procédure utilisée a consisté à effectuer une analyse spectrale puis à calculer la puissance captée dans cette bande. L'affectation des bandes de fréquence aux différents services, radiodiffusion, services fixes, services mobiles..., pose le problème du choix de la bande d'analyse. On a finalement choisi pour cette étude des bandes d'analyse de 500KHz, 1MHz et 5MHz sans chercher à distinguer les différents services, d'une part pour effectuer une statistique sur un plus grand nombre de points, d'autre part à cause de la répartition relativement uniforme des stations et de leurs puissances d'émission comparables. On a cependant traité de façon séparée les bandes englobant les stations de radiodiffusion dont les puissances émises sont notablement plus élevées que les autres.

Les résultats obtenus sont illustrés par les diagrammes des niveaux reçus de la figure 6 qui représente les résultats d'une évolution journalière typique et enregistrés les 18 et 19 février 1986.

Les bandes d'analyse choisies sont de 1MHz et les niveaux indiqués représentent la puissance reçue normalisée à une largeur de bande de 3KHz, résolution de l'analyse spectrale.

Ces résultats montrent :

- Une évolution journalière du niveau d'interférence qui fait apparaître que la puissance reçue augmente la nuit pour les fréquences faibles alors qu'elle diminue pour les fréquences élevées. Ces résultats sont évidemment liés directement aux conditions ionosphériques qui, de jour entraînent une absorption importante des fréquences faibles dans la région D, et de nuit abaissent les fréquences maximales utilisables à cause de la réduction des fréquences critiques de la région F.

- Pour toutes les fréquences, le niveau d'interférence est nettement plus élevé vers l'Est que vers l'Ouest. On note une différence moyenne de 10 à 15dB particulièrement marquée à 18 heures.
Ce résultat traduit en premier lieu une densité d'émetteurs plus grande à l'Est qu'à l'Ouest, constat qui frappe l'évidence au simple vu de la carte d'Europe occidentale. En second lieu, l'accroissement des niveaux d'interférence à 18 heures peut être interprété par la réduction de l'absorption par la région D pour les émissions issues de l'Est.

Ces résultats montrent que les interférences en Europe occidentale sont essentiellement crées par les émetteurs situés à l'Est, résultats attendus, et qu'une augmentation de 10 à 15dB doit être prise en compte dans l'élaboration des systèmes. L'utilisation d'antennes directives, voire d'antennes adaptatives doit être considérée comme un moyen de lutte efficace contre les interférences.

III.3. Corrélation spatiale des interférences

La corrélation spatiale des interférences est importante pour les raisons signalées précédemment. Des mesures effectuées simultanément sur différents sites ont permis de rechercher les coefficients de corrélation des interférences en fonction des distances séparant les sites.

Les mesures ont été effectuées en analysant, dans la gamme 3-30MHz, un ensemble de bandes de 5 à 30KHz d'étendue selon les cas. On dénombre dans cette gamme une répartition de 95 bandes affectées aux différents services. Le nombre de bandes traitées a été choisi soit égal à 95 avec des étendues d'analyse de 30KHz - cas le plus complet -, soit 50 bandes d'étendues égales à 5KHz.

La figure 7 représente quatre séries de spectres relevés simultanément sur deux sites, l'un à la station STUDIO 1 de Paris, l'autre à la station STUDIO 2 située au sud de STUDIO 1 à une distance de 450 km. Ces spectres sont relevés dans les 95 bandes définies précédemment et chaque enregistrement, effectué sur une durée de 1 heure, représente 15 spectres relevés toutes les 4 minutes.

Ces enregistrements montrent :

- Pour chaque site, la bonne corrélation temporelle que traduit la faible dispersion des courbes.

- L'absorption des fréquences basses en milieu de journée.

- La bonne corrélation des spectres entre les sites STUDIO 1 et STUDIO 2.

Cette corrélation apparaît également dans les variations temporelle des niveaux des signaux relevés dans les deux stations. Les résultats de la figure 8 montrent des enregistrements effectués durant 60 heures sur les sites STUDIO 1 et STUDIO 2 avec des bandes d'analyse de 30KHz dans les bandes suivantes :

- Radiodiffusion :	7 200 KHz
- Service maritime mobile :	8 657,5 KHz
- Service fixe et mobile :	6 882,5KHz
- Bande amateur fixe :	7 050 KHz

Ces résultats montrent qu'il existe une corrélation entre les niveaux reçus quel que soit le type de service observé. On peut noter cependant que la corrélation est fortement marquée pour les bandes de radiodiffusion où l'on note aisément la mise en service d'émetteurs à heure fixe. Cette corrélation plus forte peut être interprétée par le fait que dans ces bandes la densité des émetteurs est plus faible et les puissances plus fortes .

Pour évaluer quantitativement la corrélation entre les niveaux reçus on a utilisé les résultats d'une expérience menée dans des bandes de 5KHz peu contaminées par les interférences.

Une étude systématique de la corrélation entre différents sites a été effectuée à l'aide de 3 stations situées au sommet d'un triangle dont les côtés mesurent 500, 700 et 1000 km [2]. Un choix de 50 bandes de fréquences, de 5KHz de largeur, a d'abord été effectué parmi les bandes repérées comme statistiquement peu contaminées par les interférences. Pour discriminer l'effet azimutal décrit précédemment de l'effet présent, on a utilisé des antennes directives orientées vers le nord. Les mesures ont été effectuées simultanément dans les trois stations avec une précision supérieure à la seconde.

Les mesures ont été exploitées en analysant le niveau S_k du signal reçu à la station k {1,2,3} et en définissant par tranche d'analyse de 2 heures (figure 9) :

- Le nombre N_{ck} de périodes, appelées périodes claires, pendant lesquelles le seuil S_k demeure constamment inférieur à un seuil Sc choisi parmi 15 seuils distants de 3dB. On appelle T_{k1} la durée de la k ième période claire.

- La durée totale du temps T_{ck} des périodes claires

$$T_{ck} = \sum_{i} T_{ki} \qquad (6)$$

- L'écart type σ_{ek} de la durée des périodes claires.

- La durée moyenne D_{ek} des périodes claires

$$D_{ek} = \frac{T_{ci}}{N_{ei}} \qquad (7)$$

De la même façon, on a déterminé les périodes claires T_{ck1k2} durant lesquelles les niveaux S_{k1} et S_{k2} des stations k1 et k2 étaient simultanément inférieurs à S_c, ce qui a conduit à définir les durées moyennes D_{ck1k2} et l'écart type σ_{ck1k2} des périodes claires des stations prises 2 à 2 et enfin la durée moyenne D_{c123} et l'écart type σ_{c123} des périodes claires simultanées des trois stations.

A partir de ces définitions, il a été déterminé les probabilités $P_k(S_c)$ d'obtenir un seuil inférieur à S_c dans chaque station et la probabilité $P_{k1k2}(S_c)$ d'obtenir simultanément des niveaux inférieurs à S_c dans les deux stations.

On a alors défini un coefficient d'indépendance $l_{k1k2}(S_c)$ entre les stations qui prend la valeur 0 lorsque les variables aléatoires deviennent indépendantes. Les résultats obtenus, représentés sur la figure 10 montrent que le coefficient d'indépendance dépend du seuil S_c et qu'il est supérieur à 0,9 pour un niveau supérieur à - 100dB.

Il apparaît donc, au vu de ces résultats, que sur des distances de 500, 700 et 1000 km la corrélation entre les interférences est faible et pratiquement la même quelle que soit la distance entre les stations.

Ce résultat, obtenu sur des canaux sélectionnés pour le faible niveau statistique des interférences, trouve une application dans la transmission d'information à haute fiabilité [2] utilisant une base de réception diversifiée de grande dimension.

IV. - EVOLUTION TEMPORELLE DES INTERFERENCES -

Les données nécessaires à cette étude ont été obtenues à la station STUDIO, grâce au système de mesure fixe précédemment décrit.

IV.1. Acquisition des données

Plutôt que d'analyser la totalité du spectre H.F., nous avons jugé préférable d'étudier quatre bandes de 500KHz, représentatives de la diversité des services de télécommunication rencontrés. Le tableau 11 indique pour chacune des bandes retenues le découpage en sous-bandes et les utilisations assignées par l'U.I.T. (Union Internationale de Télécommunications).

Les mesures sont faites avec une résolution de 1KHz. Chaque spectre de largeur 500KHz enregistré résulte du moyennage de 8 spectres individuels consécutifs. Les 4 bandes sont analysées en moins de 2mn, ce qui permet d'enregistrer un bloc de 30 échantillons spectraux en 1 heure. Ces blocs sont acquis au rythme d'un toutes les 3 heures (figure 12). Si on fixe la durée d'une campagne de mesure à 15 jours, celà permet de constituer une banque de données comportant près de 58 000 000 points de mesure, ce qui constitue une bonne base de travail pour l'analyse numérique de l'évolution temporelle des interférences.

Il est important de noter d'autre part, que le système de mesure est étalonné automatiquement avant chaque cycle d'une heure, afin d'assurer une qualité de mesure maximale.

La figure 13 représente quelques exemples de spectres relevés à 00h et à 12h T.U. le 24 juin 1989, qui mettent en évidence les variations du plancher de bruit et de la dynamique des signaux.

IV.2. Diagramme d'occupation

L'intérêt des diagrammes d'occupation est essentiellement de permettre la recherche des fréquences claires. Utilisés conjointement aux prévisions de propagations, de tels diagrammes constituent une aide précieuse pour l'établissement des plans de fréquence. Il faut signaler que GOTT et Cie [3] [4] [5] ont établi un modèle de l'encombrement spectral dans la bande H.F., tenant compte de la fréquence de travail, du seuil d'interférence toléré et de l'indice d'activité solaire. Ce modèle cependant, ne permet que des prévisions à long terme, car il n'inclus pas les variations saisonnières des interférences. Le système de mesure que nous avons développé est, quant à lui, plutôt destiné à l'étude fine de l'encombrement dans un nombre limité de canaux.

Pour caractériser les évolutions temporelles de l'occupation, nous avons évalué, pour une fréquence f, un seuil S et une heure h donnés, le nombre n de fois où le niveau du signal dépasse le seuil fixé, cela sur une période de 15 jours. Le taux d'occupation croît bien évidemment lorsque le seuil diminue, comme le confirme les illustrations de la figure 14. L'exemple choisi montre que l'émetteur étudié a cessé d'émettre entre 12h24 et 12h54, et que lorsqu'il fonctionnait le niveau du signal capté était supérieur à -95dBm environ. La remontée de l'occupation au seuil -110dBm correspond au bruit d'appareillage et est donc non significative.

IV.3. **Probabilité de clarté**

Les diagrammes d'occupation fournissent une probabilité moyenne d'encombrement sur une période donnée, mais ne renseignement pas sur l'évolution temporelle de l'encombrement. Afin de combler cette lacune, nous avons étudié la probabilité pour qu'une bande soit claire pendant J jours consécutifs, au niveau S, et à une heure donnée.

Les échantillons d'amplitude relatifs à l'heure "h" sont rangés dans un tableau à 3 dimensions $A(N_j, N_t, N_f)$ d'éléments a(j, t, f), avec :

- N_j nombre de jours d'observation ($N_j = 15$)

- N_t nombre d'échantillons dans une heure ($N_t = 30$)

- N_f nombre de points de fréquence dans la bande analysée.

Le seuil S étant fixé, on forme un second tableau $B(N_j, N_t, N_f)$ dont les éléments b(j, t, f) sont tels que :

$b(j, t, f) = 0$ si $a(j, t, f) \geq S$
$b(j, t, f) = 1$ si $a(j, t, f) < S$

La probabilité pour que la bande soit claire J jours consécutifs au niveau S, est alors donnée par :

$$P(S,J) = \frac{\sum_{j=1}^{N_j-J+1} \sum_{i=1}^{N_t} \sum_{f=1}^{N_f} \prod_{i=j}^{j+J-1} b(i,t,f)}{(N_j - J + 1)\, N_t\, N_f} \qquad (8)$$

L'exemple de la figure 15 est typique de l'évolution de la probabilité de clarté. On notera que la valeur correspondant à J = 0 représente la probabilité de non-clarté, ce qui correspond au complément de la probabilité de clarté sur 1 jour (J = 1).

On observe d'autre part, pour $J \geq 1$, une décroissance régulière et monotone de la probabilité de clarté qui suggère une loi exponentielle. Une analyse simple montre qu'il en est bien ainsi. Posons :

- n_j le nombre de canaux clairs J jours consécutifs dans la bande analysée.

- n_{j+1} le nombre de canaux clairs J+1 jours consécutifs.

Il est évident que n_{j+1} est proportionnel à n_j, car les canaux encore clairs le J + 1 ème jour ne peuvent être issus que de l'ensemble des canaux clairs J jours consécutifs. D'autre part, le nombre de canaux clairs diminue nécessairement avec J. On peut donc écrire :

$$dn_j = - A_o \, . \, n_j \, . \, d_j \qquad (9)$$

où la constante positive A_o traduit le taux d'extinction des canaux clairs. La solution de cette équation différentielle est :

$$n_j = \exp(-A_o J - B_o) \qquad (10)$$

et comme la probabilité de clarté est $P(S, J) = n_j/N_f$, on obtient finalement :

$$P(S, J) = \exp(- A J - B) \qquad (11)$$

Cette loi a été vérifiée avec une excellente précision pour l'ensemble des données disponibles. De plus, l'analyse numérique des coefficients A et B en fonction du seuil S, a montré que ceux-ci suivent aussi une loi exponentielle (voir figure 16), ce qui permet d'exprimer finalement la probabilité de clarté sous la forme :

$$P(S, J) = \exp\left[- \exp(- A_1 S - A_2) \cdot J \ - \exp(- B_1 S - B_2)\right] \qquad (12)$$

dans laquelle les coefficients A_1, A_2, B_1, B_2 dépendent de l'heure et de la bande de fréquence considérés.

Nous avons aussi étudié l'évolution de la probabilité de clarté, pour un seuil donné, en fonction de l'heure. La figure 17 montre que la probabilité de clarté est, pour un seuil donné, maximale à 12h00, et minimale vers 18h00. Les résultats obtenus traduisent à la fois l'évolution des conditions de propagation (MUF, absorption), et les conditions d'utilisation des émetteurs H.F. (arrêt de certaines émissions la nuit, émissions plus nombreuses en fin de journée). Les mêmes graphes montrent d'autre part qu'une période de mesure de 7 à 10 jours suffit pour permettre l'évaluation des paramètres A et B entrant dans la formule donnant la probabilité de clarté.

IV.4. <u>Prévision des fréquences claires</u>

La modélisation des interférences permet d'envisager une prévision à court terme des fréquences claires. Une méthode simple consiste à mesurer l'encombrement spectral pendant J jours, puis à utiliser l'information acquise pour estimer l'encombrement au jour J + 1.

Dans ce but, nous avons défini la probabilité C(S, J) pour qu'une bande claire J jours consécutifs soit claire le J + 1 ème jour. Pour calculer l'indice de clarté C(S, J), définissons :

- EJ : l'évènement : bande claire J jours consécutifs au seuil S
- EJ+1 : l'évènement : bande claire J+1 jours consécutifs au seuil S
- DJ+1 : l'évènement : bande claire le jour J+1 au seuil S lorsque les J jours précédents sont clairs.

On a alors :

$$C(S,J) = \text{Prob}(D_{J+1} \text{ si } E_J) = \frac{\text{Prob}(D_{J+1} \text{et } E_J)}{\text{Prob}(E_J)} \qquad (13)$$

or :

$$\text{Prob}(D_{J+1} \text{ et } E_J) = \text{Prob}(E_{J+1}) \qquad (14)$$

d'où

$$C(S,J) = \frac{\text{Prob}(E_{J+1})}{\text{Prob}(E_J)} = \frac{P(S,J+1)}{P(S,J)} \qquad (15)$$

La figure 18 montre comment évolue l'indice de clarté en fonction du seuil pour plusieurs valeurs de J. Pour des niveaux de seuil faibles, les courbes sont limitées par le bruit plancher. Le décalage de l'origine des courbes provient du fait que la probabilité de clarté diminue lorsque le nombre de jours diminue.

Si les courbes sont très différentes pour les seuils faibles, elles ont tendance à se rapprocher pour les seuils élevés. Les différences observées s'expliquent par le fait qu'aux faibles niveaux c'est le bruit qui joue le rôle principal, alors qu'aux forts niveaux ce sont les interférences. Or le modèle adopté ne s'applique qu'aux interférences. Si on se limite à la gamme de validité du modèle, et si on introduit dans C(S,J) l'expression établie précédemment pour P(S,J), il vient :

$$C(S,J) = \exp(-A) = \exp[- \exp(-A_1 S - A_2)] \qquad (16)$$

Il apparaît que l'indice de clarté est indépendant du nombre de jours d'observation, et ne dépend que du seuil S, pour une bande de fréquence et une heure données. La figure 19, qui représente les graphes calculés et théoriques (expression 16) de C(S,J) pour S compris entre -110 et -85dBm, confirme la validité du modèle précédent.

V. - CONCLUSION -

La bande des ondes décamétriques demeure encore l'une des plus utilisées ; cependant, la qualité des liaisons dépend du rapport signal à bruit plus interférences et non du seul rapport signal à bruit.

Si des prévisions ionosphériques permettent de prévoir avec une bonne probabilité l'intensité du signal reçu, peu d'études ont été menées sur la possibilité d'effectuer des prévisions sur les interférences.

Comme pour les prévisions ionosphériques, il est possible d'envisager des prévisions à long, moyen, ou court-terme. L'étude présentée montre qu'il existe à la fois une corrélation spatiale et une corrélation temporelle des interférences. Pour chacune d'elles, des lois de prévision peuvent être établies au vu des mesures effectuées. Les relations empiriques obtenues sont susceptibles d'être interprétées à partir de modèles théoriques simples.

Ces résultats doivent être considérés comme un encouragement à poursuivre les études de modélisation et de prévision des interférences, à l'échelle locale comme à l'échelle planétaire. Pour que ces études aboutissent, il sera nécessaire de définir en premier lieu un maillage pour l'échantillonnage spatial, et une périodicité pour l'échantillonnage temporel, déterminés à partir de la précision des prévisions souhaitées.

BIBLIOGRAPHIE

[1] C. GOUTELARD, J. CARATORI - *Prévisions à très court terme par modélisation du brouillage et de l'ionosphère à l'aide de la rétrodiffusion* - AGARD, C.P. N° 453 - San-Diego - 1989

[2] C. GOUTELARD, J.P. VAN UFFELEN - *Influence des variations ionosphériques sur les systèmes H.F. à haute fiabilité utilisant de grandes bases* - AGARD, C.P. N° 441 - Münich - 1988

[3] G.F. GOTT, N.F. WONG, S. DUTTA - *Occupancy measurements across the entire H.F. Spectrum* - AGARD, C.P. N° 332 - Issy-les-Moulineaux - 1982

[4] P.J. LAYCOCK, M. MORELL, G.F. GOTT, A.R. RAY - *A model for H.F. Spectral occupancy* - IEE - London - 1988

[5] G.F. GOTT, E.D.R. SHEARMAN - *H.F. Systems operating beneath the ionosphere* - AGARD, L.S.P N° 172 - Issy-les-Moulineaux - 1990

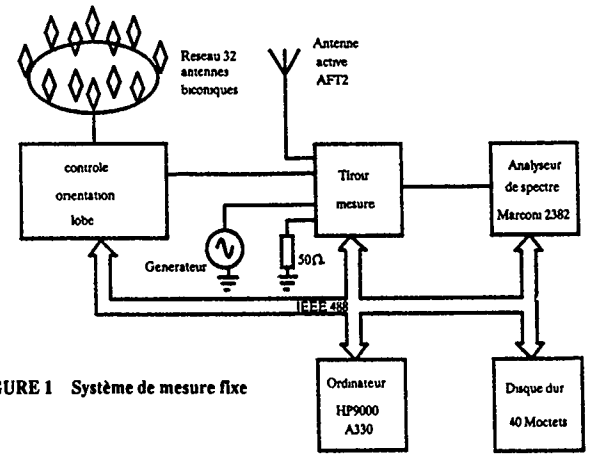


FIGURE 1 Système de mesure fixe

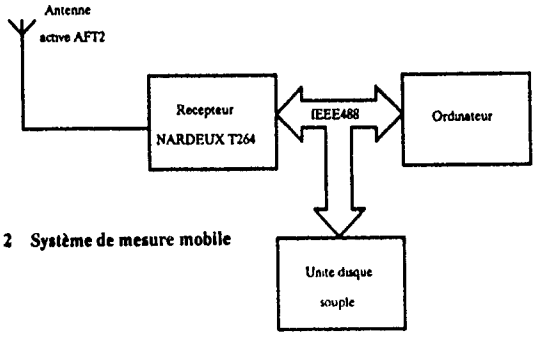


FIGURE 2 Système de mesure mobile

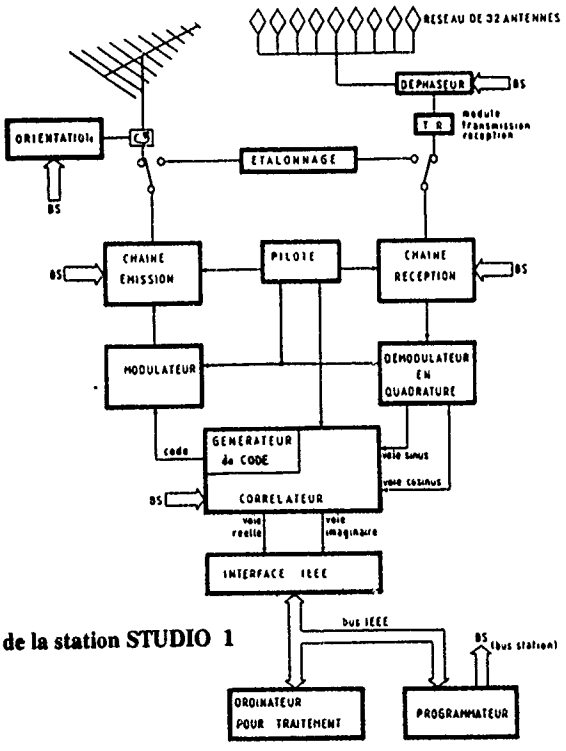


FIGURE 3 Synoptique de la station STUDIO 1

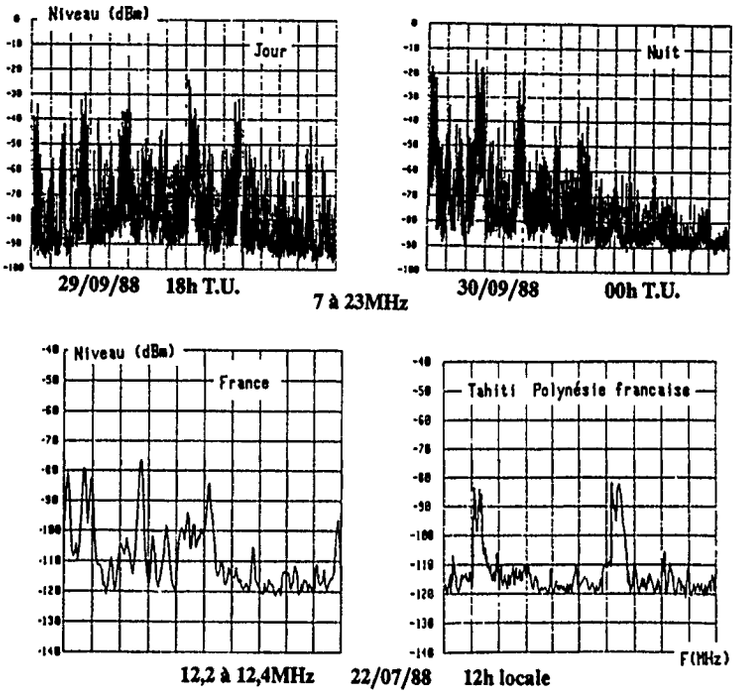


FIGURE 4 (a) Spectre H.F. en Europe Occidentale
(b) Variations géographiques du spectre H.F.

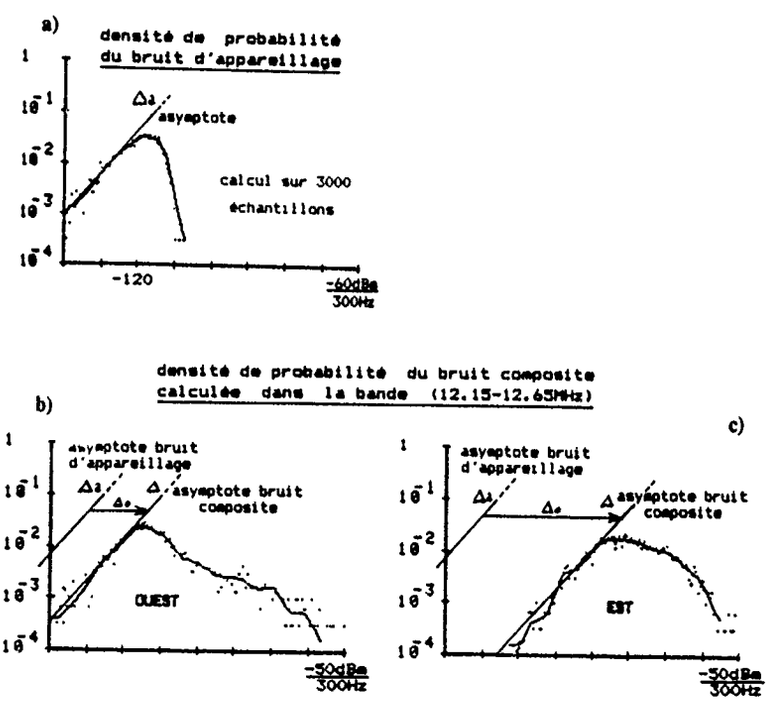


FIGURE 5 Densités de probabilité de bruit mesurées

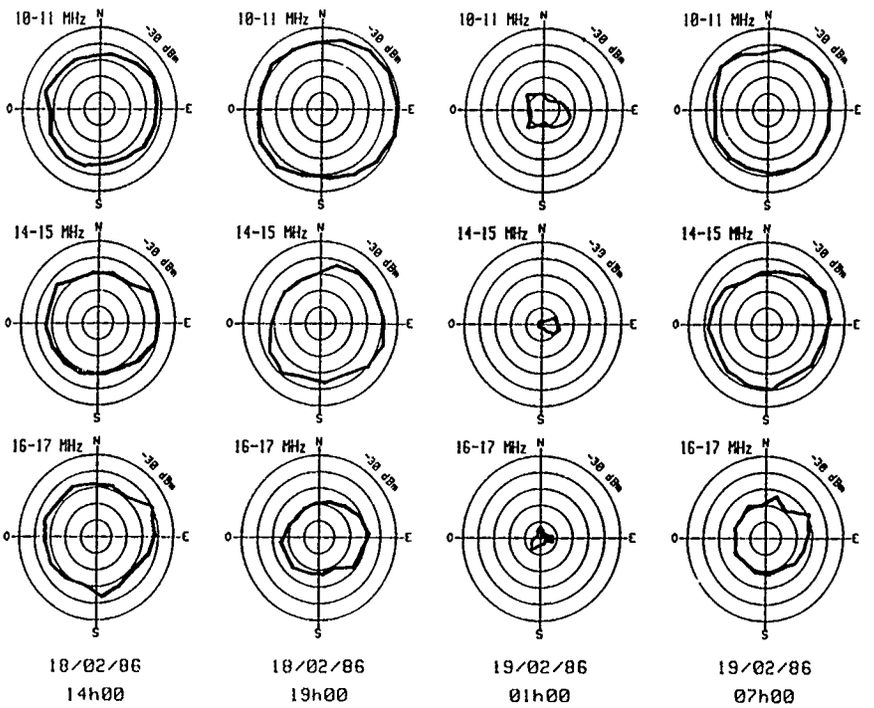


FIGURE 6 Répartition azimutale de la puissance reçue dans une bande de 1MHz, rapportée à 3KHz

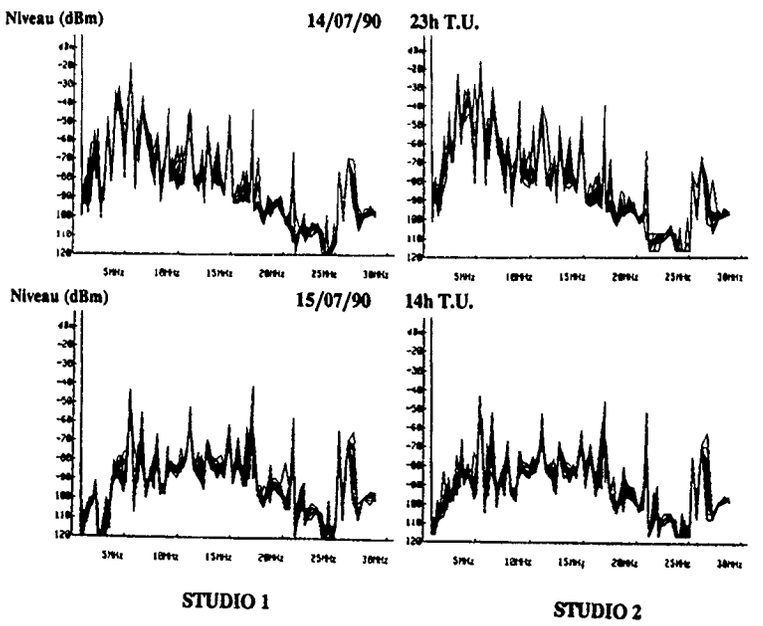


FIGURE 7 Corrélation spatiale du spectre H.F.

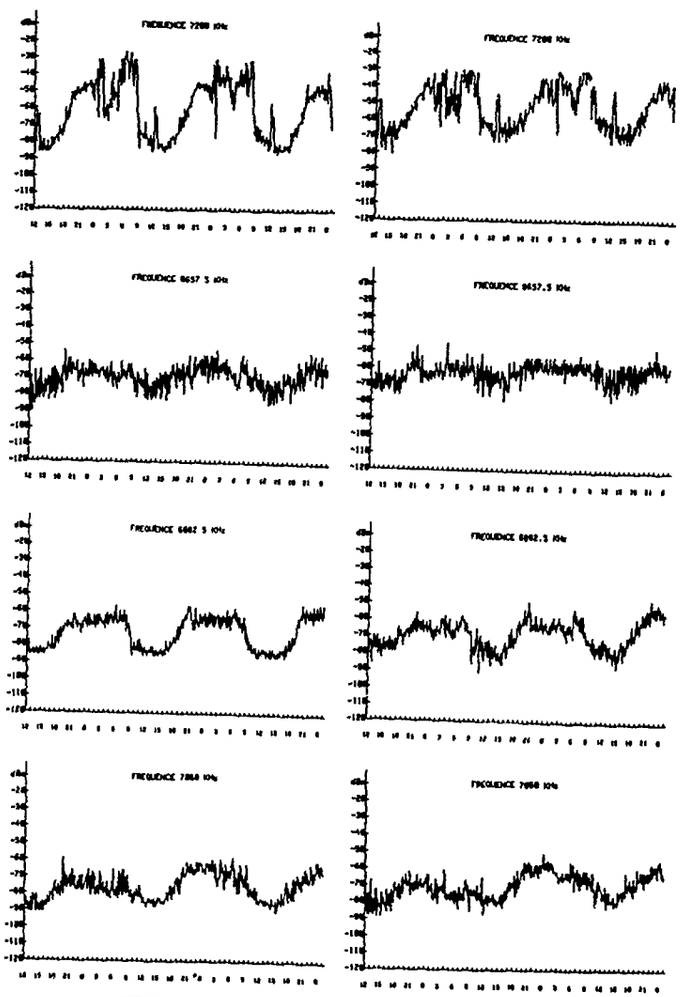


FIGURE 8 Corrélation spatiale et temporelle du spectre H.F.
Jour de début : 31.08.90 à 12h00 T.U.

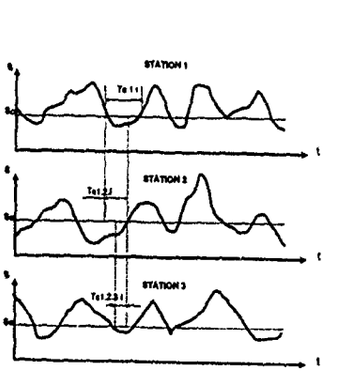


FIGURE 9 Analyse du niveau d'interférence

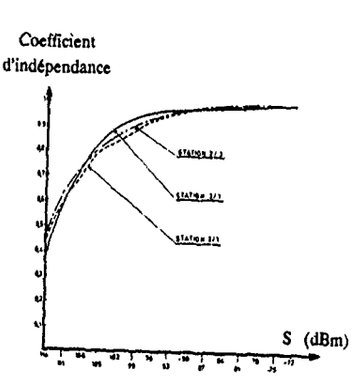


FIGURE 10 Indépendance entre stations

Bande 9.8 —— 10.3 MHz

9.800 —— 10.000 MHz : service fixe
10.000 —— 10.100 MHz : service aéromobile
10.100 —— 10.300 MHz : service fixe

Bande 11.9 —— 12.4 MHz

11.900 —— 11.975 MHz : service radiodiffusion
11.975 —— 12.330 MHz : service fixe
12.330 —— 12.400 MHz : service mobile maritime

Bande 15.4 —— 15.9 MHz

15.400 —— 15.450 MHz : service radiodiffusion
15.450 —— 15.900 MHz : service fixe

Bande 21.1 —— 21.6 MHz

21.100 —— 21.450 MHz : service amateur
21.450 —— 21.600 MHz : service radiodiffusion

TABLEAU 11 Composition des bandes analysées

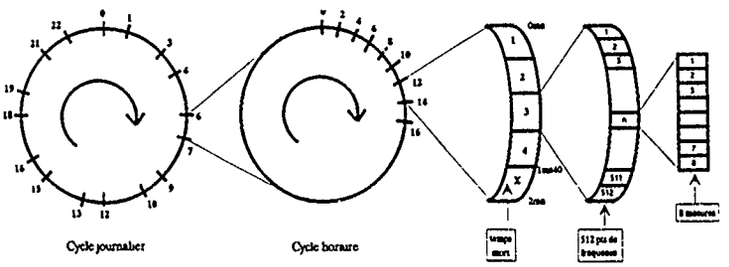


FIGURE 12 Structure d'un cycle de mesure

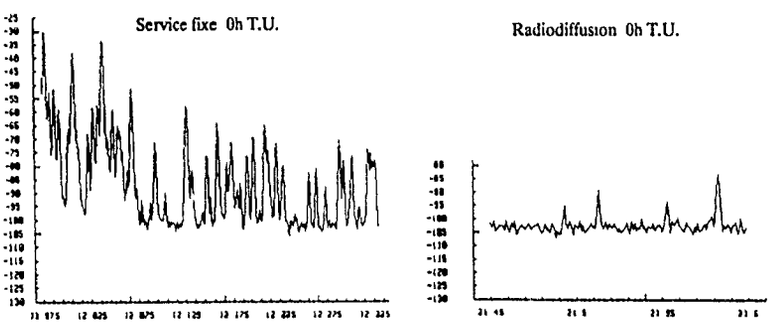


FIGURE 13 Evolution temporelle du spectre H.F.

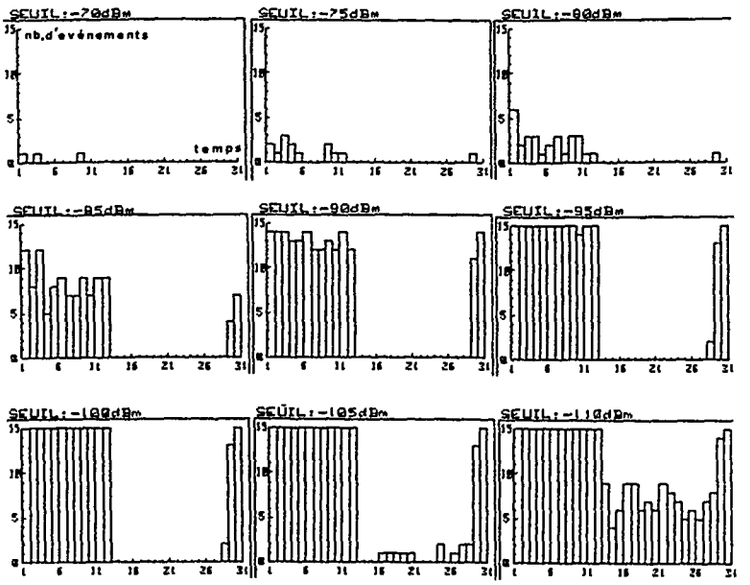
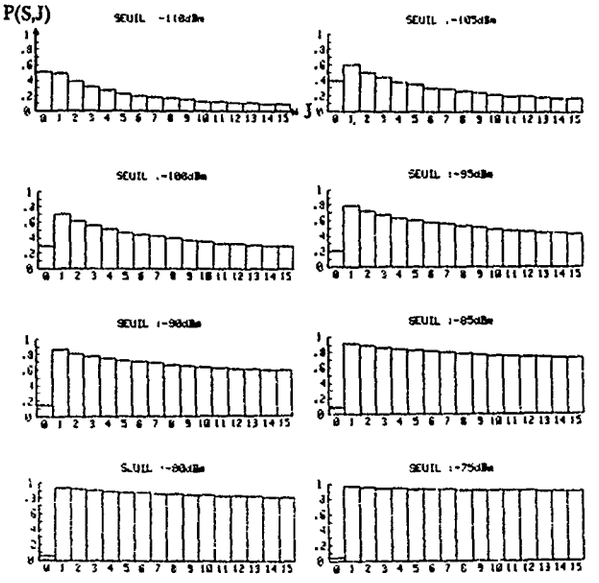


FIGURE 14 Diagrammes d'occupation.
Bande de radiodiffusion F = 11,936MHz

12h T.U. à 13h T.U.
Période du 24/06/89 au 09/07/89
Echelle des temps : 2mn/division



Bande fixe 11,975MHz à 12,33MHz
12h T.U. Période du 24/06/89 au 09/07/89

FIGURE 15 Probabilité de clarté en fonction du nombre de jours d'observation J

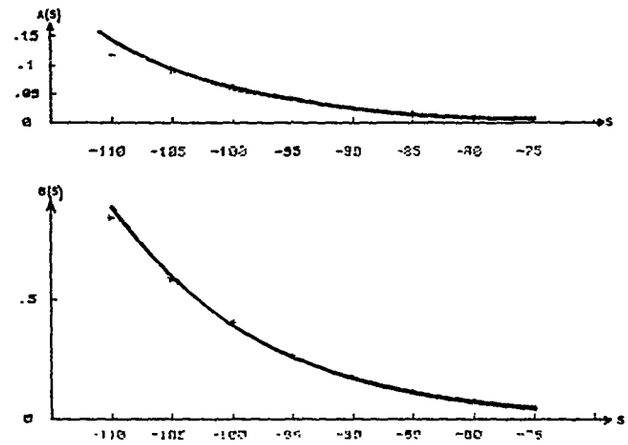


FIGURE 16 Variations des paramètres A et B en fonction du seuil S

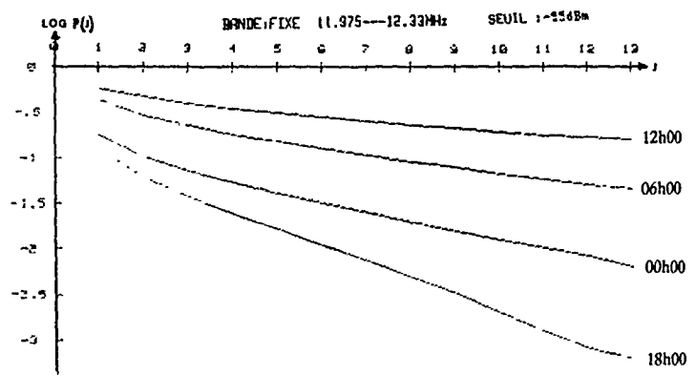


FIGURE 17 Variations du logarithme de la probabilité de clarté en fonction du nombre de jours J et de l'heure

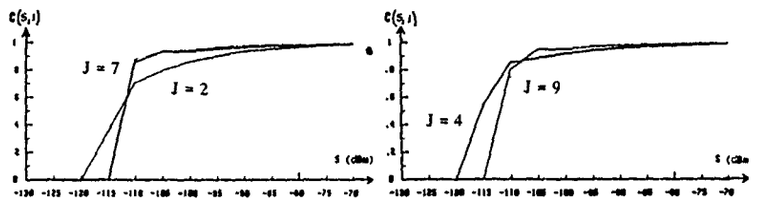


FIGURE 18 Variations de l'indice de clarté en fonction du seuil S pour différentes valeurs de J

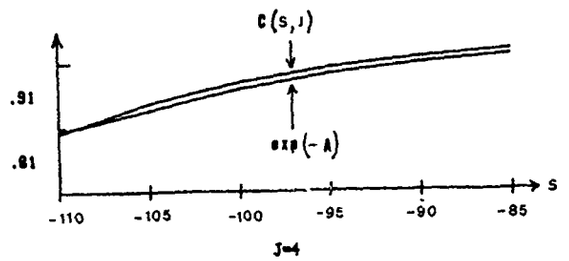


FIGURE 19 Graphes calculés et théoriques de l'indice de clarté en fonction du seuil S pour J = 4

DISCUSSION

George H. HAGN

Data in your paper indicates there is a chance for a regional model for congestion and for "anti-congestion". However, I am surprised that the decay of clear channels is as slow as it is across the 15 days for which you showed data.

Could you comment on whether you also were surprised by this relative stability ?

AUTHOR'S REPLY

Il n'existe pas de canaux clairs durant plusieurs jours consécutifs, ni même durant un jour, dans la bande HF.

Les résultats présentés concernent la clarté obtenue sur une période d'observation définie, dans l'exemple présenté, cette période est de 1 heure. D'autres périodes ont été évidemment analysées.

Topology-Selective Jamming of Fully-Connected, Code-Division Random-Access Networks

Andreas Polydoros
Communication Science Institute
PHE414, EE Dept/Systems
University of Southern California
Los Angles, CA 90089-0272

and

Unjeng Cheng
JPL
4800 Oak Grove Drive
Pasadena, CA 91109

Abstract — The purpose of this paper is to introduce certain models of topology-selective stochastic jamming and examine its impact on a class of fully-connected, spread-spectrum, slotted ALOHA-type random-access networks. The theory covers dedicated as well as half-duplex units. The dominant role of the spatial duty-factor is established, and connections with the dual concept of time-selective jamming are discussed. The optimal choices of coding rate and link-access parameters (from the users' side) and the jamming spatial fraction are numerically established for DS and FH spreading.

1. Introduction

The throughput/delay performance of Code-Division Random-Access (CDRA) spread-spectrum networks has received wide attention in the past decade, particularly because of the military interest in mobile packet radio [KaGrBuKu78], [DaGr80], [Rayc81], [Jubi85], [ShTo85], [Purs87], [PoSi87]. Accordingly, the survivability of a Packet Radio (PR) network under jamming attack is an important issue. Conceptually, a network could be attacked on three layers of importance, namely the end-to-end network, link, and physical layers. In such a jamming game, many factors could affect the results. The communicators have the choices of routing algorithms, channel quality monitoring schemes, and the network information-exchange schemes. On the other hand, jammer choices include *temporal* features (such as static, dynamic or follower jamming) as well as *topological* features (i.e., selection of nodes or links to be jammed).

The purpose of the present paper is to introduce certain concepts on *topology-selective jamming* and further examine its impact on a certain class of "local" or "monohop" or "fully-connected" networks, which we specify below. This particular class of networks has been selected because there exists a convenient analytic vehicle which adequately describes their performance [PoSi87]; thus, it provides a good starting model upon which topological jamming can be defined and assessed with a reasonable degree of analytic ease. We note, however, that the following definitions of topological selectivity are quite general and can be applied to other models as well. We also note that topological selectivity can be perceived as a complementary notion to *temporal selectivity*, which pertains to jamming patterns with different time-domain profiles but homogeneous with regard to space. In other words, the jammer can be ON or OFF in a deterministic or stochastic way but, when ON, all nodes are jammed; the reader is referred to [PrPo87] for a discussion of the latter case.

As mentioned, jamming with any particular selectivity feature manifests itself on all three layers. Thus, it is not immediately clear how the "local" or "monohop" results of this paper ought to be interpreted in a larger multihop environment, especially in view of the fact that there does not seem to exist a unique, widely accepted analytic tool for performance evaluation in this case. Here, we choose to focus on the monohop model because (a) certain networks are indeed quite adequately described by this model and (b) it allows for certain conclusions and assessment to be made. In particular, multi-receiver satellite systems employing CDRA techniques [Wu84] or terrestrial networks which either operate in smaller ranges or employ hierarchical structures with dedicated repeater-nodes would fit well the present model.

The paper is organized as follows: first, in section 2, we review the basic features of the network models. As we explain, we will consider two distinct classes of networks: those with dedicated transmitters and receivers, where no switching between those two functions takes place (sections 3 – 5), and those with non-dedicated or half-duplex units, where nodes indeed switch between the two functions sequentially in time (sections 6 – 8). The reason for the distinction is that the jamming parameters must be somewhat different for the two possibilities. So, a stochastic jamming model for the first case is presented in section 3, followed by the corresponding throughput/delay analysis in section 4. For illustrative purposes, two particular examples (scenarios) of topology-selective jamming are considered in detail. In this section, an efficient combinatorial algorithm is also presented for the recursive evaluation of certain important probabilistic parameters, an enhancing feature of the theory in [PoSi87] in its own merit. Numerical results for this section are presented in section 5. The alternative network and the concomitant jamming model is explained in section 6, while its analysis is presented in subsections 6a and 6b. Numerical results and comparisons are included in section 7. The paper concludes with the discussion in section 8.

2. Network Models

We identify in this section the particular monohop network which we employ in the present study. As mentioned, the monohop network environment under consideration is either identical or closely related to the one analyzed in [PoSi87]. In particular, we assume an arbitrary collection of potential transmitters (TRs) and receivers (RCVRs) in a statistically symmetrical setup, namely, each TR faces the same probabilistic circumstances in the channel; an analogous symmetry holds from each RCVR's viewpoint. Furthermore, every TR can be heard from every RCVR; in this sense, the network is fully-connected. Time is slotted. The total number of nodes or units under consideration is denoted by U (see Table 1, which is borrowed from [PoSi87] and summarizes the notation and terminology throughout this paper also).

We should distinguish, at this point, between the concepts of *potential* and *active* TRs or RCVRs. Any unit that *can* and, at some time, *will* transmit, belongs to the set of potential TRs whose fixed size is N_T; the same goes for potential RCVRs (size N_R). The term "active" is reserved for those units which indeed act in either capacity at a particular slot. For random access, the number of *active* TRs in any slot is a random variable (r.v.), denoted by M_T; thus, in general, $M_T \neq N_T$. On the other hand, the picture regarding the number of active RCVRs is a bit more complicated: if the network is built with exclusively-functioning TRs and RCVRs, then *active* and *potential* RCVRs are identical notions, i.e., $M_R = N_R$; we will call this the *dedicated* model and address it in the following section. The aforementioned satellite random-access application would mostly fit in this category. Let us note that *full-duplex* units constitute a special case of dedicated networks with $U = N_T = N_R$. In general, $N_T \neq N_R$, such as in the single-star [DaGr80] or connected-star topology. In many terrestrial applications it is much more economical to deploy *half-duplex* modems, which means that a unit will successfully alternate between the receiving mode and the transmitting mode; it is not dedicated solely to either one of the two functions, and it cannot do both at the same time. This comprises the *nondedicated*-modem network model, and we consider it in section 6. Of course, hybrid models can be created with some dedicated and some nondedicated units, but their analysis would be a straightforward extension of the present one. Let us note again that, in all cases, the number of active TRs M_T is a random variable ($0 \leq M_T \leq N_T$) with an unconditional pdf $f_{M_T}(m)$ (see Table 1).

The access protocol is of the slotted-ALOHA type with p_0, p_r denoting the new-packet and backlogged-packet transmission

probabilities, respectively [KlLa75]. We shall assume a finite number of TR's. No buffering will be assumed; thus, transmitters are always *busy* (never *idle* with an empty buffer), implying that they are either in the *originative* mode (with a new packet) or in the *backlogged* mode (with a packet that failed sometime previously); this is the standard model of [KlLa75]. For a network under jamming stress, the assumption of never-idle users is quite reasonable and corresponds to the interesting case of the jammer being sufficiently strong to keep the network busy most of the time. Furthermore, regardless of the mode they are in, the TRs will be either *active* or *inactive* (nontransmitting) in a particular slot, as we indicated previously. If the units are half-duplex, they can receive during the inactive periods; otherwise, they do nothing. A small summary of the nomenclature and status-classification of the units appears in Figure 1.

Finally, regarding the spreading code distribution, we assume either a common code or TR-based code system, thus specifically excluding RCVR-based codes [SoSi84], [SuLi85], [Purs87]. The implication is that a TRs packet can be successfully received by more than one RCVR, although it could potentially contribute only one unit to the total success (or throughput) count. We have termed such scenario *competitive* to account for the TRs "effort" to secure *some* RCVR's attention. This is in contrast to the *paired-off* scenario (which can be thought of as a simultaneous TR *and* RCVR-based system) where TRs and RCVRs form distinct pairs and suffer only from secondary multi-user interference (plus jamming), but where RCVR-attention is warranted for each individual packet. We will show comparisons between the competitive and paired-off cases within the jamming framework in section 5.

Let us now proceed with the development of the dedicated-unit jamming model; the half-duplex case will just be an appropriate modification thereof.

3. Jamming of Dedicated-Unit Networks

Let us consider an arbitrary topology involving N_T fixed transmitters and N_R fixed receivers under the aforementioned probabilistic symmetry condition. We adopt the following probabilistic model for the jamming action: the slotted timing is perfectly known to the jammer, so that perfect slot synchronization exists (the cost of removing such an assumption can be dealt with in a way similar to [PrPo87]). The jammer performs an independent probabilistic action on a slot-by-slot basis. In each slot, a random number of receivers, M_R^J, is selected for jamming out of the N_R total. The selection of the specific jammed subset is independent from slot-to-slot, so that no memory can be exploited by the users. Note that the jammer knows exactly the location of the dedicated receivers. Also, the jamming pattern is assumed independent of the channel state (e.g., the number of backlogged users), a realistic scenario for a rapidly changing environment. However, extensions to a channel-dependent scheme would also be of interest, possibly necessitating a different probabilistic jamming model than the one herein.

A particular topology-selective jamming strategy manifests itself in the way the M_R^J jammed receivers are selected. If we define the binary-valued r.v.'s A_i^J; $i=1,\ldots,N_R$ as

$$A_i^J = \begin{cases} 1, & \text{if RCVR}_i \text{ is jammed} \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

then we can quantify topological selectivity by the joint probability mass distribution function

$$\Pr\left[\underline{\Delta}^J = \underline{a}\right] \triangleq \Pr\left[A_1^J = a_1, \ldots, A_{N_R}^J = a_{N_R}\right] \tag{2}$$

where $a_i = 0$ or 1; $i = 1,\ldots,N_R$. Thus, the experiment performed by the jammer consists of drawing a random vector $\underline{\Delta}^J$ in each slot, which then identifies the targeted RCVR's. The assumption of independent trials (per slot) plus any specific distribution as per (2) completely determine the underlying probability space.

Associated with the above are two important parameters, namely (a) what we shall call the *spatial duty factor*, ρ_{sp}, defined as (here, $\mathcal{E}\{\cdot\}$ stands for expectation)

$$\rho_{sp} \triangleq \frac{\mathcal{E}\left\{M_R^J\right\}}{N_R} \tag{3}$$

signifying the average fraction of jammed RCVRs and (b) the jamming power per attacked RCVR per slot, J_{RCVR}, assumed fixed throughout, which relates to the *average* jamming power J_{av} as

$$J_{RCVR}(\rho_{sp}) \triangleq \frac{J_{av}}{\mathcal{E}\left\{M_R^J\right\}} = \left(\frac{J_{av}}{N_R}\right)\rho_{sp}^{-1} \tag{4}$$

In the following analysis we shall assume that the jammer is average-power-limited, i.e., J_{av} is constant and given. Thus, varying the spatial duty factor induces the classical tradeoff between percentage of jammed RCVR's versus the power J_{RCVR} (ρ_{sp}) directed to each one of them. The functional dependence of equation (4) on ρ_{sp} is meant to enhance this point.

Definition (2) is fairly general and can serve as a starting point for different probabilistic jamming options. Here, we focus on two special cases which we shall call scenario 1 and 2, respectively:

Scenario 1: Each RCVR is jammed with probability p_J, independent of any other RCVR. Then,

$$\Pr[\underline{\Delta}^J = \underline{a}] = p_J^{\Sigma a_i}(1-p_J)^{N_R - \Sigma a_i} \tag{5}$$

where $\Sigma a_i \triangleq m_R^J(\underline{a})$ is the standard Hamming weight of the binary vector $\underline{a}$. Note that $m_R^J(\underline{a})$ is simply the value of the r.v. M_R^J in that slot. Consequently, $\mathcal{E}\left\{M_R^J\right\} = p_J N_R$, which simply implies from (3) that

$$\rho_{sp} = p_J \quad \text{(Scenario 1)} \tag{6}$$

Scenario 2: The number of RCVR's jammed is fixed to N_R^J (a given constant), although the specific subset changes randomly every slot. Then,

$$\Pr\left[\underline{\Delta}^J = \underline{a}\right] = \begin{cases} \binom{N_R}{N_R^J}^{-1}, & \text{if } m_R^J(\underline{a}) = N_R^J \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

regardless of the exact location of 1's in $\underline{a}$. Clearly,

$$\rho_{sp} = \frac{N_R^J}{N_R} \quad \text{(Scenario 2)} \tag{8}$$

Note that ρ_{sp} of (8) is restricted to multiples of N_R^{-1}, while ρ_{sp} of Scenario 1 can take on any value in (0,1].

4. Performance Analysis

A network under the jamming model of the previous section can be analyzed by generalizing the concepts in [PoSi87]. In particular, we shall assume that the probabilistic symmetry conditions of the above model are still true and that the probability distribution of the number of successes per slot can be uniquely determined, once the number of attempted transmissions $M_T = m$ and the specific jamming pattern $\underline{\Delta}^J = \underline{a}$ are given for that slot. We can evaluate the throughput β (packets/slot) as the expected number of successful packet transmissions S, i.e.,

$$\begin{aligned} \beta = \mathcal{E}\{S\} &= \mathcal{E}_{M_T,\underline{\Delta}^J}\left\{\mathcal{E}\left\{S|M_T,\underline{\Delta}^J\right\}\right\} \\ &= \sum_{\underline{a}} \Pr[\underline{\Delta}^J = \underline{a}]\, \mathcal{E}_{M_T}\left\{\mathcal{E}\left\{(S|M_T,\underline{a})\right\}\right\} \\ &= \sum_{\underline{a}} \Pr[\underline{\Delta}^J = \underline{a}] \sum_m m p_s^T(m,\underline{a})\, f_{M_T}(m) \end{aligned} \tag{9}$$

where $f_{M_T}(m)$ is the composite slot traffic (Table 1) and $p_m^T(m,\underline{a})$ is the probability of success from a typical TR's viewpoint, given another $(m-1)$ packets and a specific jamming pattern $\underline{a}$ in the slot. In deriving (9), we have used the fact that the r.v. M_T (attempted transmissions) is independent of $\underline{\Delta}^J$, since there is no

coupling between the users' and jammer's actions within a slot. Upon interchanging the summations in (9) and defining the *jamming-average* probability of success from a TR's viewpoint, conditioned on another (m – 1) packets, as

$$\overline{p_s^T(m)} \triangleq \sum_{\underline{a}} \Pr[\underline{\Delta}^J = \underline{a}]\, p_s^T(m,\underline{a}) \tag{10}$$

we arrive at

$$\beta = \sum_m m\, \overline{p_m^T(m)}\, f_{M_T}(m) \tag{11}$$

which is a direct generalization of (42) in [PoSi87]. In order to proceed, we need to evaluate $\overline{p_m^T(m)}$ and $f_{M_T}(m)$ for any specific jamming strategy as per (2).

4a. Evaluation of $\overline{p_m^T(m)}$

In principle, $\overline{p_m^T(m)}$ can be evaluated in accordance with *Proposition 1* of [PoSi87] (which is quite general and also holds in this jammed scenario) as

$$\overline{p_s^T(m)} = \frac{1}{m} \sum_{s=1}^{m} s\, \overline{p_{s|m}} \tag{12a}$$

where

$$\overline{p_{s|m}} \triangleq \sum_{\underline{a}} \Pr[\underline{\Delta}^J = \underline{a}]\, p_{s|m,\underline{a}} \tag{12b}$$

is the average (over the jamming strategy) of the probability

$$p_{s|m,\underline{\Delta}^J} = \Pr[\text{s successes, given m attempted transmissions and jamming pattern } \underline{\Delta}^J] \tag{13}$$

For the nonjammed scenario, [PoSi87] provides examples of how to calculate $p_{s|m}$ in a variety of situations in terms of the number of RCVR's, their statistical dependence, etc. This can be a very complicated procedure for arbitrary models; however, a simplification occurs if one assumes that, conditioned on a specific jamming pattern in a slot, each RCVR accepts packets in a statistically independent fashion from other RCVR's. This assumption, which would obviously be valid if receiver thermal-noise were the only deterrent, can be argued to be numerically satisfactory even in the presence of multi-user noise because of random spreading patterns, fading, random distances, ground propagation and formation, etc. Then $\overline{p_{s|m}}$ can be evaluated in a recursive way which we shall present in the following section, since it is also required in the evaluation of $f_{M_T}(m)$. Let us just note here that, once this evaluation is completed, $\overline{p_m^T(m)}$ follows from (12).

Under the independence assumption outlined above, a shorter path for evaluating $\overline{p_m^T(m)}$ is as follows: Let $p_A^R(m,J)$ and $p_A^R(m;J^c)$ denote the probabilities that a typical RCVR will accept a packet in the presence of another (m – 1) contending ones, given that this RCVR is jammed (event J) or not (event J^c), respectively. Note that $p_A^R(m;J^c)$ is identical to $p_A^R(m)$, as introduced in [PoSi87]. Here, because of the possible jamming, we need to distinguish further and define these conditional quantities. Then, using (10) and an immediate generalization of *Proposition 2* of [PoSi87], we get

$$\begin{aligned} p_m^T(m) &= \sum_{\underline{a}} \Pr[\underline{\Delta}^J = \underline{a}] \\ &\quad \cdot \left[1 - \left(1 - \frac{p_A^R(m;J^c)}{m}\right)^{N_R - m_R^J(\underline{a})} \left(1 - \frac{p_A^R(m;J)}{m}\right)^{m_R^J(\underline{a})}\right] \\ &= \sum_{m_R^J=0}^{N_R} \Pr\left(M_R^J = m_R^J\right) \\ &\quad \cdot \left[1 - \left(1 - \frac{p_A^R(m;J^c)}{m}\right)^{N_R - m_R^J} \left(1 - \frac{p_A^R(m;J)}{m}\right)^{m_R^J}\right] \end{aligned} \tag{14}$$

where M_R^J is the Hamming-weight transformation (a r.v.) of the random vector $\underline{\Delta}^J$. Note the impact of the aforementioned symmetry assumption, which implies here that all jammed receivers suffer from interference of the same *average* power.

To illustrate, consider again the two scenarios of section 3. We immediately have that

$$\Pr\left(M_R^J = m_R^J\right) = \Delta\left(m_R^J, N_R, \rho_{sp}\right) \quad ; \text{ scenario 1} \tag{15a}$$

and

$$\Pr\left(M_R^J = m_R^J\right) = \begin{cases} 1, & \text{if } m_R^J = N_R^J \\ 0, & \text{otherwise} \end{cases} \quad ; \text{scenario 2} \tag{15b}$$

which, when substituted in (14), yield

$$\overline{p_s^T(m)} = \begin{cases} 1 - \left[1 - \dfrac{\rho_{sp}\, p_A^R(m;J) + (1-\rho_{sp})\, p_A^R(m;J^c)}{m}\right]^{N_R} & \text{(scenario 1)} \\ 1 - \left[\left(1 - \dfrac{p_A^R(m;J)}{m}\right)^{\rho_{sp}} \left(1 - \dfrac{p_A^R(m;J^c)}{m}\right)^{1-\rho_{sp}}\right]^{N_R} & \text{(scenario 2)} \end{cases} \tag{16}$$

It should be kept in mind that $p_A^R(m;J)$ above depends on ρ_{sp} via the jamming level $J_{RCVR}(\rho_{sp})$. Note also that the result for scenario 1 depends only on the *jamming-average acceptance probability*

$$\overline{p_A^R(m)} \triangleq \rho_{sp}\, p_A^R(m;J) + (1-\rho_{sp})\, p_A^R(m;J^c) \tag{17}$$

which is an intuitively expected result, because RCVRs are jammed independently for this case. Obviously, this is not true for scenario 2.

For the specific protocol choice $p_0 = p_r = p$ (which can be described as "uncontrolled-ALOHA"), the composite traffic $f_{M_T}(m)$ is simply the binomial distribution $\Delta(m, N_T, p)$, regardless of any jamming action. For this limited case, equations (11) and (16) suffice to determine throughput. The more general and interesting case, however, is when $p_0 \neq p_r$, which we examine below.

4b. Evaluation of $f_{M_T}(m)$ and $\overline{p_{s|m}}$

Following closely the steps in [PoSi87, Appendix A], we can evaluate $f_{M_T}(m)$ as

$$f_{M_T}(m) = \sum_{b=0}^{N_T} \pi_B(b)\, f_{M_T|B=b}(m|b) \tag{18}$$

where $f_{M_T|B=b}(m|b)$, the conditional composite traffic given b backlogged users at the beginning of the slot, is evaluated from

$$f_{M_T|B=b}(m|b) = \sum_{\max(0,\, m-b) \le n \le \min(N_T - b, m)} \Delta(m-n, b, p_r)\, \Delta(n, N_T - b, p_0) \tag{19}$$

with (p_0, p_r) the first transmission and retransmission probabilities, respectively. Note that (19) is true regardless of the jamming action. In (18), $\pi_B(b)$ is the appropriate given vector of the jamming-average matrix $\overline{P} = \{\overline{p_{ij}}\}$, i.e.,

$$\pi_B(b) = \pi_B(b)\, \overline{P} \tag{20}$$

where

$$\overline{p_{ij}} = \sum_{m_0 = \max(0,\, j-i)}^{N_T - i} \Delta(m_0, N_T - i, p_0) \cdot \sum_{m_b = \max(0,\, i-j)}^{i} \Delta(m_b, i, p_r)\, \overline{p}_{\,i+m_0-j+m_b} \tag{21}$$

It is clear from (18), (20), (21) that knowledge of the set $\{\overline{p_{slm}}\}$ is fundamental in the evaluation of the composite traffic. Assuming, as in (13), that all patterns $\underline{a}$ with the same Hamming weight $m_R^J(\underline{a})$ result in the same $p_{slm,\underline{a}}$, we can rewrite (12b) as

$$\overline{p_{slm}} = \sum_{m_R^J} \Pr\left[M_R^J = m_R^J\right] p_{slm,m_R^J} \tag{22}$$

In order to evaluate p_{slm,m_R^J}, we shall assume that the N_R receivers operate independently, conditioned on the jamming pattern. Let $\mathcal{R}^J \triangleq \{RCVR_k,\ k=1,...,m_R^J\}$ and $\mathcal{R}^{J^c} \triangleq \{RCVR_k,\ k=m_R^J+1,...,N_R\}$ indicate that jammed and non-jammed sets of receivers, respectively. Let $p_A^{R_k}$ generally stand for the probability of acceptance from the k^{th} RCVR's viewpoint. Under the symmetry assumption for each of the sets $\mathcal{R}^J, \mathcal{R}^{J^c}$, we then have that

$$p_A^{R_k} = \begin{cases} p_A^R(m;J), \text{ if } RCVR_k \in \mathcal{R}^J\ (k \le m_A^R) \\ p_A^R(m;J^c), \text{ if } RCVR_k \in \mathcal{R}^{J^c}\ (k > m_R^J) \end{cases} \tag{23}$$

Note that, in the present probabilistic framework, which individual RCVRs belong to the sets is irrelevant, and their ordering can be arbitrary; it is the set size $m_R^J = \dim \mathcal{R}^J$ that counts.

A recursive way to calculate p_{slm,m_R^J} is the following: let $p^{(k)}_{slm,m_R^J}$ denote the above probability of s successes, given m attempts and m_R^J jammed receivers, which is due to the first k receivers, $k = 0,...,N_R$. Clearly, the sought probability is simply $p_{slm,m_R^J} = p^{(N_R)}_{slm,m_R^J}$. We can first define

$$p^{(0)}_{slm,m_R^J} = \begin{cases} 1 & , s = 0 \\ 0 & , \text{otherwise} \end{cases} \tag{24a}$$

Now consider the first receiver $RCVR_1$. Then

$$p^{(1)}_{slm,m_R^J} = \begin{cases} 1 - p_A^{R_1} & , s = 0,\ m \ge 1 \\ p_A^{R_1} & , s = 1,\ m \ge 1 \\ 0 & , s \ge 2 \text{ or } m = 0 \end{cases} \tag{24b}$$

where $p_A^{R_1}$ has been defined in (23) as a function of m and m_R^J. Generalizing to the k^{th} RCVR, we can use a standard combinatorial method to conclude that

$$p^{(k)}_{slm,m_R^J} = \begin{cases} (1-p_A^{R_k})\, p^{(k-1)}_{0|m,m_R^J} & , s = 0 \\ \left[(1-p_A^{R_k}) + \left(\frac{s}{m}\right) p_A^{R_k}\right] p^{(k-1)}_{slm,m_R^J} & \\ + \left(\frac{m-s+1}{m}\right) p_A^{R_k}\ p^{(k-1)}_{s-1|m,m_R^J}, & 1 \le s \le \min(m,k) \\ 0 & , s > \min(m,k) \end{cases} \tag{24c}$$

The recursion stops when $k = N_R$. Note that two different quantities will be used in the place of $p_A^{R_k}$ above, depending on whether $k \le m_R^J$ or $k > m_R^J$, as per (23).

The above recursion can be used with any specific jamming number m_R^J to produce p_{slm,m_R^J} and then, via (22), $\overline{p_{slm}}$. It is also useful for calculating $\overline{p_{slm}}$ in a nonjammed environment (thermal and multi-user noise only), thus augmenting the theory of [PoSi87]. Regarding our previous two scenarios, we note that

$$\overline{p_{slm}} = p^{(N_R)}_{slm,N_R^J} \quad ; \qquad \text{scenario 2} \tag{25}$$

which is a direct result of (15b), (22) and (24). For scenario 1, it can be shown that $\overline{p_{slm}}$ can be obtained directly from the recursion (24) (i.e., substitute $p^{(k)}_{slm,m_R^J}$ for $\overline{p^{(k)}_{slm}}$ and identify $p^{(N_R)}_{slm,m_R^J} \equiv \overline{p_{slm}}$) so long as we use the jamming-average acceptance probability $\overline{p_A^R(m)}$ of (17) instead of $p_A^{R_k}$. This is also an intuitively appealing conclusion, in view of the fact that all RCVRs are mutually independent and statistically identical under scenario 1, each described probabilistically by $\overline{p_A^R(m)}$.

It is instructive to compare the above analysis for topology-selective jamming with the dual concept of *temporal selectivity*. We take this issue up in the following section.

4c. Comparison with Temporal Jamming

In our terminology, *temporal* implies a pulsed (blinking), two-level (ON-OFF) jamming pattern which, when ON, covers all local receivers under consideration with the same power J_{RCVR}. Let ρ_t indicates the temporal duty-factor of the jammer. It represents the long-term fraction of time that the jammer is ON, as well as the probability that a randomly observed slot will be found in the jamming state. As explained in [PrPo87], a variety of jamming waveforms can be constructed which have the same ρ_t, but different sample paths. In essence, one can vary the length (in slots) of the ON or OFF sessions by different probabilistic mechanisms, yet keep ρ_t fixed.

The two extreme cases, from a temporal variation or slot-correlation viewpoint, are (a) a long-term jammer, which stays in the same ON or OFF state for very long (practically infinite) intervals of time and (b) a slot-by-slot independent jammer with jamming probability ρ_t. The mathematical model for (a) is that of initially choosing between a "good" channel and a "bad" channel with probability ρ_t and $(1-\rho_t)$, respectively, and staying there forever. In all cases, we can express total throughput as

$$\beta = \rho_t \beta_J + (1-\rho_t)\beta_{J^c} \tag{26}$$

where the different temporal jamming strategies manifest themselves in the way we evaluate the conditional throughputs β_J and β_{J^c}. For the long-term jammer, it is immediate that

$$\beta_{\mathcal{J}} = \sum_m m p_s^T(m;\mathcal{J})\, f_{M_T}(m|\mathcal{J}); \quad \mathcal{J} = J \text{ or } J^c \tag{27}$$

where the symbol $\mathcal{J}$ is used as a jamming index; thus, (26) is easily evaluated, virtually by analyzing a multiple-access channel in two different interference levels.

It can be argued that (27) holds for all cases with finite jamming block size, exactly as (26) does, except that $f_{M_T}(m|\mathcal{J})$ should be interpreted as the conditional stationary distribution of transmitting users, evaluated for the state-$\mathcal{J}$ slots only ($\mathcal{J} = J$ or J^c). In other words, $f_{M_T}(m|\mathcal{J})$ represents the probability of having $M_T = m$ transmissions in a "typical" or randomly-chosen slot, assuming we only look at those slots of status $\mathcal{J}$. Unfortunately, evaluation of these two conditional stationary distributions is not trivial, necessitating the solution of a composite Markov chain of size which grows very quickly with complexity; the interested reader is referred to [PrPo87]. There is, however, one case which is significantly simpler, namely the slot-by-slot independent jamming of case (b) above. Then, a little thought will reveal that

$$f_{M_T}(m|J) = f_{M_T}(m|J^c) = f_{M_T}(m) \tag{28}$$

meaning that any slot is a "typical" slot, regardless of whether it belongs to a jammed or nonjammed block. This is precisely so by virtue of the memoryless property of the jamming mechanism from one slot to the next. Substituting property (28) into (26) and (27) (recall that the latter two equations hold for *any* scenario), we arrive at

$$\beta = \sum_m m\, \overline{p_s^T(m)}\; f_{M_T}(m) \tag{29a}$$

where

$$\overline{p_s^T(m)} \triangleq \rho_t p_s^T(m;J) + (1-\rho_t)\, p_s^T(m;J^c) \tag{29b}$$

is the *temporal* jamming-average probability of success.

Equation (29) is formally identical to (10)-(11), the apparent difference being that averaging here is performed over the temporal profile of the jamming process, as opposed to the spatial jamming profile of Section 3. The root of the similarity is, of course, the fact that the jamming decisions are independent from slot to slot. Furthermore, it is not hard to see that this special case of memoryless temporal jamming can be reformulated in the topology-selective framework by letting

$$\Pr[\Delta^J = \underline{a}] = \begin{cases} \rho_t & \text{, if } a_i = 1 \text{ for all } i \\ 1-\rho_t & \text{, if } a_i = 0 \text{ for all } i \\ 0 & \text{, otherwise} \end{cases} \tag{30}$$

since a substitution of (30) into (10) will immediately yield (29b).

The last step for case (b) regards the evaluation of $f_{M_T}(m)$ in (29a). If we incorporate the above conceptual linkage of equation (30) into the procedure outlined in section 4b for spatial jamming, we immediately conclude that $f_{M_T}(m)$ can be obtained from (18) – (21), with the temporal average

$$\overline{p_{stm}} = \rho_t p_{stm,J} + (1-\rho_t) p_{stm,J^c} \tag{31}$$

resulting from substituting (30) into (22). This is quite a convenient simplification which, as we mentioned, does not occur for jamming patterns with slot memory.

5. Numerical Results

As we mentioned, our main interest in this study is the topological aspects of the jamming threat. Thus, in the subsequent discussion, we shall assume some simple signal and jamming formats, namely full-band jamming (which can either be noise or spread tone) with (a) coded Direct-Sequence BPSK modulation and (b) Frequency Hopping MFSK modulation. The probabilities of a acceptance $p_A^R(m;J)$ and $p_A^R(m;J^c)$ are given by

$$p_A^R(m;\mathcal{J}) = \sum_{\ell=0}^{e} \binom{L}{\ell} p_{cs}^{\ell}(m;\mathcal{J})\,(1-p_{cs}(m;\mathcal{J}))^{L-\ell}; \quad \mathcal{J} = J \text{ or } J^c \tag{32}$$

for an e-error correcting code of block length L (also assumed to equal the packet length) and hard-decision, bounded-distance decoding. Here, $p_{cs}(m;\mathcal{J})$ is the channel symbol error rate, given by (for case (a))

$$p_{cs}(m;\mathcal{J}) = Q\left[\sqrt{\frac{2E_{cs}}{N_{oq}(m;\mathcal{J})}}\right] \tag{33}$$

where

$$\frac{E_{cs}}{N_{oq}(m;J)} = \frac{G_{baud}}{\gamma_{in}^{-1} + (m-1)\alpha_{ma} + J_{RCVR}(\rho_{sp})/S} \tag{34a}$$

$$\frac{E_{cs}}{N_{oq}(m;J^c)} = \frac{G_{baud}}{\gamma_{in}^{-1} + (m-1)\alpha_{ma}} \tag{34b}$$

and

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-z^2/2}\, dz \tag{35}$$

In the above equations,

$$G_{baud} \triangleq W_S T_{baud} \tag{36}$$

is the time-bandwidth product per baud or symbol, γ_{in} is the input signal-to-thermal-noise ratio over the whole spread bandwidth W_S

$$\gamma_{in} \triangleq \frac{S}{N_{0,th} W_S} = \frac{E}{N_{0,th} G_{baud}} \tag{37}$$

and α_{ma} is the multiple-access coefficient, which depends on the cross-correlation between the particular multiple-access codes in use. The numerical results are computed for $\alpha_{ma} = 1$ and $L = 1024$. For the extended BCH code of block length 1024, the number of correctable errors e and the coding rate can be approximately related by [Rayc 81]

$$r \cong 1 - \frac{11e}{1024} \tag{38}$$

In Figures 2 through 4, we consider *scenario 2* and plot the normalized throughput $r\beta(r)$ versus the spatial duty factor ρ_{sp}, with p_0 and p_r as parameters. The ten-error correcting code of

rate r = 0.89 is used. We see that if G_{band} = 13 dB, the worst-case ρ_{sp} approaches 1 as p_0 and p_r approach 0.5, whereas if G_{band} = 20 dB, the worst-case ρ_{sp} stays around 0.5 for p_0 and p_r less than 0.5. These observations can be explained as follows: The multiple-access (muti-user) noise remains small in comparison to the noise level, whenever G_{band} is large or, if G_{band} is small, when the channel-traffic is light (implying small values of p_0 and p_r). In such an environment, the worst-case ρ_{sp} is less than one (i.e., selective jamming) because a significant amount of power is needed per targeted RCVR in order to cause decoding errors, despite the fact that the remaining non-targeted RCVRs will mostly decode correctly. For instance, at the minimum value of Figure 4, $p_A^R(5,J) = 1.9 \times 10^{-2}$ whereas $p_A^R(5,J^c) = 1$. On the other hand, if the average multiple-access interference is large and is sufficient to cause many errors on its own, the jammer is better-off spreading the power evenly upon all RCVRs (ρ_{sp}=1), bringing about the most damage on the total network with the small additional jamming power allotted per RCVR. Of course, the same uniform jamming strategy will be optimal for large jamming-power levels regardless of the multi-user noise level.

In Figures 2 through 4, we show the normalized throughput versus the spatial duty factor ρ_{sp}, parameterized by p_0. The retransmission probability p_r is optimized for the controlled protocol in these three figures. We see that the effectiveness of the worst-case jammer is slightly reduced for many cases by optimizing p_r.

In Figure 5, we show the normalized throughput versus the jammer spatial duty factor ρ_{sp} with $p_0 = p_r = 0.3$ for several error-correction code rates. In the same figure, we also show the normalized throughput for the paired-off case. Notice that the competitive scenario is more robust against the worst-case jamming than the paired-off scenario.

In Figure 6, we show the normalized throughput versus the jammer spatial duty factor ρ_{sp}, parameterized by p_0, for *jamming scenario 1*. The retransmission probability p_r is optimized in this figure. Compared with Figure 4, we see that the worst-case jammer in scenario 1 is slightly more effective than that of scenario 2.

In Figure 7, we show the average packet delay versus the jammer spatial duty factor ρ_{sp} for jamming scenario 2. The controlled protocol is considered and the retransmission probability p_r is optimized in this figure.

In Figure 8, we show the normalized throughput versus the jammer spatial duty factor ρ_{sp} for a Frequency-Hopping, TR-based, multiple-access network. We note that certain constants have been neglected in the normalized throughput $r\beta(r)$; in other words, this should not be equated with the system utilization, nor should the Frequency-Hopping performance be compared to the direct-sequence one.

Let us now switch attention to the non-dedicated or half-duplex scenario.

6. Half-Duplex Nodes

The model we consider here is of the TR-*priority* type, meaning that whenever a unit has a packet scheduled for channel access, the transmitter-function takes over and an active transmission ensues, blocking the receiving capability of the unit under consideration at that point in time. If there is no transmission scheduled for a particular slot, so that the unit is not an active TR in that slot, then it is set in the receiving mode. It follows that the random numbers of active TRs (M_T) and active RCVRs (M_R) are related by

$$M_R = U - M_T \tag{39}$$

Note that if $U = M_T$ in a particular slot, implying that all units act as TRs , then there will be no active RCVRs available to capture the transmitted packets. The implications of such events will be analyzed below.

We will proceed with our analysis in two steps: first, we discuss network performance when there is only multi-user interference plus thermal noise (part 6a). Then, we shall also address the jamming case in part 6b.

6a. Noise-Only Performance

When thermal noise is the only deterrent in addition to the omni-present multi-user interference, then the required analytical expressions are a straightforward extension of the results in [PoSi87], with N_R substituted by M_R as per equation (39). For instance, the basic equation (4) of [PoSi87, Proposition 1] for the throughput β remains the same:

$$\beta = \sum_{m_T=1}^{N_T} m_T \, p_s^T(m_T) \, f_{M_T}(m_T) \tag{40}$$

where now

$$p_s^T(m_T) = 1 - \left(1 - \frac{p_A^R(m_T)}{m_T}\right)^{U-m_T} \tag{41}$$

assuming, as in [PoSi87, Proposition 2], independently operating receivers and success per packet at most one. The analytic procedure for computing $f_{M_T}(m_T)$ via the appropriate Markovian model is identical to [PoSi87, Appendix A], except that the fundamental set of probabilities $\{p_{s|m_T}\}$ should be evaluated based on $m_R = U - m_T$. The recursive algorithm of section 4b also holds, except that the quantity N_R should be simply substituted by m_R which, given the quantity m_T in a particular slot, can be considered fixed.

6b. Noise Plus Jamming

We adopt here the same stochastic, slot-by-slot independent, topology-selective jamming model of section 3. Again, the jamming status of the network in a particular slot is summarized by the indicator vector $\underline{A}_U^J$, whose (0,1)-valued components A_i^J ; i=1,..,U indicate whether a unit is jammed or not. The jamming strategy is again manifested in the joint probability mass distribution function (pdf) $\Pr[\underline{A}_U^J = \underline{a}_U]$ of $\underline{A}_U^J$. However, there is a significant modeling difference in this half-duplex case compared to the dedicated-RCVR case: we will assume that the jammer does not know which units will be the active TRs in a given slot; consequently, he cannot know which are the active RCVRs in order to target only those. This is very different from the previous case, where the set of dedicated RCVRs was assumed fixed, known, and exclusively targeted by the jammer. In the absence of similar information, the jammer will randomly select a subset of units to jam in each slot, despite the fact that some of that power might go wasted on active TRs within that slot. The jammer has no alternative but to suffer this random loss per slot. Of course, the jammer still has at his disposal for optimization the probabilistic law $\Pr[\underline{A}_U^J = \underline{a}_U]$, and we examine below this optimization aspect.

The first consequence of the aforementioned modeling difference is the fact that the spatial duty factor ρ_{sp} of equation (3) cannot be defined explicitly here, simply because the number of active RCVRs per slot is not fixed but random. Instead, we define the *half-duplex spatial duty factor* ρ_{sp}^{hd} as the average fraction of the units jammed per slot, i.e.,

$$\rho_{sp}^{hd} = \frac{\mathcal{E}\{M_U^J\}}{U} \tag{42}$$

where M_U^J is the r.v. signifying the number of such jammed units per slot. This quantity is a direct byproduct of the stochastic jamming law and can be chosen by the jammer in an optimal way. It can be argued, of course, that to every

probabilistic jamming choice (and its concomitant ρ_{sp}^{hd}) there corresponds an *effective* jamming duty factor, which would be the average fraction of the jammed *active* RCVRs per slot. This can be computed after the system of equations for throughput, composite traffic, etc, has been solved and the average number of active transmissions per slot has been determined. However, this is an indirect result of the total model and not something that can be determined apriori, as ρ_{sp}^{hd} can; hence, we opt to express everything in terms of the latter.

Again, a number of different probabilistic jamming scenarios can be devised which correspond to the same value of ρ_{sp}^{hd}. Presumably, each such scenario would produce a different throughput for the system, and an exhaustive search would require optimizing over ρ_{sp}^{hd} as well as the underlying jamming pdf. However, in view of the numerical conclusion of section 5, the difference in throughput between the two scenarios was rather insignificant, once the jamming duty factor was kept the same. We will only consider the Bernoulli model, since it is somewhat more amenable to analysis. Thus, we assume that each unit is jammed with probability p_j independently of all others. It follows that $\mathcal{E}\{M_U^J\} = p_j U$, so that $\rho_{sp}^{hd} = p_j$.

A second major analytical consequence of the considered jamming model is that the jamming status of the set of active RCVRs in each slot is *not* statistically independent of the number of active TRs M_T, as in the previous dedicated case, simply because M_R depends on M_T via (39). We must, therefore, rework the derivation of throughput in section 4 in order to account for this fact.

In particular, let $\Delta_{U-M_T}^J$ denote the jamming indicator vector within the random set of $M_R = U - M_T$ active RCVRs in a slot, with $\Pr[\Delta_{M_R}^J = a_{M_R}^J | M_T] \triangleq \Pr[A_1^J = a_1, \ldots, A_{U-M_T}^J = a_{U-M_T} | M_T]$ denoting the conditional probability distribution function of that vector, given M_T active TRs in the slot. Following the steps of equation (9), we can write

$$\begin{aligned}\beta = \mathcal{E}\{S\} &= \mathcal{E}_{M_T,\,\Delta_{U-M_T}^J}\left\{\mathcal{E}\{S|M_T,\Delta_{U-M_T}^J\}\right\} \\ &= \mathcal{E}_{M_T}\left\{\mathcal{E}_{\Delta_{U-M_T}^J|M_T}\,\mathcal{E}\{S|M_T,\Delta_{U-M_T}^J\}\right\} \\ &= \sum_{m_T=0}^{N_T} m_T f_{M_T}(m_T) \\ &\quad \cdot \sum_{a_{m_R}\,(m_R=U-m_T)} \Pr\left[\Delta_{m_R}^J = a_{m_R} | M_T = m_T\right] p_s^T(m_T, a_{m_R})\end{aligned} \tag{43}$$

where $p_s^T(m_T, a_{m_R})$ is the probability of success from a typical TR's viewpoint, given another $(m_T - 1)$ packets and a specific jamming pattern $a_{m_R} = a_{U-m_T}$ in a slot. Again, if we define the conditional version of equation (10), namely

$$\overline{p_s^T(m_T)} \triangleq \sum_{a_{U-m_T}} \Pr\left[\Delta_{U-m_T}^J = a_{U-m_T}^J | m_T\right] p_s^T(m_T, a_{U-m_T}) \tag{44}$$

we arrive at the same equation (11). It follows that equation (12a) remains the same, where now

$$\overline{p_{slm_T}} = \sum_{a_{U-m_T}} \Pr\left[\Delta_{U-m_T}^J = a_{U-m_T} | m_T\right] p_{slm_T, a_{U-m_T}} \tag{45}$$

Furthermore, utilizing the approximation of conditional independence between RCVRs, we arrive at a slightly modified version of (14):

$$\begin{aligned}&\overline{p_s^T(m_T)} \\ &= \sum_{m_R^J} \Pr\left[M_R^J = m_R^J | m_T\right] \\ &\cdot \left[1 - \left(1 - \frac{p_A^R(m_T;J^c)}{m_T}\right)^{U-m_T-m_R^J}\left(1 - \frac{p_A^R(m_T;J)}{m_T}\right)^{m_R^J}\right]\end{aligned} \tag{46}$$

To proceed, we must evaluate the conditional quantity $\Pr[M_R^J = m_R^J | m_T]$ in (46). For any specific stochastic jamming law, this is the solution to a combinatorial problem which can be addressed with the help of the Vennian diagram of Figure 9 and the total-probability law

$$\Pr[M_R^J = m_R^J | m_T] = \sum_{m_U^J} \Pr[M_U^J = m_U^J]\,\Pr[m_R^J | m_T, m_U^J] \tag{47}$$

where $\Pr[M_U^J = m_U^J]$ depends solely on the jamming stochastic model. Although this can be a fairly complicated task, things simplify significantly in a Bernoulli jamming scenario, because then

$$\Pr\left[M_R^J = m_R^J | m_T\right] = \begin{cases} \Delta\left(m_R^J, U-m_T, \rho_{sp}^{hd}\right) & ; m_R^J \le m_R = U - m_T \\ 0 & ; \text{otherwise} \end{cases} \tag{48}$$

The combination of (46) and (48) yields an expression identical to (16) for scenario 1, except that ρ_{sp} and N_R are substituted by ρ_{sp}^{hd} and $U - m_T$, respectively. We note that Scenario 2 would be considerably more complicated to analyze, although the basic theoretical path we have provided herein should be adequate for this task.

As we conclude in section 4a, the above results suffice to calculate the throughput in (11) for case where $p_0 = p_r = p$. The more general case $p_0 \neq p_r$ requires the evaluation of $\overline{p_{slm_T}}$ and $f_{M_T}(m_T)$ as per section 4b. A careful review of the analysis therein for Scenario 1 will convince us that all the algorithms remain intact, except that N_R should be substituted by $U - m_T$. In other words, the recursion in equation (24) will indeed produce $\overline{p_{slm_T}}$, if we replace $p_{slm,m_R^J}^{(k)}$ by $p_{slm_T,m_R^J}^{(k)}$ and $p_{slm,m_R^J}^{(U-m_T)}$ by $\overline{p_{slm_T}}$, as long as we use the average acceptance probability

$$\overline{p_A^R(m_T)} = \rho_{sp}^{hd}\,p_A^R(m_T;J) + (1 - \rho_{sp}^{hd})\,p_A^R(m_T;J^c)$$

instead of $p_A^{R_k}$.

7. Numerical Results for Half-Duplex Networks

The theory developed herein has been applied to a fully-connected network with U=10 half-duplex units. Figures 10 and 11 show the normalized throughput $r\beta(r)$ versus the jamming duty-factor, parameterized by the new-packet transmission probability p_0. In Figure 10, the uncontrolled ($p_0 = p_r$) ALOHA protocol is considered. In Figure 11, the controlled ($p_0 \neq p_r$) ALOHA protocol is considered, where the retransmission probability p_r has been optimized. The system parameters of both figures are identical to those in Figure 4.

Certain interesting conclusions can be drawn from these figures: first, there is again an optimal ρ_{sp}^{hd} which, for these parameter values, is around $\rho_{sp}^{hd} = 0.5$. Furthermore, comparing

Figures 10 and 11 with Figure 4, we conclude that half-duplex has consistently lower throughput than the full-duplex version, as expected. The throughput in Figures 10 and 11 reaches its asymptotic maximum at $p_0 \approx 0.5$ (for the whole range of ρ_{sp}^{hd}); in fact, no noticeable gain exists in the range $p_0 \approx 0.3 \sim 0.5$, particularly around the minimum value. The throughput for the uncontrolled protocol eventually decreases as p_0 is larger than 0.6. This is because when p_0 and p_t are large, only a few units can be RCVR's; and thus the throughput is reduced. On the other hand, the throughput for the controlled and optimized protocol keeps increasing as p_0 increases. This is because when p_0 becomes larger, the corresponding optimum p_t becomes smaller; and thus the throughput is kept essentially constant.

In Figure 12, we show that the worst case ρ_{sp}^{hd} is a function of the signal-to-jammer power ratio when the other parameters are fixed.

8. Conclusions

The problem of jamming fully-connected CDMA networks has been addressed by devising appropriate models for stochastic jamming alternatives. A procedure for analyzing the throughput/delay performance of such networks has been proposed, and employed to identify optimal parameter values both from the user's and the jammer's viewpoint. The optimized variables were (a) the spatial duty factor from the jamming side and (b) the coding rate plus link-access parameters of an ALOHA protocol, from the users' side. The impact of half-duplex units versus dedicated TRs and RCVRs was also examined.

Certain other types of networks would more or less fit the present model and can be analyzed by similar methods, such as RCVR-based codes, buffered users, etc. Major departures from this model which are of interest are (a) multihop topologies with end-to-end performance measures and (b) adaptive schemes, where both users and jammer rely on channel observables to adjust their strategies. In general, the superposition of CDMA techniques and jamming can be expected to have an impact on the reliability of such observables, which makes the analysis and design of these techniques an even more challenging problem.

References

[DaGr80] D. H. Davis and S. A. Gronemeyer, "Performance of Slotted ALOHA Random Access with Delay Capture and Randomized Time of Arrival," IEEE Trans Comm., vol. COM-28, pp. 703–710, May 1980.

[Jubi85] J. Jubin, "Current Packet Radio Network Protocols," Proceedings of INFOCOM '85, pp. 86 – 92, April 1985.

[KaGrBuKu78] R. Kahn, S. Gronemeyer, J. Burchfiel, and R. Kunzeman, "Advances in Packet Radio Technology," Proc. IEEE, vol. 66, pp. 1468 – 1496, November 1978.

[KlLa75] L. Kleinrock and S. Lam, "Packet Switching in a Multiaccess Broadcast Channel: Performance Evaluation," IEEE Trans. Comm., vol. COM-23, pp. 410 – 423, April 1975.

[PoSi87] A. Polydoros and J. Silvester, "Slotted Random Access Spread-Spectrum Networks," IEEE Journal on Selected Areas in Comm., Vol. SAC-5, No. 6, July 1987.

[PrPo87] N. Pronios and A. Polydoros, "Spread-Spectrum, Slotted-ALOHA, Fully-Connected Networks in Jamming," submitted to IEEE Trans. Comm.; see also preliminary version in MILCOM '87.

[Purs87] M. B. Pursley, "The Role of Spread Spectrum in Packet Radio Networks," Proceedings of the IEEE, vol. 75, pp. 116 – 134, January 1987.

[Rayc81] D. Raychaudhuri, "Performance Analysis of Random Access Packet Switched Code Division Multiple Access System," IEEE Trans. Comm., vol. COM-29, pp. 895–901, June 1981.

[ShTo85] N. Shacham and J. D. Tornow, "Future Directions in Packet Radio Technology," Proceedings of INFOCOM '85, pp. 93 – 98, April 1985.

[SoSi84] E. S. Sousa and J. A. Silvester, "A Spreading Code Protocol for a Distributed Spread Spectrum Packet Radio Network," in Proc. GLOBECOM '84, Atlanta, GA, December 1984.

[SuLi85] S. L. Su and V. O. K. Li, "An Iterative Model to Analyze Multi-Hop Packet Radio Networks," in Proc. Allerton Conf., Oct. 1985, pp. 545–554.

[Wu84] W. W. Wu, Elements of Digital Satellite Communications, Computer Science Press, Rockville, Maryland, 1984.

Acknowledgement

This work was performed for Axiomatix Corp., sponsored by ARO, under Contract No. DAAL03-87-C-0007.

Table 1. Key Parameters.

U	Total number of radio units in the local channel (fixed)
N_T	Maximum number of potential transmitters in a slot (fixed)
N_R	Maximum number of potential receivers in a slot (fixed)
M_T	Number of active transmitters in a particular slot (r.v.)
M_R	Number of active receivers in a particular slot (r.v.)
B	Number of backlogged users at the beginning of a slot (r.v.)
M_B	Number of backlogged users retransmitting in a slot (r.v.)
M_0	Number of new users transmitting in a slot (r.v.)
S	Number of "channel successes" in a slot (r.v.)
$f_{M_T}(m)$	Prob{$M_T = m$}, unconditional pdf
$\Delta(\xi, \Xi, p)$	Pr{ξ successes in Ξ trials}, binomial with parameter p

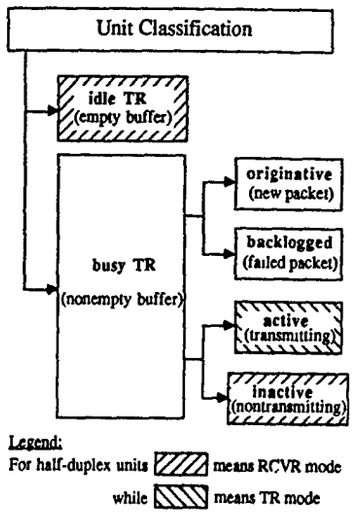


Figure 1. General classification of packet-radio unit status.

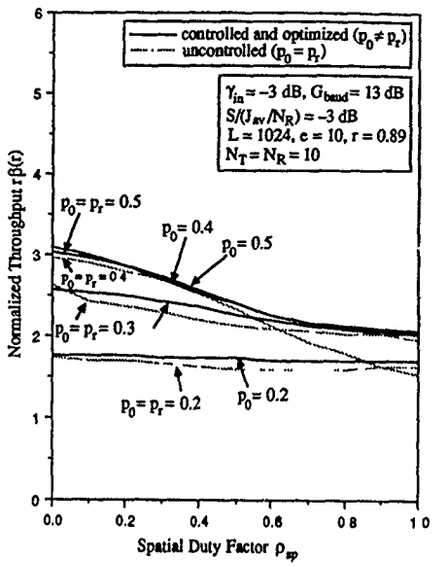


Figure 3. Normalized throughput versus spatial duty-factor for full-duplex units, scenario 2.

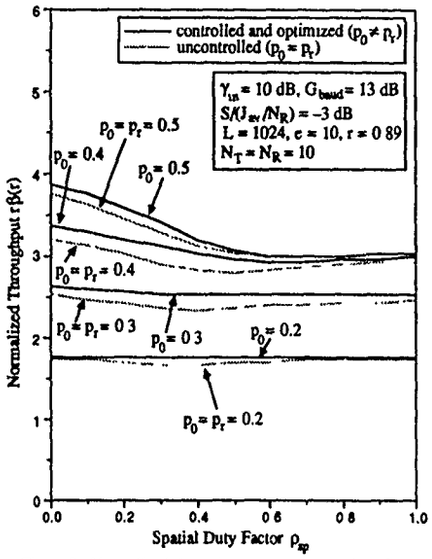


Figure 2. Normalized throughput versus spatial duty-factor for full-duplex units, scenario 2.

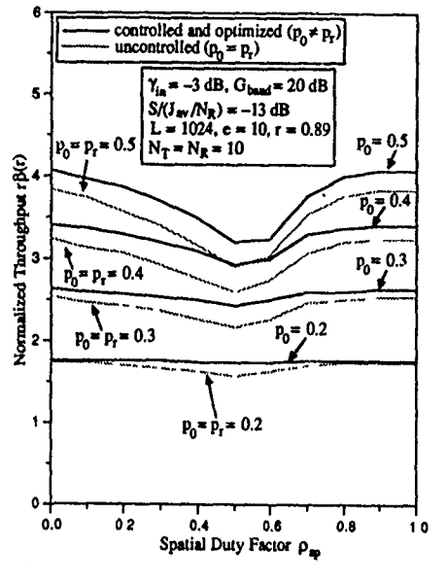


Figure 4. Normalized throughput versus spatial duty-factor for full-duplex units, scenario 2.

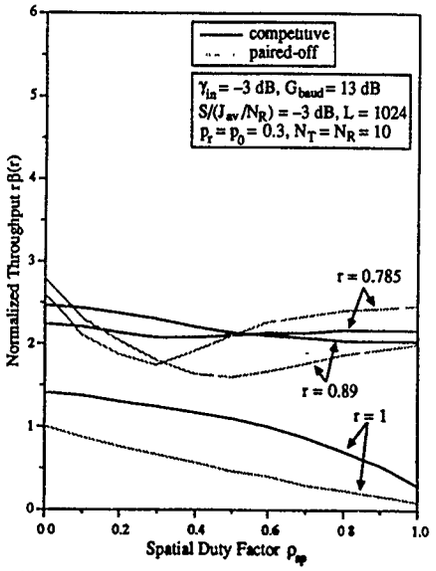


Figure 5. Normalized throughput versus the spatial duty factor ρ_{sp} for several error-correction codes for jamming, scenario 2.

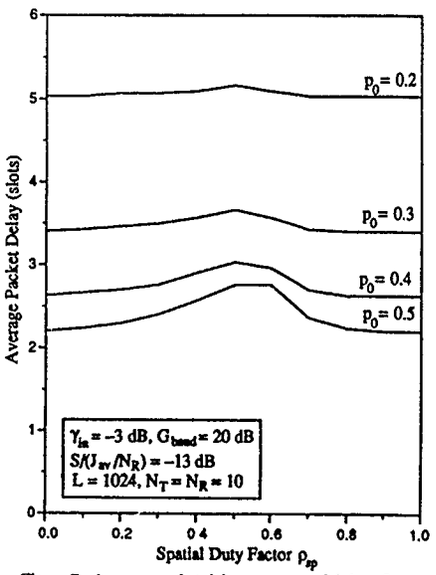


Figure 7. Average packet delay versus spatial duty-factor for full-duplex units, scenario 2, control and optimized ($p_0 \neq p_r$).

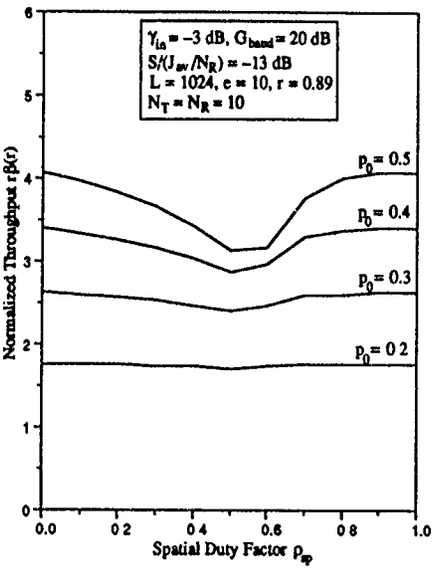


Figure 6. Normalized throughput versus the spatial duty factor ρ_{sp} for full-duplex units, scenario 1, control and optimized ($p_0 \neq p_r$).

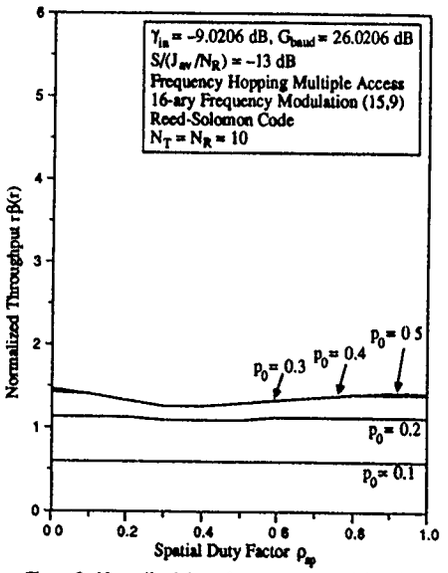


Figure 8. Normalized throughput versus spatial duty-factor for full-duplex units, scenario 1, control and optimized ($p_0 \neq p_r$).

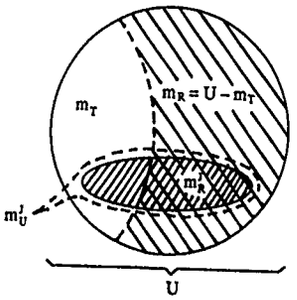


Figure 9. Vennian Diagram for Evaluating $Pr[m_k^I | m_T, \bar{m}_U^I]$.

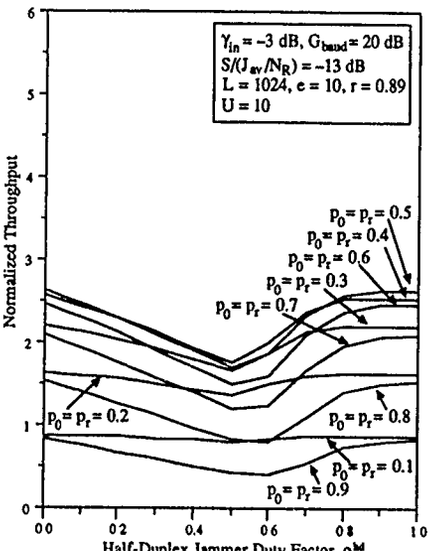


Figure 10. Normalized throughput versus spatial duty-factor for half-duplex units, scenario 1, uncontrolled ($p_0 = p_r$).

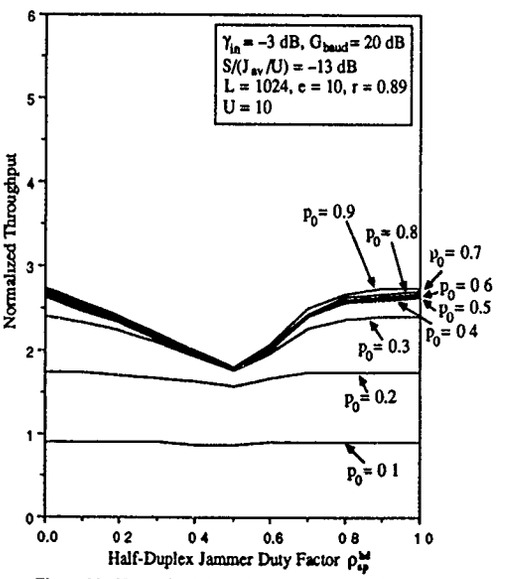


Figure 11. Normalized throughput versus spatial duty-factor for half-duplex units, scenario 1, controlled and optimized ($p_0 \neq p_r$).

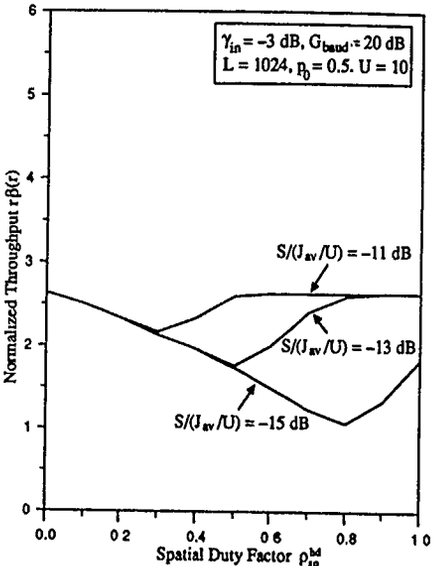


Figure 12. Normalized throughput versus spatial duty-factor for half-duplex units, scenario 1, controlled and optimized ($p_0 \neq p_r$).

DISCUSSION

D. YAVUZ

In your generic development you invoked the independence of "packet acceptance" conditioned on a given jamming strategy. Given a "jamming type", the packet reception/acceptance of the jammed modes will be totally correlated, i.e. no reception at all for that subset of modes, hence not independent ! Could you clarify ?

AUTHOR'S REPLY

Dr YAVUZ, what I meant is that, once we condition on α jammming pattern in a specific slot, $A^J = a$, then jamming is not stochastic for that slot. What remains is randomness due to all other nuicauses (noise, multi-user interference, fading etc...). The statement means that the conditional acceptance event is independent from one receiver to the other, due only to these other disturbances. Unconditional acceptance is highly correlated, as you noted above.

C. PHILIPPIDES

Have you considered encoding techniques which may be self synchronized based on modulation character of the transmitting signal which show channel reception improvement ?

And if so have you noticed if applying encryption techniques minimise jamming of the channel ?

AUTHOR'S REPLY

The model is very general with respect to the particular modulation/synchronization/encryption mecanism used. Each such combination will induce a different Pr (packet acceptance). The analytic framework we proposed can be used to study all these combinations (please contact me if that is of interest to you !).

SOME CONSIDERATIONS CONCERNING LOW NOISE RADIO RECEIVING SYSTEMS

by

Knut N. Stokke
Norwegian Telecommunications
Regulatory Authority
Parkveien 57
N-0203 OSLO 2
Norway

INTRODUCTION

In order to get a good signal to noise ratio at very high radio frequencies (higher than 20 - 40 MHz), low noise amplifiers are often used. Such amplifiers should have a very low total noise factor (noise figure). Normally it is the noise factor of the first stage of an amplifier which is decisive for the total noise factor.

Losses in a system will also influence the noise factor of a system. Consequently the losses in cables and waveguides must be taken into account when considering the total noise factor.

The thermal noise radiation from the surroundings may increase the noise temperature of an antenna. For reflector antennas, such as parabolic reflector antennas, the noise temperature is dependent on the elevation angle, and also dependent on the form of the surroundings, mountains, hills, buildings, etc.

The measurements referred to in this article, were done by the Radio Interference Division in the Norwegian Telecom.

TECHNICAL CONSIDERATIONS

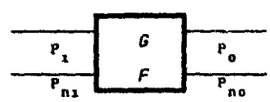


Figure 1. Four-terminal network.

In Figure 1 is sketched an element of a communication system. The noise factor F for this two-port may be defined as the ratio between the signal to noise ratio at the input, and the signal to noise ratio at the output. We then have:

$$F = \frac{P_i/P_{ni}}{P_o/P_{no}} \qquad \text{1.}$$

where P_i is the input signal power, P_o is the output signal power, P_{ni} is the noise power at the input, and P_{no} is the noise power at the output.

The power gain of this element is:

$$G = \frac{P_o}{P_i} \qquad \text{2.}$$

and if Equation 2 is introduced into Equation 1, we have:

$$F = \frac{1}{G} \cdot \frac{P_{no}}{P_{ai}} \qquad \text{3.}$$

that is, the noise factor is inversely proportional to the power gain of an element. This means that it is necessary to have a reasonable gain in an element when we want a low noise factor.

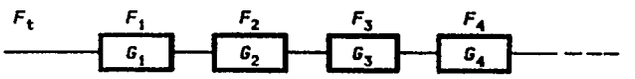


Figure 2. Two-ports in series.

When we have several two-ports in series, the total noise factor is (Reference 1):

$$F_t = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 \cdot G_2} + \frac{F_4 - 1}{G_1 \cdot G_2 \cdot G_3} + \text{---} \qquad \text{4.}$$

where F_1, F_2, F_3, --- are the noise factors and G_1, G_2, G_3, --- are the power gains of the elements.

If we want to use the formula on a receiving system, we may have conditions as indicated in Figure 3.

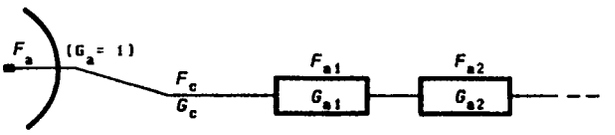


Figure 3. Antenna, cable, and amplifiers.

In Figure 3, F_a is the antenna noise factor. The noise temperature of the antenna is due to the total thermal noise radiation into the antenna, and there is no gain ($G_a = 1$) concerning noise considerations.

G_c is the power gain of the cable or waveguide, that is, G_c is less than 1 which means attenuation.

F_c is the noise factor of the cable, and is equal to the attenuation, $1/G_c$. This may be explained by the fact that even if the signal through the cable is attenuated by $1/G_c$, the thermal noise remains the same because of constant impedance.

F_{a1} is the noise factor and G_{a1} is the power gain of the first amplifier, F_{a2} is the noise factor and G_{a2} is the power gain of the second amplifier, and so on.

If we introduce these values into Equation 4, we have:

$$F_t = F_a + \frac{F_c - 1}{1} + \frac{F_{a1} - 1}{G_c} + \frac{F_{a2} - 1}{G_c \cdot G_{a1}} + \text{---} \qquad \text{5.}$$

If the cable attenuation in dB is called α_c, we have:

$$\alpha_c = 10 \log 1/G_c \quad \text{or:} \quad G_c = \frac{1}{10^{\alpha_c/10}}$$

and consequently:

$$F_c = 10^{\alpha_c/10}$$

We also have that (Reference 1):

$$T_a = (F_a - 1)T_0$$

$$F_a = \frac{T_a + T_0}{T_0} \qquad \text{---------- 6.}$$

where T_a is the antenna noise in °Kelvin, T_0 is the reference temperature in °K, and F_a is the antenna noise factor.

If the noise factors are low and the gain of the amplifiers is relatively large, we may with good accuracy only take account of the influence of the first amplifier on the total noise factor. And if we then use 290°K (17°C) as reference temperature, we have for the conditions given in Figure 3:

$$F_t = \frac{T_a + 290}{290} + 10^{\alpha_c/10} - 1 + \frac{F_{a1} - 1}{\frac{1}{10^{\alpha_c/10}}}$$

$$F_t = \frac{T_a}{290} + F_{a1} \cdot 10^{\alpha_c/10} \qquad \text{---------- 7.}$$

If the low noise amplifier is inserted just after the antenna as indicated in Figure 4, th. total noise factor is (see Equation 4):

$$F_t' = \frac{T_a + 290}{290} + \frac{F_{a1} - 1}{1} + \frac{10^{\alpha_c/10} - 1}{G_{a1}}$$

$$F_t' = \frac{T_a}{290} + F_{a1} + \frac{10^{\alpha_c/10} - 1}{G_{a1}} \qquad \text{---------- 8.}$$

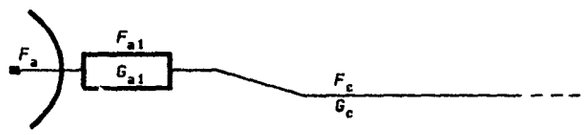


Figure 4. Antenna, amplifier, and cable.

In Figure 5 are given curves for Equations 7 and 8 for different amplifier noise factors F_{a1}. G_{a1} is here 16 (12 dB). The cable attenuation α_c varies from 0 dB to 1.5 dB. The antenna noise temperature is here 30°K, which may be an average antenna noise temperature in nordic countries for national

broadcasting satellite reception in the 12 GHz band (Reference 2). We observe that losses before the low noise amplifier may considerably increase the total noise factor, even if the losses are relatively low.

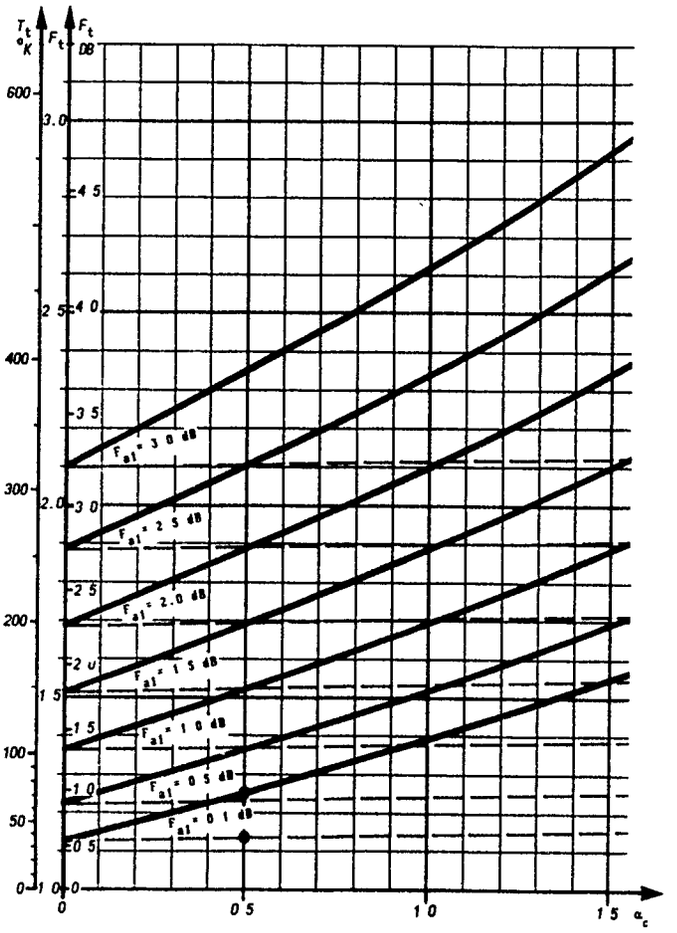


Figure 5. Curves for Equations 7 and 8 when the antenna noise temperature is 30°K. (Equation 7: ————, Equation 8: - - - - - - -).

Especially when the amplifier noise factor is very low (cooled amplifiers), for example 0.1 dB, losses before the amplifier may spoil the possibilities. If we have a loss of 0.5 dB after the amplifier, the noise temperature T_t for the system is about 35°K. But if there is a loss of 0.5 dB before the low noise amplifier, the noise temperature T_t for the system will be about 70°K.

This example also indicates that it is advantageous to use low noise preamplifiers near the antenna when we want to improve the signal to noise ratio at the higher radio frequencies. And when the losses are of the order of the order of 1 dB in the waveguide, which may often be the situation at a radio link station, it may be more advantageous to use an amplifier with noise factor F_{a1} = 1 dB near the antenna than an amplifier with F_{a1} = 0.1 dB after the waveguide.

Until now we have looked at losses in cables and waveguides. However, other types of losses, as for instance mismatch losses, have the same destruktive influence on the total noise factor.

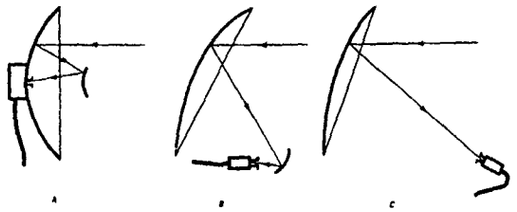


Figure 6. Cassegrain antenna and offset antenna.

Every object brought into an electromagnetic field represents a loss. For instance an antenna sub-reflector in a Cassegrain antenna (Figure 6A) is a lossy element, and at the same time the sub-reflector disturbes the electromagnetic field. In addition such a sub-reflector is not uniformly illuminated, and a part of the energy passes outside the sub-reflector causing a reduction in the signal strength. We see here that it is difficult to separate between a loss in signal strength and an increase in thermal noise. But whatever the reason for the losses or reduction in signal strength, they have an influence on the total signal to noise ratio.

An antenna with offset sub-reflector, as indicated i Figure 6B, will not disturb the incoming field, and the sub-reflector may be used to get a better illumination of the antenna reflector. However, the effect of the illumination factor of the sub-reflector is still there.

It is important to get the signal from the antenna to the low noise amplifier without unnecessary losses. For a reflector antenna it would be advantageous to let the signal go directly to the waveguide input of the amplifier. Previously these amplifiers had rather large dimensions, and they had to be placed at the rear of the antenna. Cassegrain antennas were therefore well fitted under such conditions. However, low noise amplifiers (and converters) are now so small that they may be placed in the front of the antenna without causing any severe disturbance in the electromagnetic field.

An offset antenna may use the amplifier/converter directly in the focusing area (Figure 6C), thus giving an effective antenna. However, we have to remember that the construction of an offset antenna reflector is more complicated than that of a parabolic antenna reflector. And a simple parabolic reflector antenna with preamplifier gives so good results that it is not so easy to get higher efficiency (illumination factor for a normal receiving parabolic antenna is about 0.5, may be optimized to 0.55 - 0.60).

There is a type of antenna which may have higher efficiency, namely the horn antenna. The illumination factor may here be more than 0.7. Large horn antennas are expensive, and because of the space needed for such antennas, they are used mostly in special cases. However, for frequencies higher than about 20 GHz even small horn antennas may have adequate gain for receiving systems.

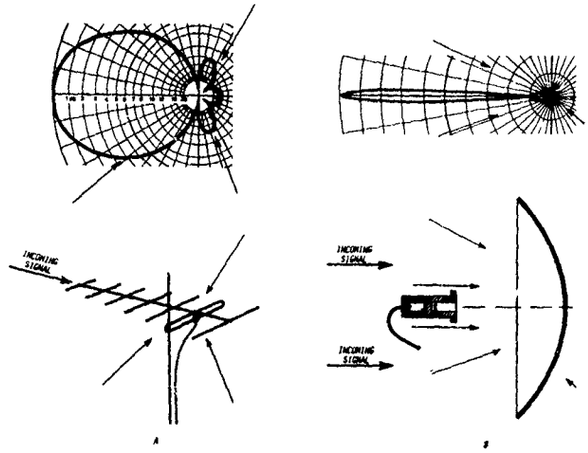


Figure 7. Yagi antenna and parabolic reflector antenna.

Concerning the thermal noise coming in to an antenna, the thermal noise pattern for the antenna may be somewhat different from the ordinary electric field pattern dependent on the conditions at the measuring site. A Yagi antenna has a very broad main lobe and the first sidelobes may not be more than 10 - 15 dB below the maksimum, as indicated in Figure 7A. The noise temperature for such an antenna in the VHF - UHF band is therefore about the same as the total noise temperature in the surroundings at these frequencies, that is about 300°K.
In order to illustrate how the total noise factor F_t varies when the antenna noise temperature is 300°K, curves for Equations 7 and 8 are given in Figure 8. These curves are parallel to the curves in Figure 5 (parallel displaced by 300°K - 30°K = 270°K).

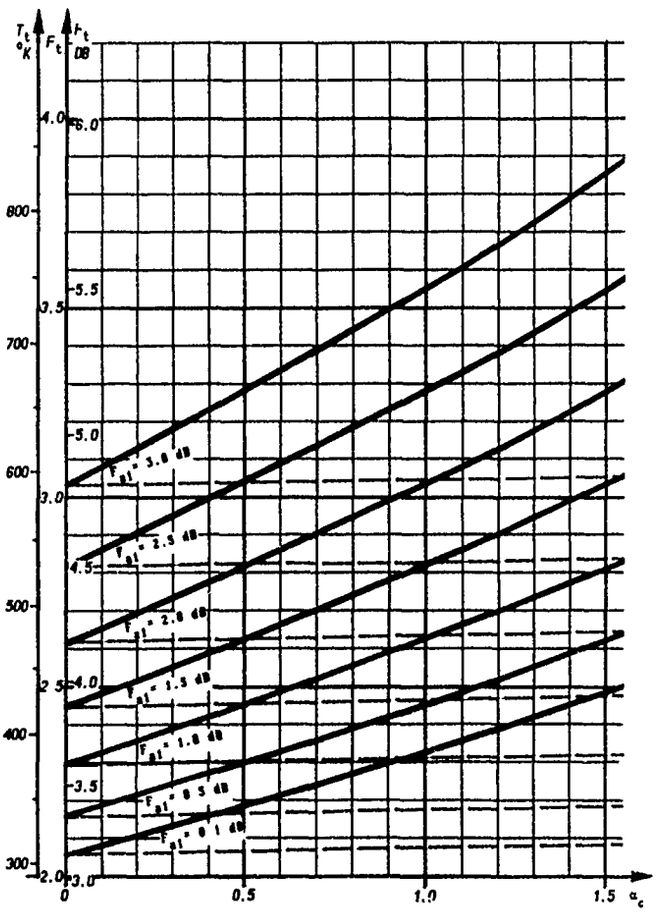


Figure 8. Curves for Equations 7 and 8 when the antenna noise temperature is 300°K. (Equation 7: ————, Equation 8: -------).

Parabolic reflector antennas have a very good directivity and small sidelobes, and the antenna noise temperature varies with the elevation angle. The noise temperature for an antenna in the 12 GHz band may have variations as shown in Figure 9 (Reference 2). When directed upwards to clear cold sky the antenna noise temperature may be as low as 8 - 10°K. Directed to the horizon, the noise temperature of a reflector antenna may increase to about 250°K.
The curve in Figure 9 is for flat earth. When there are hills, buildings, etc., near the antenna, the conditions may be changed. It has often been said that narrow beam reflector antennas may be mounted very low near the ground. This is true when we receive high signal intensities. However, if we want to receive signals near the noise level, we have to carefully avoid additional noise.

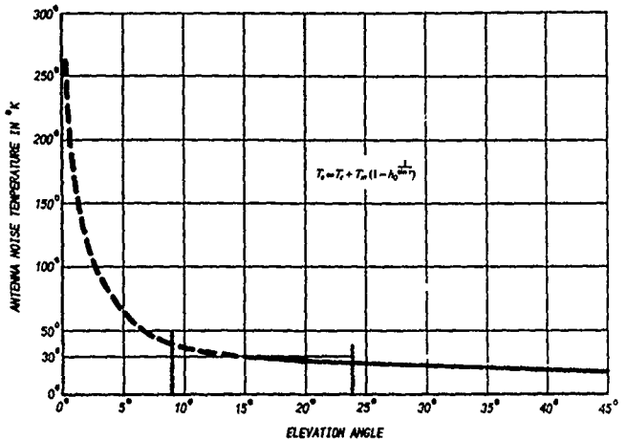


Figure 9. Curve for the antenna noise temperature at 12 GHz.

Even if the sidelobes of parabolic reflector antennas may be more than 20 - 30 dB below the maximum of the main lobe, the high intensity thermal noise radiation from near objects may have considerable influence on the total noise conditions. Radiating objects in the surroundings may be at different distances, and the influence on the measured noise pattern may therefore be different in different directions.

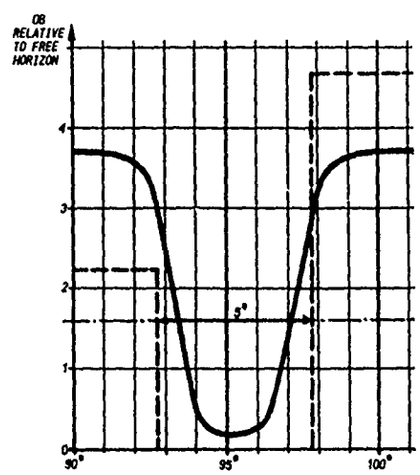


Figure 10. Influence of buildings on the thermal noise in reflector antennas.

The thermal noise radiation from buildings may also reduce the signal to noise ratio. As indicated in Figure 10, we had about 0.2 dB more noise halfway between two buildings 5° apart than at free horizon. The antenna used was a 1.5 m parabolic reflector antenna, and the elevation angle was about 10°. The noise factor of the amplifier/converter was about 1.3 dB.

0.2 dB is a very low value and seems to be of no importance. As mentioned before, that is true when we have high signal strengths. But when we are working at low signal strengths, very low additional noise may have some influence, especially when we have signal to noise ratios near the limit between usable and not usable. An when we have a very low noise factor in the first amplifier, this effect will be even more pronounced.

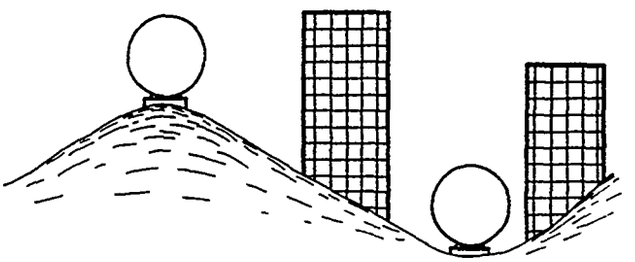


Figure 11. Parabolic reflector antennas at different sites.

Measurements at two different sites, as indicated in Figure 11, gave higher noise levels in the valley than on the hill, dependent on elevation angle of the antenna. Therefore it may be important to have rather high antenna sites when the elevation angle is low. And when there are trees in the surroundings, we have to be aware of the fact that trees also are thermal noise radiators.

MEASUREMENT OF THERMAL NOISE CAUSED BY A BUILDING.

In order to get information about the influence of the surroundings on the thermal noise pattern of an antenna, the thermal noise caused by a building was measured at an elevation angle of 25°. The elevation angle to the top of the building was about 45°. The building and the noise variations referred to the maximum noise level, are shown in Figure 12.

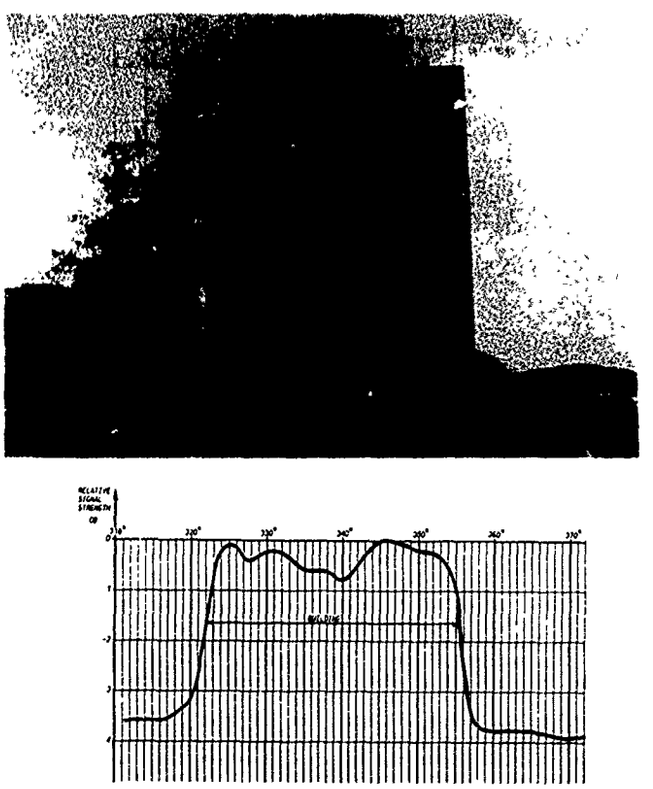


Figure 12. Thermal noise caused by a building.

We see here that the influence of the building is about 2 - 3° to each side of the building, and we reach the maximum noise values when the antenna is directed about 2° into the building.

We then tried to find a narrow object which could radiate adequate thermal energy to give acceptable distance between the normal noise in the surroundings and the thermal noise caused by the object. In Figure 13 is shown a chimney at an industrial plant. The distance from the antenna to the chimney was such that the horizontal width was less than 2°. The elevation angle of the antenna was 30°, and the elevation angle to the top of the chimney was about 50°.

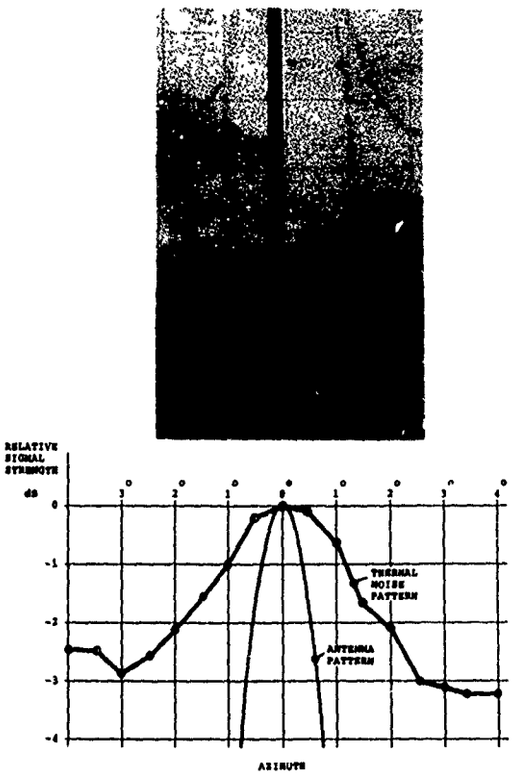


Figure 13. Measurement of the thermal noise pattern for a 1.5 m parabolic reflector antenna at 12 GHz using a chimney as radiating source.

The relative noise pattern for the 1.5 m antenna is also shown in Figure 13. If there had been no influence from the surroundings, the pattern should be almost the same as the normal electric field strength pattern where the sidelobes are more than 25 dB below the maximum gain.

INFLUENCE OF THE SUN COMPARED WITH RADIATION FROM A BUILDING.

In order to get an impression of the influence of the radiation from the sun, the 1.5 m antenna was rotated around the horizon at a constant elevation angle (about 20°) at a relatively open site. The sun was then at about 35° elevation angle. Almost no change in the noise level was observed when the antenna was rotated 360°.

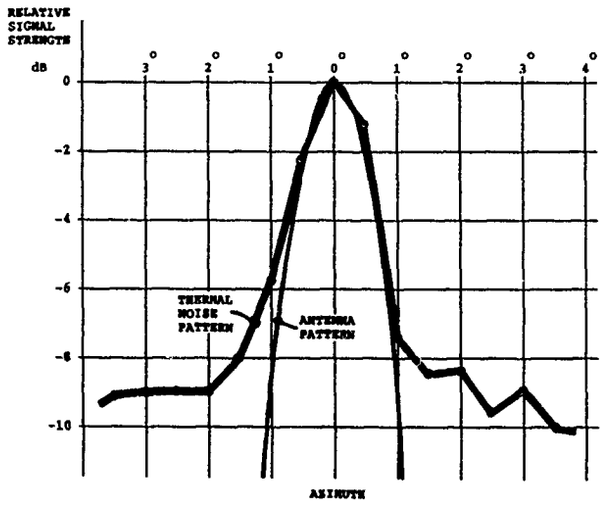


Figure 14. Measurement of the thermal noise pattern for a 1.5 m parabolic reflector antenna at 12 GHz using the sun as radiating source.

The next measurement was done when rotating the antenna main lobe through the middle of the sun. The elevation angle was then about 23°. In Figure 14 is shown the relative pattern for the thermal noise. The increase in thermal noise relative to the thermal noise at clear sky (and the surroundings) at 23° elevation angle was 10.1 db. If we compare with the maximum thermal noise radiation from the chimney in Figure 13 and from the building in Figure 12, the difference is 6 - 6.5 dB.

A measurement at a very low elevation angle (about 6°) half a kilometre from a 60 m high building, gave 2.3 db increase in the received noise. These results are rather surprising because of the relatively high thermal noise radiation from buildings and other objects.

If we use Equation 7, we may find an expression for the thermal noise registered for instance on a spectrum analyser. We will always have a certain loss between the antenna and the first amplifier, and this loss was in the order of 0.25 dB. Referring to Figure 3, we assume that the amplification of the pre-amplifier/converter is of such a magnitude (often more than 50 dB) that we may ignore the noise factor of the spectrum analyser.

The noise factor of the first amplifier was 1.3 dB, that is, $F_{s1} = 1.35$. We have:

$$F_t = \frac{T_a}{290} + F_{s1} \cdot 10^{\alpha_c/10} = \frac{T_a}{290} + 1.35 \cdot 10^{0.25/10}$$

For measurement of the noise caused by the sun, the elevation angle was 23°. From Figure 9 we see that the antenna noise temperature for clear sky is 25°K. The total noise factor is then:

$$F_t = \frac{25}{290} + 1.35 \cdot 10^{0.25/10} = 1.52 \ (\approx 1.8 \text{ dB})$$

which corresponds to a noise temperature of $T = (F - 1)T_o = 151°K$.

From $P = kTB$ we see that the noise power is proportional to the noise temperature. If the power is increased by 10.1 dB, the noise temperature of the sun measured with a 1.5 m parabolic antenna should be 1545°K.

If we consider the radiation from the building in Figure 12, the elevation angle was 23°. The noise was increased by 3.8 dB, which means that the noise temperature of this building was 362°K. The real noise temperature of such a building is higher than the absolute temperature for such a building, which normally is 290 - 295°K.

Even if we look at the measurements for a 60 m high building at half a kilometre distance, we also have a higher noise temperature than the normal temperature of a building. From Figure 9 we have that the antenna noise temperature at 6° elevation is about 60°K, and we have:

$$F_t = \frac{60}{290} + 1.35 \cdot 10^{\frac{0.25}{10}} = 1.64$$

which corresponds to a noise temperature of 185°K. 2.3 dB increase gives a noise temperature of 314°K for the building half a kilometre away.

It is surprising that the radiation from the sun does not give more than 6 - 7 dB higher noise level than the noise radiation from a building about 100 m away. However, the measurements were done in the autumn and at a low elevation angle (23°).

Relatively few measurements were done concerning thermal noise radiation from different object. The results in this article are only indications of what may be expected under the conditions described.

The measurements described in this article were done in the 12 GHz band. If we shall use higher frequencies, the thermal noise will be even more important. For example for broadband systems in the cable-TV distribution, a frequency band in the 40 GHz range may be used. A 2 GHz band in the 80 GHz range (84 - 86 GHz) has already been allocated to satellite broadcasting. However, at such frequencies the noise factors for receiving equipment will be rather high until there has been some development in these bands.

For instance the noise factors for normal receiving equipment in the 12 GHz range was about 6 dB in 1977. In 1984 a noise factor of 3 dB could be achieved, and now it is possible to have noise factors as low as 0.5 dB. A similar development is expected for the higher frequency bands.

CONCLUSIONS.

It is important to have as low losses as possible between the antenna and the first amplifier in a receiving system. The use of very low noise pre-amplifiers near the antenna is advantageous when we want to receive low signal levels.

The thermal noise radiation from buildings seems to be higher than indicated by the absolute temperature. When working at very low signal levels and at low elevation angles, we have to be aware of the possibility of an increase in the thermal noise because of buildings and other objects in the surroundings or at the horizon. We should also try to have better clearance angle to buildings, etc., than seen from the normal field strength pattern for the antenna. This should especially be taken into account when it is very important to have low noise factors, as for instance in radio link systems and in satellite receiving systems.

REFERENCES:

(1) K. N. Stokke: Fjernsyn, mikrofon, kamera, sender, mottaker. (In Norwegian). Universitetsforlaget, Oslo (1983).

(2) K. N. Stokke: Senderteknikk. (In Norwegian). Vett & Viten, Sandvika (1989).

DISCUSSION

George H. HAGN

When considering the clearance of a building or an obstacle one normally computes the first Fresnel zone to check the clearance when aligning the antenna. Now it seems that we may have a different alignment rule required when siting a receiving antenna for low noise operation.
Would you comment ?

AUTHOR'S REPLY

The power gain G (refered to an isotropic antenna) of a parabolic reflector antenna is :

$G \approx 5(\frac{d}{\lambda})^2$

where d is the diameter of the antenna and is the wavelength.
At 12GHz (λ = 2.5 cm) a 1.5 m antenna has a gain of :

$G \approx 5(\frac{150}{2,5})^2 = \underline{18\,000}$ (= 42.6dB)

The angle from maximum gain to the - 3dB point of the main lobe for such an antenna is :

$$\varphi = \frac{1}{2}\sqrt{\frac{27843}{G}} = 0,62°$$

and the angular distance to the first minimum (= the angular distance to the end of the main lobe) is about 3 times this value, that is about 1.8°.
If we look at the radius h_1 of the first Fresnel zone, we have :

$h_1 \approx \sqrt{\lambda r_1}$

where r_1 is the distance from the receiver to the actual shielding object. This equation is an approximation when we are relatively near the receiver (that is, near the receiving antenna).
At for instance 100m distance we have :

$h_1 \approx \sqrt{0.025 \cdot 100} = 1.58$ m

which gives an angle φ_1 to the actual object of :

$$\varphi_1 = \arctan \frac{1.58}{100} = \underline{0.91°}$$

That is, well within the main lobe.
In fact, a noise radiator at that point will have a rather high gain within the main lobe.

When we refer to the strength of the received signal, we may use Fresnel zone considerations. However, when we look at the thermal noise coming into an antenna, it may not be sufficient to keep only the first Fresnel zone free. In fact, my paper shows that especially when we want to receive low signal strengths, it may not even be sufficient to keep the main lobe free because of high intensity thermal radiation from near objects into the sidelobes.

CONTRIBUTION DES GAMMES METRIQUE ET DECAMETRIQUE AU CONCEPT DE RADAR DE VEILLE : PERFORMANCES A SITE BAS ET COMPARAISON AVEC LES RADARS CLASSIQUES

Marc Lesturgie, Office National d'Etudes et de Recherches Aérospatiales, Chatillon sous Bagneux 92322, France

RESUME

Dans le contexte du problème de la détection des cibles évoluant à basse altitude, et des possiblités qu'offrent des systèmes fonctionnant en basses fréquences (gammes HF et VHF), on présente des résultats de modèles de la propagation à site bas sur la terre et sur la mer. Les performances théoriques de ces systèmes nouveaux sont comparées à celles des systèmes classiques fonctionnant en hyperfréquences. Deux exemples (radar de veille sol-air, radar 'marine' d'alerte précoce) illustrent la comparaison.

INTRODUCTION

L'évolution du radar au lendemain de la seconde guerre mondiale a été marquée par la recherche de systèmes de détection alliant compacité, résolution et précision de localisation. L'utilisation de fréquences de plus en plus élevées, les nécessités de traitement de signal rapide ont été le moteur de la recherche de pointe en électronique (développement de nouveaux composants, essor de la microélectronique). Parallèlement au contexte évolutif du radar, on a attaché de plus en plus d'importance aux contre-mesures, et aux techniques de contre-contre-mesures. Néanmoins, le problème de la cible pénétrant à basse altitude, derrière un masque de terrain constitue une stratégie de contre-mesure élémentaire et réaliste à laquelle les radars classiques fonctionnant en hyperfréquences ne savent faire face.

On peut imaginer pour pallier cette contre mesure élémentaire de développer des systèmes de détection fonctionnant dans les bandes de télécommunication (HF-VHF) qui ont la propriété intéressante d'être robustes vis à vis des masques. Il s'agit de nouveaux concepts de radar, dont nous allons établir une comparaison avec les radars de veille classiques, sur le plan de la portée à site bas. Trois exemples de dimensionnement de structures radar sont donnés et complétés par une synthèse comparative de leurs performances.

PHENOMENES DE PROPAGATION

La propagation troposphérique est régie par plusieurs phénomènes dont les principaux sont :

- la réfraction, liée aux effets de gradient d'indice
- la réflexion, liée à la présence du sol
- la diffusion et la diffraction qui traduisent la contribution électromagnétique des obstacles (trajets multiples ou fouillis)

PROPAGATION A SITE BAS

Le calcul du champ électromagnétique rayonné par un doublet situé au dessus d'un sol de caractéristiques électriques données, à courbure sphérique est un problème ancien, abordé par Sommerfeld, Bremmer et Wait.[1]

Le modèle utilisé prend en compte les effets de réfraction troposphèriques (gradient d'indice standard de -39.10^{-6}/kms). Pour le calcul du champ, on est amené à distinguer deux zones physiques , de part et d'autre de l' horizon (figure 1) :

- une zone de visibilité où la propagation est optique
- une zone d'ombre, à grande distance où la propagation résulte des phénomènes de diffraction sphérique et d'onde de surface (on n'envisage pas , ici, la propagation guidée dans les conduits troposphérique ou ionosphérique)

L'horizon radioélectrique s'exprime en fonction des altitudes des points d'émission et de réception :

$$D_{h(km)} = \sqrt{17.h_{e(m)}} + \sqrt{17.h_{r(m)}}$$

En zone de visibilité (1) le champ électrique résulte de la superposition des champs direct et réfléchi par la surface : des interférences apparaissent, et sont d'autant plus prononcées que le module du coefficient de réflexion est voisin de 1. A incidence rasante , en polarisation horizcntale, le coefficient de réflexion est proche de -1, quelque soit le type de sol. Par contre, en

polarisation verticale, il dépend du type de sol et de la fréquence. En particulier aux basses fréquences, sur la mer, son module est inférieur à 1 et sa phase inférieure à 180° de sorte que la perte de champ (interférence destructive) qui précède l'horizon est moins prononcée.

En zone d'ombre (2), au delà de l'horizon, le champ décroit plus vite qu'en espace libre, et la propagation à grande distance, qui fait intervenir des phénomènes d'onde de surface privilégie l'utilisation de la polarisation verticale sur un sol bon conducteur (cas de la mer). Le modèle utilisé calcule la série des potentiels de Hertz, selon la méthode de Bremmer [1,2].

Aux distances voisines de l'horizon, l'influence des différents paramètres n'est pas simple à expliquer, car elle résulte d'un compromis entre la qualité de la propagation optique et la qualité de la propagation dans l'ombre. De plus, dans une partie de cette zone intermédiaire, le champ est calculé par interpolation, car les modèles des zones (1) et (2), à leur frontière, peuvent diverger et ne pas se raccorder.

Les figures 2 à 5 présentent l'évolution des pertes de propagation par rapport à l'espace libre,

- à 10 Mhz, 100 Mhz et 3 Ghz
- sur la mer et sur le sol (terre) de caractéristiques:

(mer) $\varepsilon = 80$ $\sigma = 4$ (sol) $\varepsilon = 10$ $\sigma = 0.01$

- en polarisation horizontale ou verticale,

pour un doublet court devant la longueur d'onde. L'émetteur est situé à 30 mètres du sol. La zone d'interpolation est repérée à l'aide de pointillés, l'horizon par une flèche verticale.

On constate que, quels que soient le type de sol et la polarisation ,la propagation transhorizon à très grande distance (>160kms) classe par ordre d'intéret décroissant les gammes HF, VHF, et hyperfréquences.

La propagation des hyperfréquences en zone optique est bonne , aux interférences près (dont l'influence est, pratiquement, atténuée par les effets de diffusion). A la limite (cas de l'optique) la propagation est parfaite jusqu'à l'horizon, et nulle au delà.

L'utilisation de la gamme HF sur la mer, en polarisation verticale (polarisation requise pour satisfaire au mieux les conditions aux limites à l'interface conducteur) est prometteuse en deçà et au delà de l'horizon.

En gamme VHF la propagation avant l'horizon est moins bonne qu'en bande S, meilleure au delà. Aux distances moyennes (<150 kms) la propagation sur la terre, est meilleure en VHF qu'en HF

Quelques applications, dans le domaine des communications corroborent ces propriétés :

- liaisons "à vue", faisceaux hertziens
- communications marines : MF-HF
- communications sol / sol : VHF

Ce dernier exemple résulte également d'un choix faisant intervenir la robustesse d'une liaison vis à vis des masques de terrain ou obstacles macroscopiques, dont nous allons maintenant évoquer l'influence.

DIFFRACTION PAR DES MASQUES

Lorsqu'un obstacle intercepte l'ellipsoide de Fresnel d'une liaison dégagée entre deux points, la propagation se trouve atténuée (figure 6) .

La modélisation du champ électrique, dans l'ombre d'un obstacle tel qu'une colline ou une vallée encaissée est complexe . Il est possible d'obtenir des ordres de grandeur du champ rayonné en considérant le cas du dièdre semi-infini, dont la solution rigoureuse est due à Sommerfeld. Il existe également une solution approchée, qui est la solution de l'optique, indépendante de la polarisation, et qui est relativement proche des résultats de mesure de champ effectuées sur la terre.[3]

Le champ reçu par un récepteur est donné en amplitude et en phase par l'expression simple

$$\frac{E}{Eo} = \int_v^{\infty} e^{j\pi t^2/2}.dt \qquad \text{où} \qquad v = h\sqrt{\frac{2}{\lambda}\left(\frac{1}{d1} + \frac{1}{d2}\right)}$$

avec :

Eo = champ d'espace libre
λ= longueur d'onde
h = altitude de l'obstacle
d1= distance émetteur-obstacle
d2= distance récepteur-obstacle

Les figures 7,8,9 présentent la répartition du champ, rapporté au champ d'espace libre , à 10,100 Mhz et 3 Ghz. Les basses fréquences sont avantagées. Aux très hautes fréquences, le champ dans l'ombre de l'arête est négligeable.

APPLICATIONS RADAR

Les performances en détection des systèmes radar sont , dans le cas général, complexes à évaluer : outre la qualité de la propagation elles font intervenir le niveau de S.E.R. (section efficace radar) des cibles, le facteur de bruit, le type de traitement ....

En l'absence de brouilleur les performances en détection des systèmes radar sont données par l'expression du rapport signal à bruit énergétique :

$$\frac{S}{B} = \frac{P_m.T.G_e.G_r.\lambda^2.\Sigma.L^2}{(4.\Pi)^3.D^4.F.KTo} \quad \text{avec}$$

P_m = puissance moyenne émise
T = durée d'intégration
λ = longueur d'onde
Ge,Gr= gain des antennes
D = distance radar-cible
F = facteur de bruit radar
KTo = -204 dBJoules
Σ = S.E.R. de la cible
L = pertes de propagation

DIAGRAMME DE COUVERTURE

Dans le but d'isoler, dans un premier temps, l'impact des effets de propagation a site bas sur les systèmes de veille, nous considérons une classe de radars possèdant une portée en espace libre de 400 kms.

Pour une altitude h donnée, on recherche la distance D telle que

$$[D/Do]^2 = L[D,h,f,\varepsilon,\sigma]$$

avec
L = pertes de propagation
h = altitude de la cible
Do = portée espace libre
f = fréquence d'émission
ε = permittivité du sol
σ = conductivité du sol

DETECTION AU DELA DE L'HORIZON

Les figures 10 à 13 présentent les diagrammes de couverture pour trois fréquences, sur terre et sur mer, en polarisation horizontale et verticale. La courbe en pointillé représente la limite de visibilité (horizon)

En polarisation verticale, l'explication physique des phénomènes qui conduisent à ces diagrammes de couverture est simple. Plus la fréquence d'émission est basse, plus importants sont le couplage et le transfert d'énergie avec le sol. Si de plus, l'impédance de surface est faible, alors cette énergie communiquée au sol sera propagée (par onde de surface) avec une dissipation (pertes) faible.

Sur mer (figure 10) ces deux propriétés sont verifiées à 10 Mhz et conduisent à une détection au delà de l'horizon envisageable jusqu'à 160 kms. Par contre en VHF (100 Mhz), le couplage subsiste mais l'impédance de surface augmentant, la dissipation dans le sol devient importante et les propriétés de détection à site bas qui en découlent sont mauvaises. Lorsque la fréquence augmente encore (3 Ghz par exemple) le couplage avec la surface disparait et la propagation en deçà de l'horizon s'améliore : la portée radar devient alors voisine de la portée optique.

Sur le sol (terre, figure 11) de faible conductivité les ondes de surface -dont la dissipation est excessive - sont inexploitables, aux trois fréquences considérées ; tout couplage avec la la surface s'avère néfaste . Notamment en HF la portée radar est très inférieure à la portée optique. Le couplage disparaît dès que l'on utilise la gamme VHF : à 100 Mhz on rejoint (et dépasse même très légèrement - par le jeu de coefficient de réflexion en phase avec le trajet direct) la courbe de couverture optique (horizon). Le comportement des radars opérant en hyperfréquences est proche de l'optique : détection assurée théoriquement jusqu'à l'horizon, si la fréquence tend vers l'infini.

Enfin en polarisation verticale, on peut observer des fluctuations en fonction de la fréquence, de la portée autour de l'horizon, fluctuations qui apparaissent par le biais des coefficients de réflexion. Les schémas de la figure 14 résument la philosophie de l'évolution de la détectabilité à site bas en fonction de la gamme de fréquence retenue.

En polarisation horizontale (figures 12 et 13) aucun couplage bénéfique n'est à attendre du sol ou de la mer. La couverture limite est la couverture optique, atteinte théoriquement lorsque la fréquence tend vers l'infini.

DETECTION DERRIERE UN MASQUE

La figure 15 présente la couverture radar, dans l'ombre d'une arête de 400 mètres de haut, située en vue du radar, à 80 kms. Les fréquences basses (gammes HF et VHF) sont privilégiées.

DIMENSIONNEMENT DES RADARS

Jusqu'à présent nous avons décrit l'influence de la propagation sur les portées des radars à l'égard de cibles évoluant à basse altitude, en fonction de la gamme de fréquence de fonctionnement, pour une portée 'en espace libre' donnée.

En fait, l'analyse des performances des systèmes en fonction de leur fréquence d'émission, pour être objective, doit inclure également la variation d'autres paramètres , contenues dans l'expression du rapport signal à bruit énergétique :

$$\frac{S}{B} = \frac{P_m.T.G_e.G_r.\lambda^2.\Sigma.L^2}{(4.\Pi)^3.D^4.F.KTo}$$

Outre le facteur en λ^2 qui favorise l'utilisation des basses fréquences, l'expression fait apparaître différents termes dont la nature et les valeurs diffèrent selon la gamme de fréquence utilisée:

- *la durée d'intégration*, lorsqu'on effectue un traitement doppler cohérent la durée maximale d'intégration cohérente varie en $\sqrt{f}$; en effet le résidu de phase lié au mouvement de la cible et non compensé par le traitement doppler s'écrit à l'ordre 2 :

$$\delta\phi = 2.\pi.f.\beta(V,\Gamma)\, t^2$$ (β fonction scalaire de V vitesse de la cible, Γ accélération)

Ainsi, typiquement, la durée limite d'ntégration cohérente sera de :

- quelques dizaines de milisecondes en bandes L,S,C
- quelques centaines de milisecondes en VHF
- de l'ordre de la seconde en HF

Par ailleurs les radars fonctionnant en gammes HF ou VHF ne possèdent pas d'ambiguité doppler-distance (les dopplers maximum observables étant toujours inférieurs à la fréquence de répétition des impulsions)

- *les gains des aériens* sont liés aux surfaces d'antennes physiques; aux grandes longueurs d'onde on utilise des réseaux phasés plus encombrants que les antennes hyperfréquences, mais, en revanche, moins vulnérables (la destruction partielle du réseau n'entrainant pas la supression totale des performances).

- *le facteur de bruit* du radar F dépend de la qualité des chaînes de réception et du niveau du bruit atmosphérique ambiant. Dans les gammes radar classique le facteur de bruit (de l'ordre de 5 à 10 dB) est lié au deux phénomènes. Aux basses fréquences il est donné par le facteur de bruit atmosphérique, qui domine : de l'ordre de 25 à 30 dB en gamme HF.

- *les pertes de propagation* dépendent, comme étudié précédemment, de la fréquence et de la configuration choisies (polarisation, type de sol)

- *la section efficace radar* (S.E.R.) est un paramètre qui dépend de la géomètrie et de l'attitude de la cible, et de la fréquence d'émission. On distingue en général trois zones de variation de la S.E.R en fonction des dimensions relatives de la cible par rapport à la longueur d'onde (rapport a/λ). En zone de Rayleigh ($a/\lambda << 1$, typiquement les 'petits' avions en gamme HF) les niveaux de SER sont faibles, mais augmentent avec la fréquence En zone de résonance ($1< a/\lambda <10$, c'est le cas des avions en gamme VHF) les niveaux sont généralement plus forts, mais peuvent varier énormément en fonction de la fréquence (de l'ordre de 30 dB) En zone optique ($a/\lambda >> 1$, cas des cibles en hyperfréquences) les niveaux sont de plus en plus imprévisibles, du fait de l'utilisation croissante de matériaux absorbants tendant à faire chuter le S.E.R.

EXEMPLES

On décrit deux exemples d'avant projets de radars fonctionnant en basse fréquence , et que l'on compare à un radar classique fonctionnant en bande S :

- Radar 'marine' d'alerte précoce

- détection avion à 80 kms (vue de face), altitude 30m
- détection missile à 50 kms (vue de face), altitude 30m
- fréquence = 10 Mhz

Le choix de la fréquence résulte de la modélisation du rapport signal à bruit (figure 16) qui présente un maximum , du fait des effets de variation inverse de la SER et du facteur de bruit d'une part, de la qualité de la propagation d'autre part.

- Radar de veille sol-air

- détection jusqu'à 400 kms, en vue dégagée
- fréquence = 100 Mhz

Le tableau suivant résume les caractéristiques des trois radars.

- performances en détection :

La puissance à émettre en VHF est plus faible qu'en bande S, pour la même portée en espace libre (400 kms). Le radar VHF est robuste au voisinage de l'horizon et à l'égard des cibles masquées.

Le radar 'marine' fonctionnant en HF à de bonnes propriétés de détection au delà de l'horizon. On peut envisager son utilisation comme système d'alerte précoce, implanté à bord d'un navire .

- performances en localisation :

En gammes HF et VHF la directivité en gisement s'obtient au prix d'un déploiement horizontal très important: plusieurs centaines de mètres en VHF, plusieurs kilomètres en HF. Le nombre d'antennes utilisée croit alors considérablement, à moins de sous échantillonner la figure géométrique de réseau : le réseau devient alors lacunaire. La directivité en site risque également d'être médiocre si le déploiement vertical des antennes du réseau est

RADAR	MARINE	VHF	CLASSIQUE
Bande	HF	VHF	S
Fréquence	10 Mhz	100 Mhz	3 Ghz
Polarisation	V	V	H ou V
S.E.R. avion	0.2 m^2	10 m^2	1 m^2
Fact. de bruit	30 dB	10 dB	8 dB
Durée traitt	1 s	0.1 s	5-10 ms
T/R	200µs/3ms	100µs/3ms	100µs/3ms
GAIN E/R	13 dB	18 dB	40 dB
Puiss. crète	6 kw	50 kw	500kw
Puiss. moy.	400 w	1.6 kw	17 kw
PORTEE (à vue)	100 kms	400 kms	400 kms
PORTEE (30m/sol) (s/b)	80 kms (13dB)	75 kms (13dB)	65 kms (21dB)
DIMENSIONS.... antenne	Réseaux d'antennes		
	10 émetteurs 10 récepteurs 150 m	32 émetteurs 32 récepteurs 234 m	2.3m 3.5m
RESOLUTIONS			
gisement	8°	2.5°	1.16°
distance	qq 10 kms	qq 100 m	qq 10 m

faible. Le R.I.A.S. [4] - *Radar à Impulsion et Antenne Synthétiques* -, fonctionnant en VHF possède ainsi des pouvoirs de résolutions comparables à ceux des radars de veille classiques, grâce à sa lacunarité et au dénivelé important des antennes du réseau.

En gamme HF dans l'application 'marine' (où l'on s'est fixé un déploiement de 150 m, les capacités de localisation sont mauvaises, mais on peut envisager son utilisation comme système d'alerte à bord d'un navire. En revanche, un plus grand déploiement (possible sur le littoral par exemple) permettrait d'effectuer une surveillance 'marine' performante du point de vue de la localisation.

Quant à la résolution en distance, liée à la bande du signal radar, elle est à priori limitée en HF du fait de l'encombrement spectral (raies parasites et interférences). On peut néanmoins envisager un mode de fonctionnement à plusieurs fréquences, émises simultanément, aux emplacements clairs du spectre.

- <u>interférences et brouillage</u> :

Les radars opérant aux grandes longueurs d'onde ont des pouvoirs de résolution angulaires moyens , et perçoivent donc des niveaux de fouillis importants, notamment à site bas. Le fouillis de mer (application marine) et le fouillis de sol (détection derrière un masque, ou à site

bas) dépassent en général la cible d'au moins 70 décibels. L'intégration doppler est indispensable pour isoler la cible, comme c'est le cas pour le R.I.A.S , notamment pour ce qui concerne la détection des navires [5].

L'élimination des brouilleurs intentionnels, dans un contexte de contre mesures réalistes appliquées aux radars 'basses fréquence' fait l'objet de traitements adaptatifs sur le réseau de réception . Sachant que le nombre de brouilleurs éliminables est une fonction croissante du nombre de capteurs du réseau, les systèmes 'basses fréquences' constitués d'antennes réparties semblent avantagés par rapport aux systèmes radar hyperfréquences qui disposent rarement de plus de trois antennes. En revanche les brouilleurs 'basses fréquences' sont plus faciles à constituer.

CONCLUSION

Des systèmes radar futurs fonctionnant en gamme HF ou VHF semblent prometteurs vis à vis des performances en détection à site bas. De plus, ils ne sont pas sensibles aux absorbants utilisés en hyperfréquences pour diminuer les niveaux de section efficace radar. Par contre les performances de localisation , pour égaler celles des radars classiques requièrent un déploiement au sol important. Certains traitements sont indispensables : traitement doppler cohérent, traitement d'antenne (réseau). D'autres traitements , spécifiques aux antennes réseaux, ouvrent la porte à un grand nombre d'applications, limitées dans le cas des radars classiques utilisant une antenne de configuration figée.

REMERCIEMENTS

La Direction des Recherches, Etudes et Techniques a financé les travaux dont cette présentation est issue.

BIBLIOGRAPHIE

[1] H. Bremmer - Terrestrial radio waves - Elsevier Publ. Co.,1949
[2] P. Golé - Etude de la propagation des ondes électromagnétiques VHF - La Recherche Aérospatiale n°1, p13-27, jan-fév. 1983
[3] L. Boithias - Propagation des ondes radioélectriques dans l'environnement terrestre, Dunod, 1984
[4] J. Dorey, G. Garnier - Radar à Impulsion et Antenne Synthétiques, colloque international sur le radar , Paris 24-28 avril 1989
[5] M. Lesturgie, D. Thibault - Fouillis de mer et limitations introduites sur la détectabilité des cibles en VHF, colloque international sur le radar , Paris 24-28 avril 1989

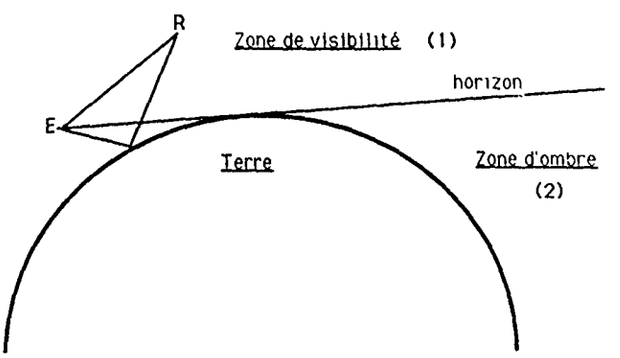


Figure 1 : Propagation à site bas

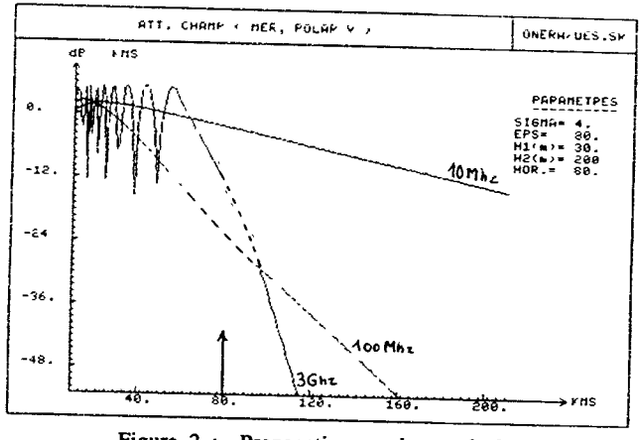


Figure 2 : Propagation au dessus de la mer [polar V]

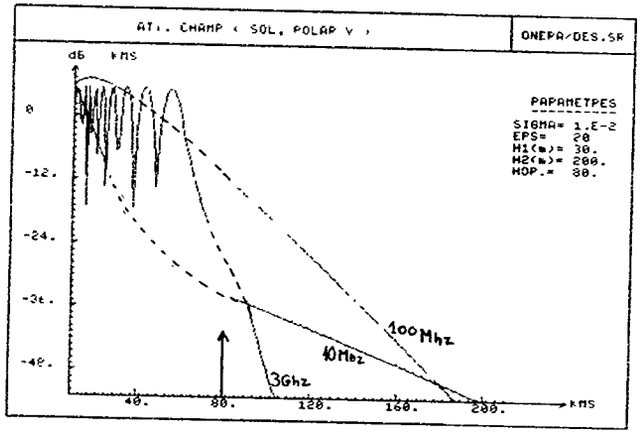


Figure 3 : Propagation au dessus de la terre [polar V]

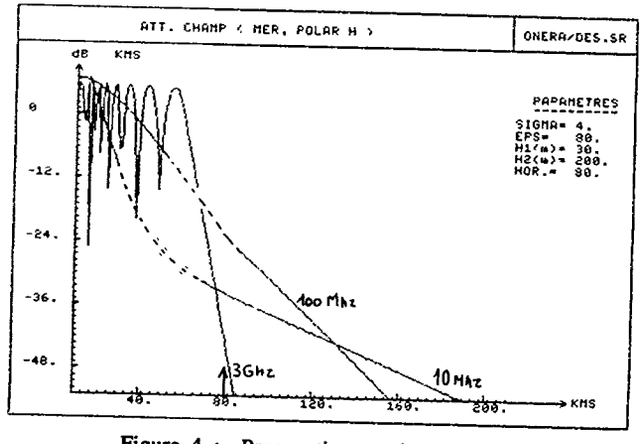


Figure 4 : Propagation au dessus de la mer [polar H]

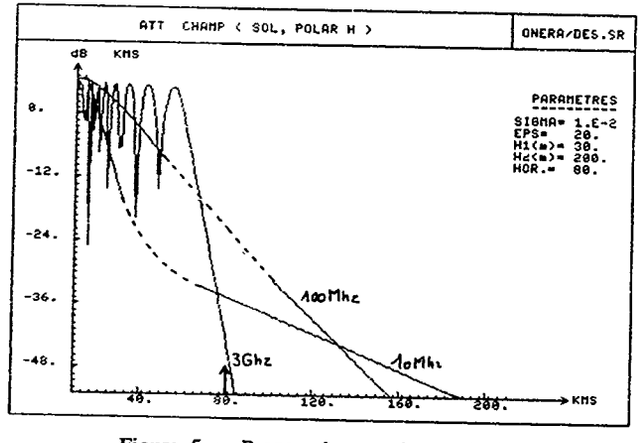


Figure 5 : Propagation au dessus de la terre [polar H]

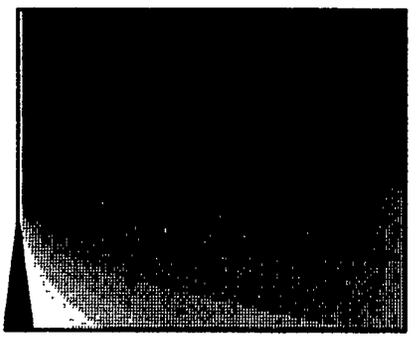


Figure 7 : Fréquence 10 Mhz

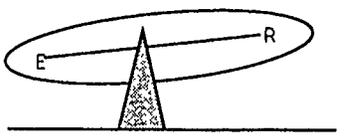


Figure 6 : Propagation derrière un masque

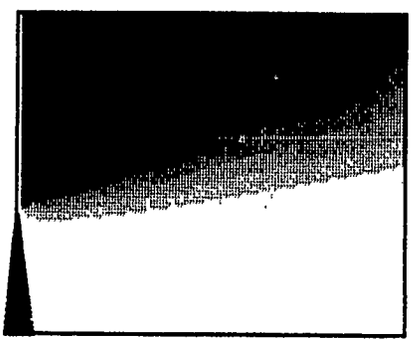


Figure 8 : Fréquence 100 Mhz

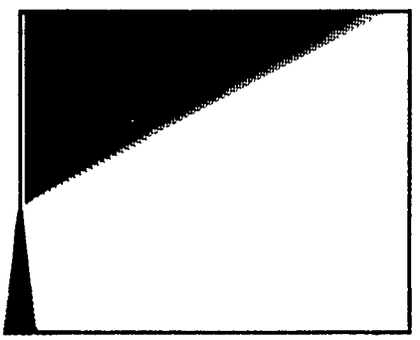


Figure 9 : Fréquence 3 Ghz

Atténuation du champ électrique derrière une arête

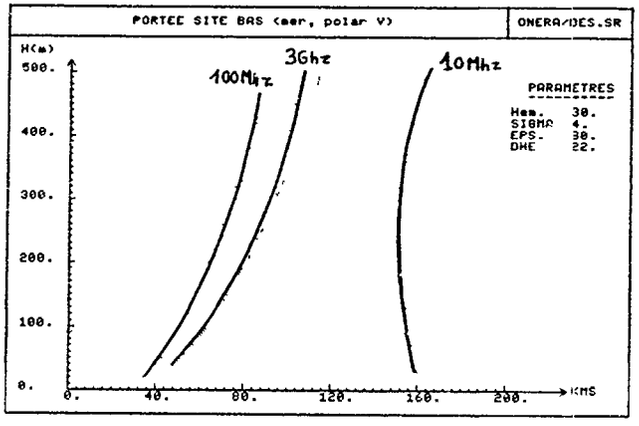


Figure 10 : Couverture à site bas [mer, polar V]

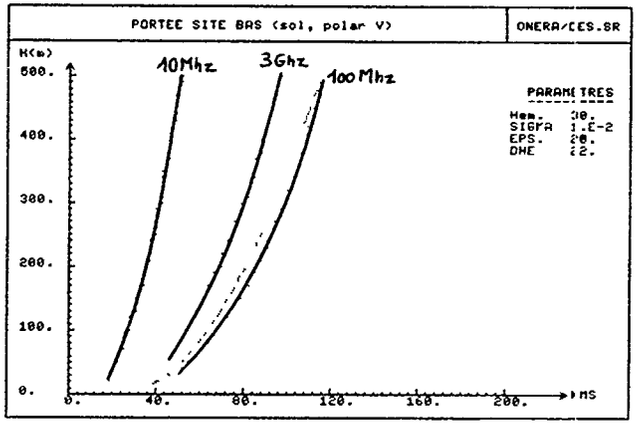


Figure 11 : Couverture à site bas [terre, polar V]

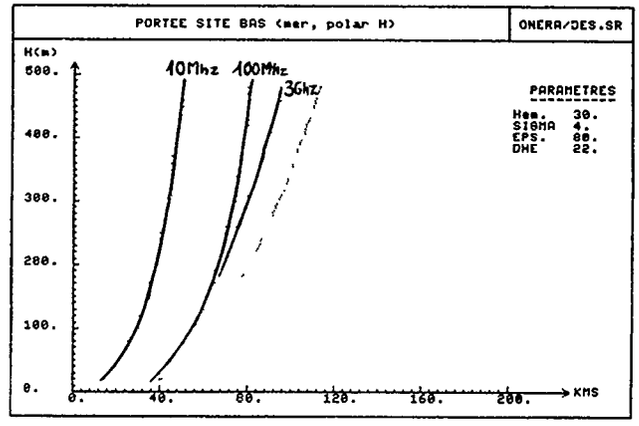


Figure 12 : Couverture à site bas [mer, polar H]

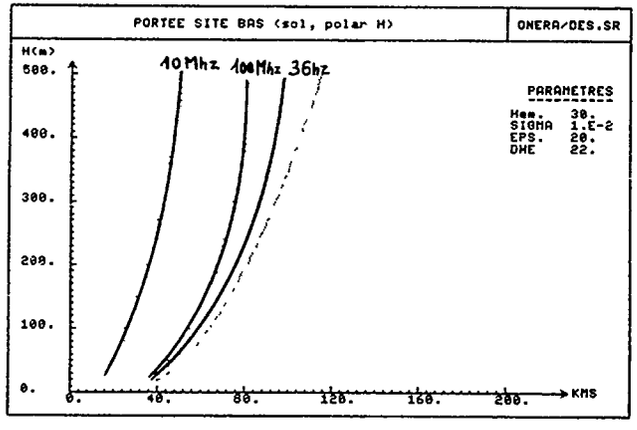


Figure 13 : Couverture à site bas [terre, polar H]

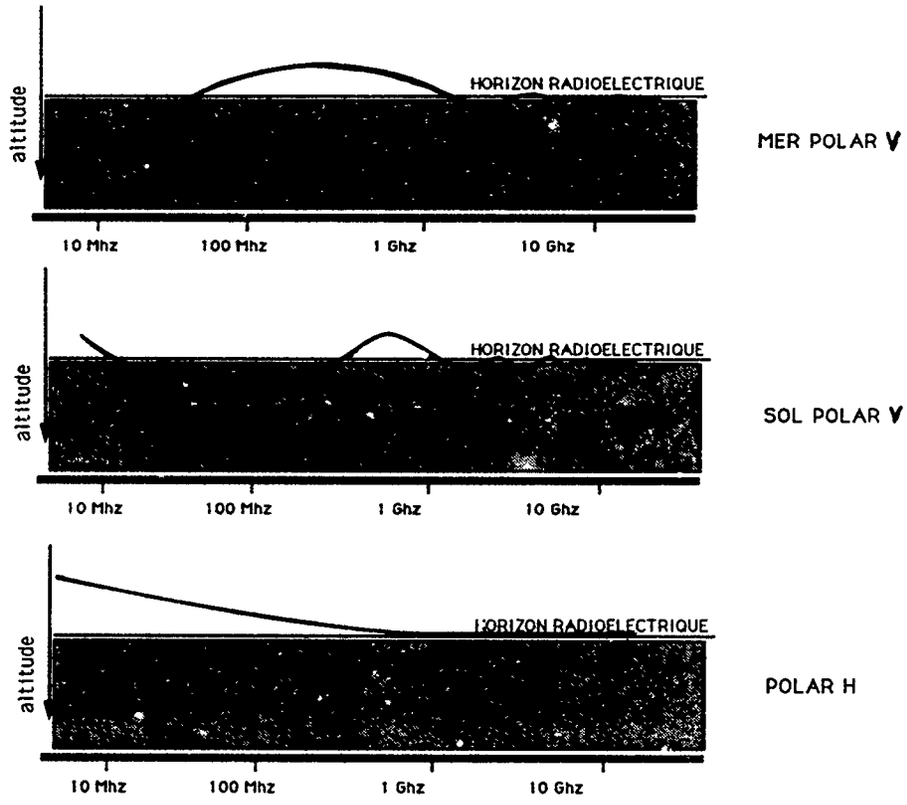


Figure 14 : Evolution de la détectabilité à basse altitude en fonction de la gamme de fréquences (à distance fixée)

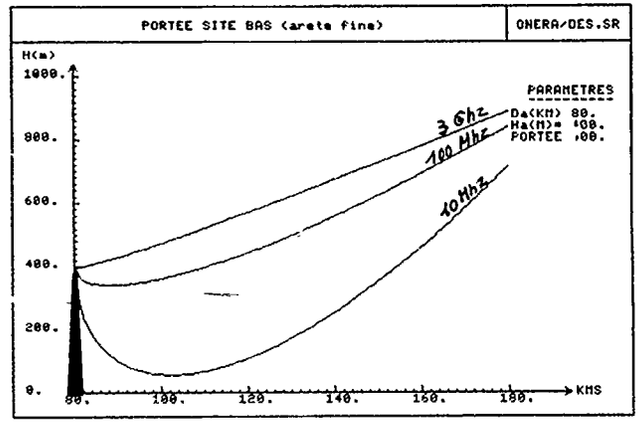


Figure 15 : Couverture derrière un masque de terrain

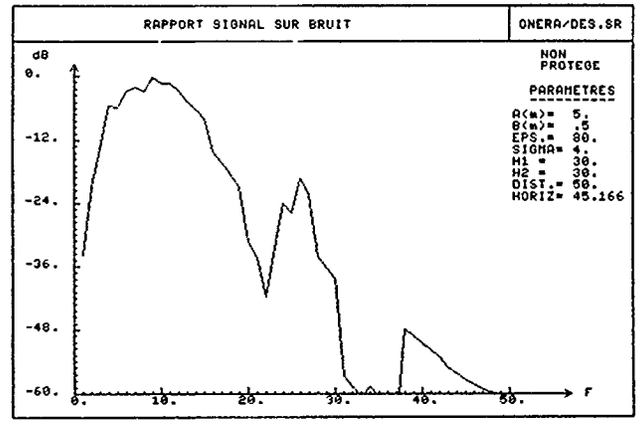


Figure 16 : Evolution du rapport signal à bruit à site bas

DISCUSSION

E. SCHWEICHER, BE

La portée de 50 km du radar marine d'alerte précoce est-elle valable pour les missiles à vol rasant (sea skimmers en anglais) ?

AUTHOR'S REPLY

La portée est valable pour les missiles évoluant à 10 ou 30 m de la surface. Les résultats sur la détection de missile sont issus d'une modélisation simplifiée de sa section efficace radar.

John S. BELROSE

You have presented classical groundwave propagation curves, comparing 10MHz, 100MHz and 3GHz for a fixed ground conductivity and dielectric constant. These parameters for land of a particular type depend upon frequency. For a detailed comparison of propagation over a particular terrain I suggest that you should have used the appropriate effective conductivity and dielectric constant parameters. This would give you different results. Have you given this matter consideration ?

AUTHOR'S REPLY

La constante diélectrique et la conductivité prises en compte sont des valeurs moyennes ; il est peu probable que la variation de ces paramètres pour un sol donné soit significative, de 10MHz à 100MHz ; par contre, la variation de (ε , σ) en fonction du type de terrain (sol sec ou rocailleux) peut être importante, bien qu'en pratique (ayant testé différentes valeurs de (ε , μ) pour le sol) on ne constate pas de changement de comportement des courbes de propagation. Sur la mer, par contre, les effets de diffusion/diffraction par les vagues dépendent fortement de la fréquence et requièrent de considérer des valeurs de conductivité équivalente variables (de 4 à 1 en fonction de la fréquence et du vent).

PROPAGATION STUDIES OF A 60GHz COMMUNICATION SYSTEM

by

Dipl.-Ing. H.-J.Ostertag
Standard Elektrik Lorenz AG
Mannheimstrasse 226
D-7530 Pforzheim
Germany

CONTENTS

1. Introduction

According to the importance of the VHF Radio System the enemy has concentrated enormous activities to reduce and degrade the system performance. This electronic threat results in a number of Electronic Counter Measures (ECM), mainly in Interception and Jamming. Automatically switched on "Set-up Receivers" allow total surveillance of the VHF-band. Besides passive surveillance, active Counter Measures can follow.

New requirements were defined according to the future electronic threat:

* Security against
  - Interception and Detection
  - Deception
  - Jamming
* Improvement and Upgrading of Communication Capacity
* Secure data transmission for Automatic Command, Control and Weapon Systems.

These tactical requirements resulted in investigation of new principles and new technical requirements for the equipement:
- digital transmission
- digital voice modulation
- secure coding methods
- Frequency economic methods
- Spread spectrum application
- encryption

Realization of such systems is feasible. Limits are set because of multipath effects and communication capacity in the VHF-range.

A new possibility for the future is the selection of new frequency ranges. Regarding possible free ranges, one comes up to ranges higher than 10GHz.

Atmospharic attenuation is a main factor influencing propagation. Resonator effects lead to high value of attenuation. There are some Maxima and Minima, where the ranges of 35GHz, 60GHz and 95GHz are of special interest.

In the case of 60GHz, atmospharic attenuation comes up to 15dB/km. The additional attenuation essentially reduce transmission range.

Assuming system values for a transmission link

Gt = Gr = 7dB Antenna Gain
Pt = 1W Transmitted Power
Pr = -90dBm Receiver Sensitivity

the range will be only 680m. Including rainfall assuming a rate of 10mm/h, the range is reduced to 520m.

For the evalution of ECM resistance of a radio system, it is necessary to consider surveillance and jamming conditions. Antenna gain of surveillance receiving equipement is assumed to 30dB. So 60GHz can be detected over a distance of about 2000m. The atmospharic attenuation of 15dB/km at 60GHz is an excellent means of achieving ECM-resistance.

The 60GHz frequency band offer broadband capacity and high resistance against Interception and Jamming. For some short radio nets 60GHz is a practical solution.

2. Measuring Equipement

Transmitted data, a pseudo random noise sequence with length of 2^{10}-1 bits, are given from a word generator. Bit rate can be selected from 0,125Mbit/s to 8Mbit/s. Switch enables to select Frequency Modulation (FM) or Phase Modulation (PM). In case of FM the Voltage Controlled Oscillator (VCO) at 150MHz is tuned by the digital signal, producing 2-FSK modulation. For PM phase of 150MHz signal is switched $\pm 90^{0}$. Modulating signal is differentially coded to produce 2-DPSK modulation. Mixing the 150MHz band and signal of a 20MHz quartz oscillator generates 2 frequency bands at 130MHz and 170MHz. One band or both of these bands can be selected for transmission, thus enabling frequency diversity.

The if at 130/170MHz is converted to 3GHz using a 2,85GHz phase locked oscillator (PLO). 3GHz band is mixed up to 61GHz. A 58GHz PLO is used as local oscillator. 61GHz signal is amplified by a 3-stage injection locked amplifier with 22dB gain. Output power at antenna plane is about 100mW.

Receiver converts 61GHz band down first to 3GHz and second to 130/170MHz using PLOs at frequencies 58GHz and 2,85GHz. The if bands at 130MHz and 170MHz are filtered, amplified and limited. A number of filters enables selection of narrow bandwidth in case of bitrates < 8Mbit/s. After demodulation, baseband signals can be selected and in case of FSK also combined. Analog baseband signal is converted to digital level and in case of DPSK differentially decoded. At video outputs of limiting if amplifiers if-power can be measured.

3. Measuring Conditions\Environment

Received signal power and bit errors are measured. In time slots of 10ms bit errors are counted and if video signals are A/D-converted. Results are transferred to a personal computer (PC). Stored data can be computed by the PC giving signal to noise ratio (S/N) and bit error rate (BER). Data can be accessed to statistical software.

Measurement equipement consist of transmitter and receiver located in two vehicles. Measurements result stationary or with slowly moved receiver in different environments. Roads with vegetation on both sides and forest roads are main selected. Distance between transmitter and receiver is varied to about 400m.

4. Measurements with Moved Receiver and Fixed Transmitter

In the first test configuration both vehicles were on a street without special obstacles (line of sight). Driving one vehicle away from the other, variations of field strength were measured. Variations come up to nearly 30dB (fig. 1). The distance between adjacent minima and maxima become larger with larger distance.

The signal to noise ratio and the bit error rate as function of distance are shown in fig.2. Two lines of S/N apply to if 130MHz and if 170MHz. Bit error rate is increasing at the minima of S/N.

Fig.3 and fig.4 show the results of measurements with FSK and DPSK at a bit rate of 8Mbit/s. For comparison, theoretical values are outlined with a parameter c. The parameter c is defined as the relation of direct received signal energy to the mean square value of multipath signals, according to models of RICE-fading.

This gives c = 0 no direct signal
c = ∞ no multipath signal

The measured BER (steady movement of one vehicle) shows following results:

- at high value of S/N, mean BER is in good agreement with RICE-fadig.
- at low value of S/N, BER exceed the values of RALEIGH-distribution. Additional errors may be caused by intersymbol interference which is not included in the models.
- DPSK results in better BER than FSK.

On grass direct received signal is main part of received signals. According to influence of grass only diffuse reflection is added. The results on grass are shown in fig.5 and 6:

- S/N values are more concentrated
- BER are much better and in good accordance to theory.

5. Measurements with Stationary Receiver and Transmitter

Varations of received signals on roads with vegetation on both sides are measured. Fig.7 shows a partition of this measurement.

In fig.8 results of amplitude variations are drawn to a Normal Probability Plot. Normal probability results to a straight line in this plot. Distribution of amplitudes are in good agreement to a Normal Probability.

6. Summary

This paper described measurement equipement for 60GHz propagation measurements. Experimental results have been presented.

Results of propagation measurements were helpful designing a 60GHz radio system. The experimental program was sponsored by German Military Authorithies.

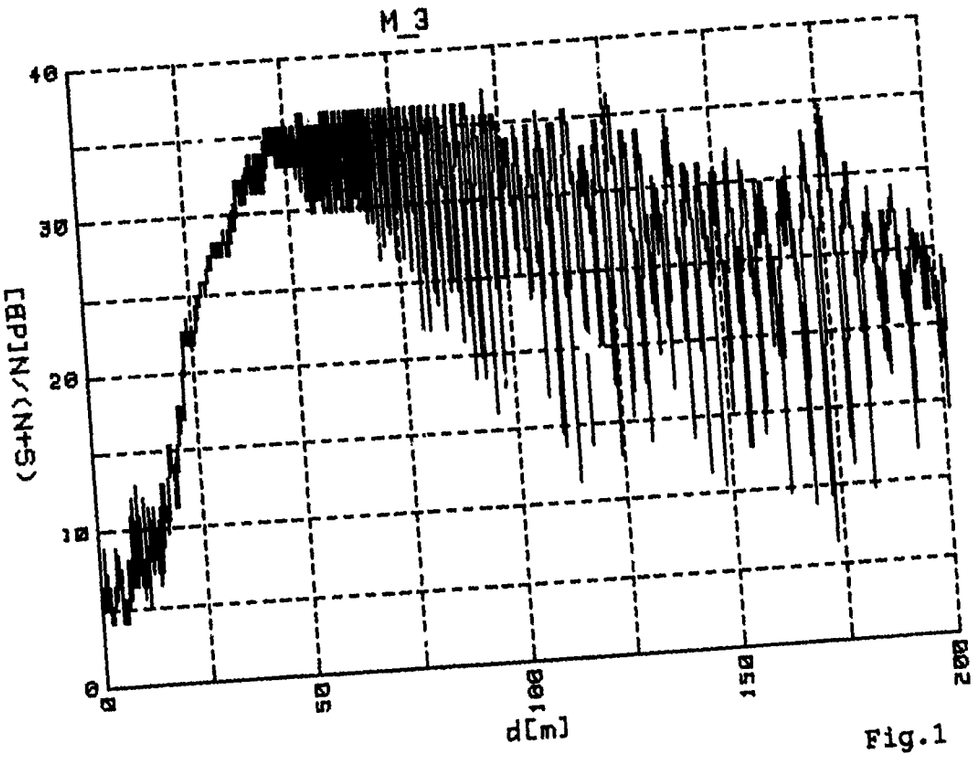

M_3
40
30
20
10
0
(S+N)/N[dB]
0
50
100
150
200
d[m]

Fig.1

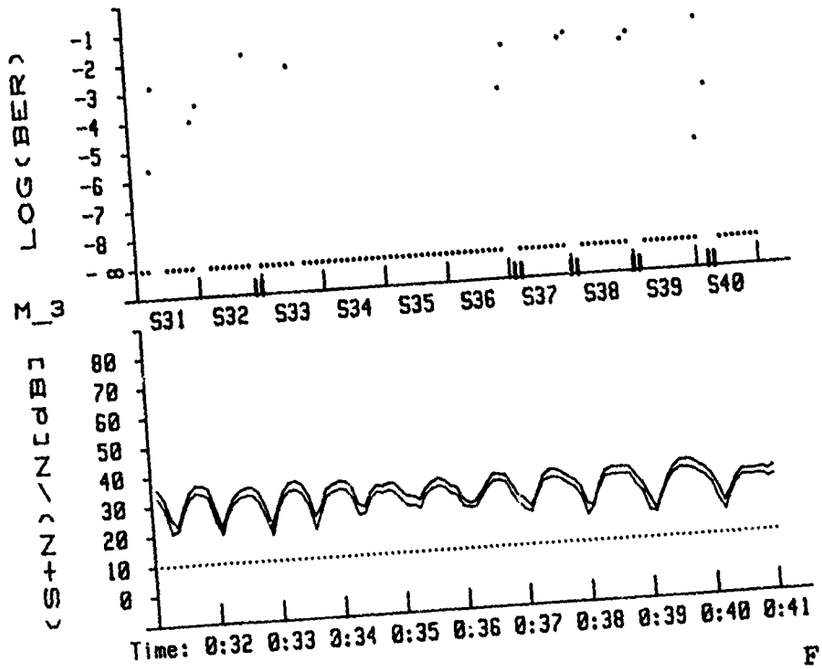

LOG(BER)
-1
-2
-3
-4
-5
-6
-7
-8
M_3
S31 S32 S33 S34 S35 S36 S37 S38 S39 S40
(S+N)/N[dB]
80
70
60
50
40
30
20
10
0
Time: 0:32 0:33 0:34 0:35 0:36 0:37 0:38 0:39 0:40 0:41

Fig.2

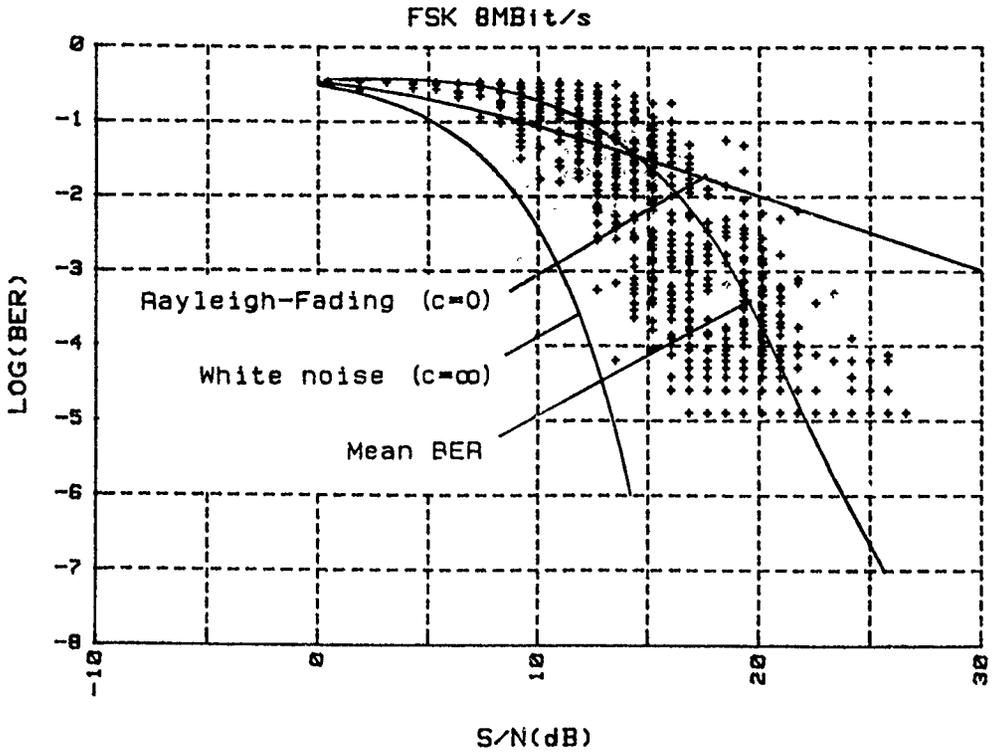

FSK 8MBit/s
0
-1
-2
-3
-4
-5
-6
-7
-8
LOG(BER)
Rayleigh-Fading (c=0)
White noise (c=∞)
Mean BER
-10
0
10
20
30
S/N(dB)

Fig.3

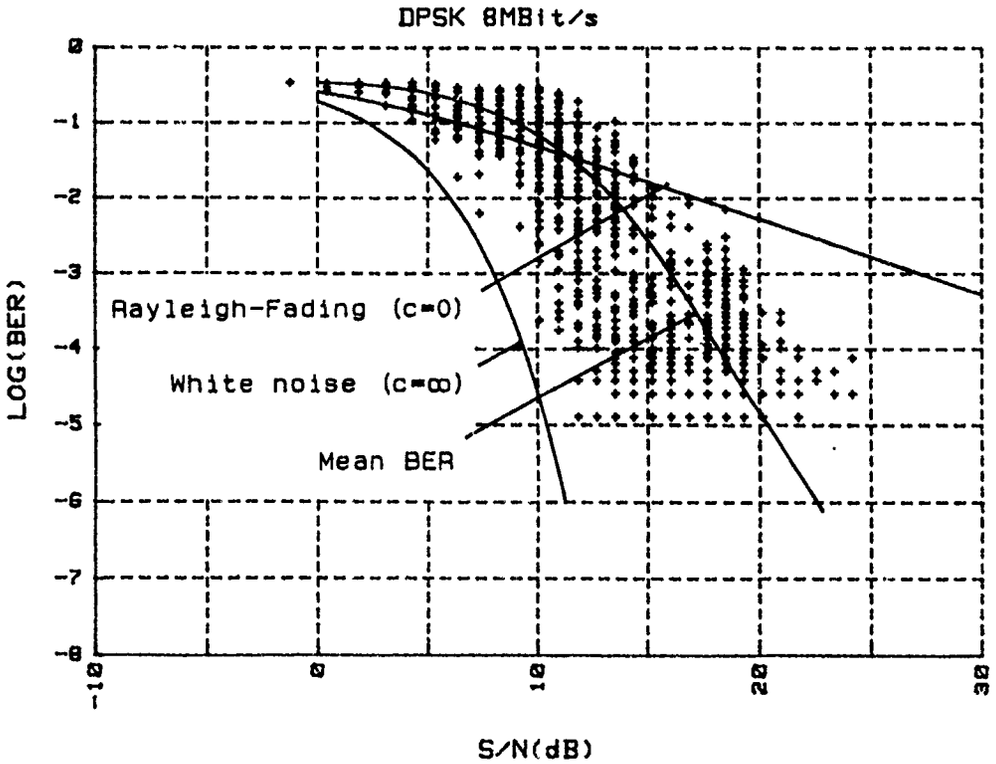

DPSK 8MBit/s
0
-1
-2
-3
-4
-5
-6
-7
-8
LOG(BER)
Rayleigh-Fading (c=0)
White noise (c=∞)
Mean BER
-10
0
10
20
30
S/N(dB)

Fig.4

FSK 8MBit/s

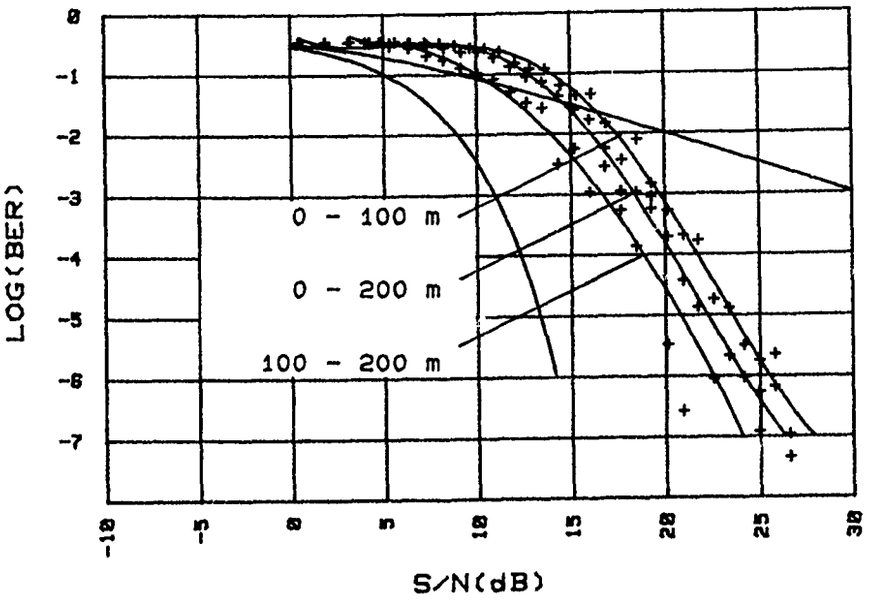


Fig.5

DPSK 8MBit/s

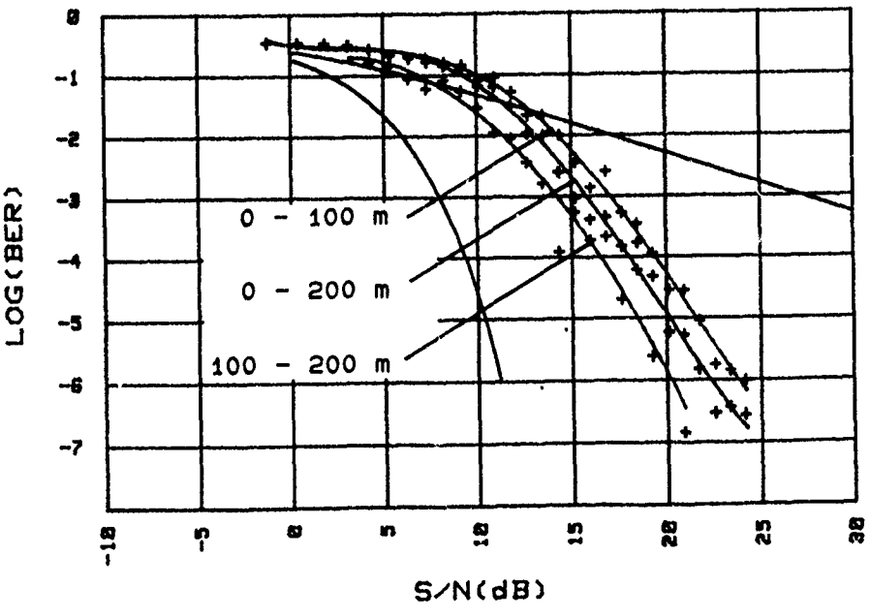


Fig.6

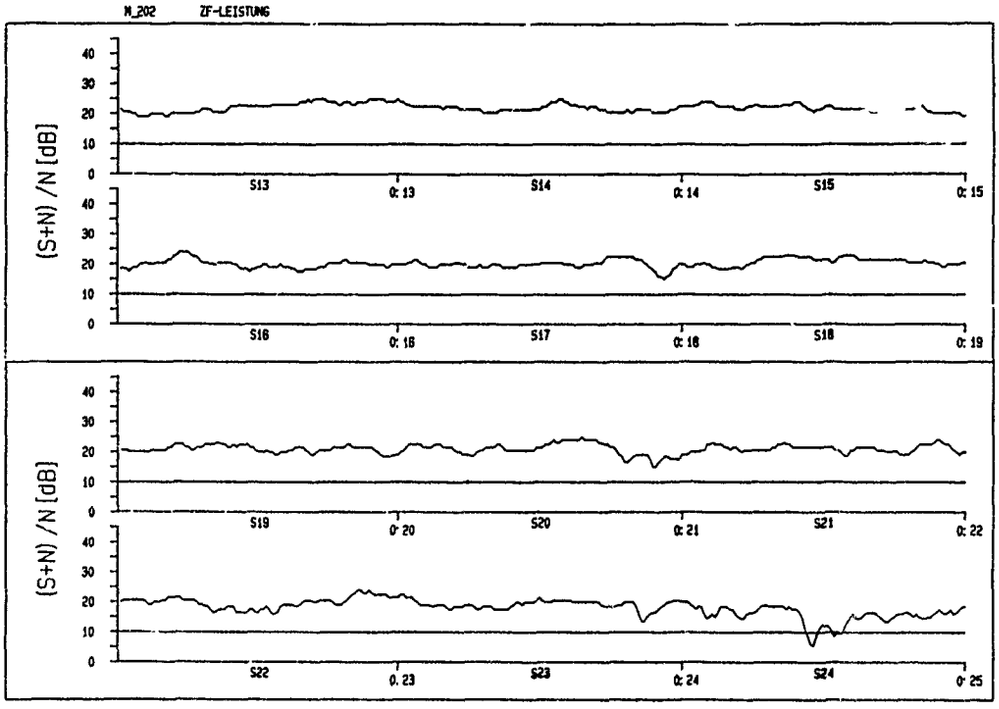

N_202 ZF-LEISTUNG
(S+N) /N [dB]
40
30
20
10
0
S13
0: 13
S14
0: 14
S15
0: 15
S16
0: 18
S17
0: 18
S18
0: 19
(S+N) /N [dB]
S19
0: 20
S20
0: 21
S21
0: 22
S22
0. 23
S23
0: 24
S24
0: 25

Fig.7

DISCUSSION

ICA FUERXER

Did you measure the impulse response of the channel ?

AUTHOR'S REPLY

Impulse response was measured. Resolution of measurement system was given to delayed signals with ≥ 1μs delay. No amplitudes of interest were found in the upper delay spectrum.

E. SCHWEICHER, BE

Is the ILO at 61GHz using a Gunn diode and an Impatt diode connected through a circulator ?

AUTHOR'S REPLY

Gunn diodes are used in each of the 3 stages. Last stage is a 2-diodes configuration.

UTILIZATION OR REDUCTION OF THE EFFECTS OF SEA CLUTTER FOR REAL AND SYNTHETIC APERTURE POLARIMETRIC RADARS

Ezekiel Bahar
Electrical Engineering Department
University of Nebraska--Lincoln
Lincoln, Nebraska 68588-0511 USA

Abstract

The presence of clutter in radar signals received over the rough sea surface presents radar engineers with a dual challenge. It can be exploited to obscure the radar signature of a friendly target. However, when the detection of the target is essential, it is necessary to devise means to minimize the effects of sea clutter on target identification.

Central to the objective to exploit or suppress the effects of sea clutter is the need for a comprehensive radio wave scattering theory that can be applied to surfaces (such as the air-sea interface) with a broad range of roughness scales. When the major contributions to the scattered fields come from the vicinity of the specular (stationary phase) points, the well known physical optics approach (based on the Kirchhoff approximations of the surface fields) provides results that are in good agreement with experimental data. However, for surfaces with very small roughness scales and slopes, the (polarization dependent) small perturbation approach of Rice is generally used. In order to exploit the salient features of the physical optics approach and the small perturbation approach, hybrid perturbation - physical optics approaches (based on a two scale model of the rough sea) have been developed. Thus the surface of the rough sea is regarded as the superposition of a small scale rough surface on a filtered surface consisting of the large scale spectral components of the sea surface. However, the wavenumber where spectral splitting is assumed to occur between the large and small scale spectral components of the rough surface is not uniquely defined. Furthermore, this approach cannot be used to determine the cross polarized backscatter cross sections.

A full wave approach, based on the complete expansion of the fields and the imposition of exact boundary conditions, is used to convert Maxwell's equations into sets of generalized telegraphists' equations for the electromagnetic field transforms. These coupled equations for the distributed system are solved iteratively to obtain explicit expressions for the single and multiple scattered fields. They are shown to satisfy the reciprocity, realizability and duality relationships in electromagnetic theory and they are invariant to coordinate transformations. In the high frequency limit when the major contributions to the scattered field come from the vicinity of the stationary phase (specular) points on the surface, they are shown to reduce to the physical/geometrical optics solutions. In the low frequency limit they reduce to the small perturbation solutions when the slopes are also negligibly small. Thus, there is no need to adopt the two scale model when the full wave approach is used.

Since the distinction between different roughness scales is made relative to the electromagnetic wavelength, the full wave approach is also applicable to scattering problems with transient or broad spectral excitations for both real and synthetic aperture radars. The full wave approach can be used to characterize the scattered field completely through the elements of the Phase/Stokes (Mueller) matrix. Thus, full wave polarimetric techniques can be readily used to suppress or enhance the effects of the sea clutter and thereby make the signature of the target more obscure or more distinct. The full wave approach can also be applied to stealth and inverse stealth technologies.

1. Introduction

Numerous scattering theories have been developed to determine the scattering and depolarization of electromagnetic waves from deterministic and randomly rough surfaces. However, the most commonly used theories, based on the physical optics and the small perturbation approaches, cannot be applied to composite surfaces, characterized by a broad range of roughness scales. In addition, several computer programs have been developed to numerically solve the integral equations for the scattered fields. These numerical solutions have limited use since they are restricted to one dimensionally rough surfaces even when supercomputers are used. Several hybrid solutions based on two scale models of the rough surfaces have also been introduced to overcome these difficulties. They are based on a combination of the physical optics (Beckmann and Spizzichino 1963), and the small perturbation (Rice 1951) approaches. Thus, a surface with small scale roughness is assumed to ride on the large scale (filtered) surface and the solution is expressed as a sum of two cross sections, one associated with the large scale surface and the other associated with the small scale surface (Valenzuela 1968, Wright 1968). The difficulties with the hybrid approach based on the two scale model, lies in the determination of the spatial wavenumber that separates the large scale surface from the small scale surface and in the underlying assumption that these two rough surfaces can be regarded as statistically independent. The large scale (filtered) surface is assumed to be sufficiently smooth such that it satisfies the large radii of curvature restriction imposed when the physical optics approach is used. The small scale surface is assumed to have a sufficiently small mean square height and slope in order to satisfy the small perturbation assumptions. These restrictions cannot in general be satisfied simultaneously and the physical optics approach does not account for the backscattered cross polarized fields. Therefore, these two scale solutions critically depend on the choice of the spatial wavenumber where spectral splitting between the large and small scale surfaces is assumed to occur (Brown 1978). Furthermore, even for surfaces that simultaneously satisfy the large radii of curvature restriction (inherent in the physical optics approach) and the small height and small slope limitation (assumed when the perturbation approach is used), the two commonly used physical optics and small perturbation solutions are not in agreement.

Since the sea surface consists of a broad range of roughness scales, it is important to employ a rough surface theory that is not encumbered by the limitations of the physical optics, perturbation or the hybrid solutions based on two scale models.

Furthermore, in order to utilize or reduce the effects of sea clutter in radar sysems, the rough surface theory must also be applicable to real and/or Synthetic Aperture Polarimetric Radars. Thus, in general, it is necessary to completely characterize the radar returns from the rough sea by the sixteen elements of the Mueller matrix that relate the incident Stokes vector to the scattered Stokes vector (Bahar and Fitzwater 1986).

The full wave approach that is based on the complete expansion of the electromagnetic fields and on the imposition of exact boundary conditions, is used to determine the singly and multiply scattered fields from surfaces that are either characterized deterministically or statistically. These full wave solutions are invariant to coordinate transformations and they satisfy the reciprocity, realizability and duality relationships in electromagnetic theory. Both the small perturbation solutions as well as the physical optics solution can be obtained directly from the full wave solutions. Furthermore, using hindsight it is possible to derive the full wave single scatter, far field solution directly from the perturbation solution by subjecting the perturbation solution to a set of coordinate transformations as well as a phase modification which correctly accounts for the surface height fluctuations. Thus, using the full wave approach, it is possible to determine the direct relationship between the small perturbation solution and the physical optics solution.

2. Review of the Full Wave Approach

For the purposes of the full wave analysis of rough surface scattering, it is necessary to define two (Cartesian) coordinate systems. The first $(\bar{a}_x,\bar{a}_y,\bar{a}_z)$, is a fixed, reference coordinate system (See Fig. 1). Since the full wave solutions are invariant to coordinate transformations, the reference coordinate system can be chosen arbitrarily, although for convenience it is usually associated with the transmitter/receiver system and/or the mean sea surface (chosen in this work to be normal to the unit vector $\bar{a}_y$). The second coordinate system $(\bar{n}_1,\bar{n}_2,\bar{n}_3)$ is associated with the local features of the rough surface. Thus $\bar{n}_2$ is the direction of the unit vector $\bar{n}$ normal to a point on the rough surface, while $\bar{n}_1$ and $\bar{n}_3$ are tangent to the rough surface at the same point (See Fig. 2).

The transverse (y,z) components of the electric and magnetic fields $\bar{E}_T$ and $\bar{H}_T$ above and below the rough surface $f(x,y,z)=y-h(x,z)=0$ are expressed completely in terms of generalized transforms, that account for the vertically and horizontally polarized waves. The complete wave spectrum consists of the radiation fields and the lateral waves (two infinite integrals) and the guided (surface) waves. Thus $\bar{E}_T$ and $\bar{H}_T$ are expressed as follows in terms of the vertically and horizontally polarized field transforms E^P and H^P (P=V vertical or H horizontal)(Bahar 1973a).

$$\bar{E}_T(x,y,z)=\sum_V \int_{-\infty}^{\infty}[E^V(x,v,w)\bar{e}_T^V+E^H(x,v,w)\bar{e}_T^H]dw \qquad (1a)$$

$$E^P(x,v,w)=\int_{-\infty}^{\infty}\bar{E}_T(x,y,z)\cdot(\bar{h}_p^T\times\bar{a}_x)dydz \quad P=V \text{ or } H \qquad (1b)$$

$$\bar{H}_T(x,y,z)=\sum_V \int_{-\infty}^{\infty}[H^V(x,v,w)\bar{h}_T^V+H^H(x,v,w)\bar{h}_T^H]dw \qquad (1c)$$

$$H^P(x,v,w)=\int_{-\infty}^{\infty}\bar{H}_T(x,y,z)\cdot(\bar{a}_x\times\bar{e}_p^T)dydz \quad P=V \text{ or } H \qquad (1d)$$

in which the symbol $\sum_V$ denotes summation over the complete wave spectrum. The vector basis functions $\bar{e}_T^P$ and $\bar{h}_T^P$ are biorthogonal to the reciprocal basis functions $\bar{e}_P^T$ and $\bar{h}_p^T$. These complete field expansions are substituted into Maxwell's equations. On employing the biorthogonal properties of the basis functions and on imposing the exact boundary conditions,

$$[\bar{n}\times\bar{E}]_{\bar{h}^-}=[\bar{n}\times\bar{E}]_{\bar{h}^+} \qquad (2)$$

$$[\bar{n}\times\bar{H}]_{\bar{h}^-}=[\bar{n}\times\bar{H}]_{\bar{h}^+} \qquad (3)$$

at the rough interface $h(x,z)$.

Maxwell's equations are converted into the following set of generalized telegraphists' equations (Scheikunoff 1955; Bahar 1973b, 1974) for the forward and backward wave amplitudes a^P and b^P respectively.

$$-\frac{da^P}{dx}-iua^P=\sum_Q\sum_{v'}\int(S_{PQ}^{BA}a^Q+S_{PQ}^{BB}b^Q)dw'-A^P \qquad (4a)$$

$$-\frac{db^P}{dx}+iub^P=\sum_Q\sum_{v'}\int(S_{PQ}^{AA}a^Q+S_{PQ}^{AB}B^Q)dw'+B^P \qquad (4b)$$

In (4) A^P and B^P are transforms of the electromagnetic sources and

$$H^V=a^V+b^V \qquad E^V=a^V-b^V \qquad (5a)$$

$$H^H=a^H-b^H \qquad E^H=a^H+b^H \qquad (5b)$$

To derive the generalized telegraphists' equation (4), it is necessary to properly take into account problems of convergence. Thus, the orders of integration (summation) and differentiation are not interchanged by employing Green's theorem. In (4), the coupling coefficients $S_{PQ}^{\alpha\beta}$ are the forward ($\alpha\neq\beta$) and backward ($\alpha=\beta$) scattering coefficients. The second subscript Q denotes the polarization of the waves incident upon the surface, while the first subscript P denotes the polarization of the scattered waves (P,Q=V,H). The telegraphists' equations may be solved numerically to account for the total scattered field (singly and multiply scattered waves). However, in this review only iterative analytical solutions are presented since they can be readily compared with other analytical single scatter solutions, such as the physical optics (Beckmann and Spizzichino) and the small perturbation solutions (Rice 1951). Higher order iterative solutions can also be derived to separately account for multiple scatter (Bahar and El-Shenawee 1990). The principal elements of the full wave solutions are shown schematically in Fig. 3.

3. Integral Expressions for the Single Scattered Radiation Field

The full wave solutions for the like and cross polarized radiation fields single scattered by two dimensionally rough surfaces $f(x,y,z)=y-h(x,z)=0$ is expressed as follows in matrix notation (Bahar 1987)

$$G^f=G_0\int_A D(\bar{n}^f,\bar{n}^i)\exp[i\bar{v}\cdot\bar{r}]dA\ G^i\equiv SG^i \qquad (6)$$

in which $\bar{n}^i$ and $\bar{n}^f$ are unit vectors in the directions of the incident and scattered fields. The vector $\bar{v}$ is

$$\bar{v}=k_0(\bar{n}^f-\bar{n}^i)=v_x\bar{a}_x+v_y\bar{a}_y+v_z\bar{a}_z \qquad (7)$$

where $k_0=\omega\sqrt{\mu_0\varepsilon_0}$ is the free space wavenumber and the $dA=dxdz/(\bar{n}\cdot\bar{a}_y)$ is the rough surface element.

The unit vector $\bar{n}$ is normal to the rough surface in the direction of ∇f and $\bar{r}$ is the position vector to a point on the rough surface. The elements of the 2x1 column matrix G^i are the incident vertically and horizontally polarized complex wave amplitudes G^{Vi} and G^{Hi} at the origin (with $\bar{n}^i x \bar{a}_y$ defined as the vector normal to the reference plane of incidence). Similarly G^f is a 2x1 column matrix whose elements are the vertically and horizontally polarized complex wave amplitudes G^{Vf} and G^{Hf} (with $\bar{n}^f x \bar{a}_y$ defined as the vector normal to the reference scatter plane) at the observation point given by the position vector

$$\bar{r}^f = x^f\bar{a}_x + y^f\bar{a}_y + z^f\bar{a}_z = r^f\bar{n}^f \quad . \tag{8}$$

Thus

$$G^i = \begin{pmatrix} G^{Vi} \\ G^{Hi} \end{pmatrix} = \begin{pmatrix} E^{Vi} \\ E^{Hi} \end{pmatrix} = \eta_o \begin{pmatrix} H^{Vi} \\ H^{Hi} \end{pmatrix} \tag{9a}$$

$$G^f = \begin{pmatrix} G^{Vf} \\ G^{Hf} \end{pmatrix} = \begin{pmatrix} E^{Vf} \\ E^{Hf} \end{pmatrix} = \eta_o \begin{pmatrix} H^{Vf} \\ H^{Hf} \end{pmatrix} \tag{9b}$$

where $\eta_o = \sqrt{\mu_o/\varepsilon_o}$ is the free space wave impedance. The coefficient G_o is

$$G_o = k_o^2 \exp(-ik_o r^f)/2\pi i k_o r^f \tag{10}$$

and a suppressed $\exp(i\omega t)$ time dependence is assumed in this work. The 2x2 rough surface element scattering matrix $D(\bar{n}^f,\bar{n}^i)$ in (4) is expressed as follows:

$$D(\bar{n}^f,\bar{n}^i) = \begin{bmatrix} D^{VV} & D^{VH} \\ D^{HV} & D^{HH} \end{bmatrix} = (-\bar{n}^i\cdot\bar{n})T^f\,F\,T^i \tag{11}$$

where $(-\bar{n}^i\cdot\bar{n}) = C_o^{in}$ and $(\bar{n}^f\cdot\bar{n}) = C_o^{fn}$ are the cosine of the local angles of incidence θ_o^{in} and scatter θ_o^{fn} for $y>h$. For $y<h$ these angles θ_l^{in} and θ_l^{fn} are given by Snell's law. The 2x2 matrix T^i transforms the incident vertically and horizontally polarized waves with respect to the fixed (reference) plane of incidence $(\bar{n}^i x \bar{a}_y)$ to the corresponding vertically and horizontally polarized waves with respect to the local plane of incidence $(\bar{n}^i x \bar{n})$ at a point $\bar{r}$ on the rough surface. The elements of the 2x2 local scattering matrix F account for the like and cross polarized scattering for a surface element dA at the point $\bar{r}$ on the rough surface (See Fig. 2). The 2x2 matrix T^f transforms the scattered vertically and horizontally polarized waves with respect to the local plane of scatter $(\bar{n}^f x \bar{n})$ at the point $\bar{r}$ on the surface back to the corresponding vertically and horizontally polarized waves with respect to the fixed plane of scatter $(\bar{n}^f x \bar{a}_y)$. The explicit expressions for the elements of the local scattering matrix F depend on the unit vectors $\bar{n}^i$ and $\bar{n}^f$ in the directions of the incident and scattered waves as well as the direction of the local normal to the rough surface. It also depends on the electromagnetic parameters (ε,μ) of the media above and below the rough interface $(f(x,y,z)=0)$ since in this work the approximate impedance boundary conditions are not employed.

$$C_o^{in}F^{VV} = \{2C_o^{in}C_o^{fn}[(\mu_r C_l^{in}C_l^{fn}\cos(\phi^{fn}-\phi^{in}) - S_o^{in}S_o^{fn}) \cdot (1-1/\varepsilon_r)+(1-\mu_r)\cos(\phi^{fn}-\phi^{in})]\} \cdot [(C_o^{in}+\eta_r C_l^{in})(C_o^{fn}+\eta_r C_l^{fn})(C_o^{in}+C_o^{fn})]^{-1} \tag{12a}$$

$$C_o^{in}F^{HH} = \{2C_o^{in}C_o^{fn}[(\varepsilon_r C_l^{in}C_l^{fn}\cos(\phi^{fn}-\phi^{in}) - S_o^{in}S_o^{fn}) \cdot (1-1/\mu_r)+(1-\varepsilon_r)\cos(\phi^{fn}-\phi^{in})]\} \cdot [(C_o^{in}+C_l^{in}/\eta_r)(C_o^{fn}+C_l^{fn}/\eta_r)(C_o^{in}+C_o^{fn})]^{-1} \tag{12b}$$

$$C_o^{in}F^{HV} = \{-\sin(\phi^{fn}-\phi^{in})2C_o^{in}C_o^{fn}\eta_r[(1-1/\varepsilon_r)C_l^{in} - (1-1/\mu_r)C_l^{fn}]\}[(C_o^{in}+\eta_r C_l^{in})(C_o^{fn}+C_l^{fn}/\eta_r)(C_o^{in}+C_o^{fn})]^{-1} \tag{12c}$$

$$C_o^{in}F^{VH} = \{\sin(\phi^{fn}-\phi^{in})2C_o^{in}C_o^{fn}\eta_r[(1-1/\mu_r)C_l^{in} - (1-1/\varepsilon_r)C_l^{fn}]\}[(C_o^{in}+C_o^{in}/\eta_r)(C_o^{fn}+\eta_r C_l^{fn})(C_o^{in}+C_o^{fn})]^{-1} \tag{12d}$$

In (12) $S_{o,l}^{in}$ and $S_{o,l}^{fn}$ are the sine of the angles of incidence and scatter and $\phi^{fn}-\phi^{in}$ is the angle between the local planes of incidence and scatter. It is interesting to note that at the stationary phase points of the integral expression for the scattered fields (5), the unit vector $\bar{n}$ normal to the rough surface is in the direction of the vector $\bar{v}$ (6),

$$\bar{n} \rightarrow \bar{n}_s = \bar{v}/v \tag{13}$$

At these stationary phase (specular) points on the rough surface, the surface element scattering matrix $F(\bar{n}^f,\bar{n}^i)$ reduces to

$$F(\bar{n}^f,\bar{n}^i)\Big|_{\bar{n}_s} = \begin{pmatrix} R_V(\theta_{op}^i) & \\ 0 & R_H(\theta_{op}^i) \end{pmatrix} \tag{14}$$

where $\cos\theta_{op}^i = -\bar{n}^i\cdot\bar{n}_s = \bar{n}^f\cdot\bar{n}_s$ and θ_{op}^i is equal to the local angle of incidence and scatter at the specular points. Thus it is readily shown that if in the integrand of the expressions for the scattered fields (4), the unit vector $\bar{n}$ is replaced by its value $\bar{n}_s$ at the stationary phase points (before integration) the full wave solutions reduce to the physical optics solutions (Beckmann and Spizzichino 1963). Therefore, for the physical optics approach to be valid in the high frequency limit, it is not only necessary for the radii of curvature to be large compared to wavelength, but the major contributions to the scattered fields must come from the vicinity of the specular points of the rough surface. This is a principal reason why the physical optics approach cannot be used to correctly predict the enhanced backscattered fields even at optical frequencies (Bahar and Fitzwater 1989). If the integral for the scattered fields is evaluated analytically using the stationary phase approximation the full wave solution reduces to the Geometric Optics solution.

It is also interesting to demonstrate the direct link between the full wave solution (4) and the low frequency small perturbation solution. Thus, it can be shown that if in the integral expression for the scattered fields (4), the unit vector $\bar{n}$ normal to the rough surface is replaced by the unit vector $\bar{a}_y$ normal to the reference (mean) plane y=o,

(negligible slopes) and in addition, if the exponential function $\exp(i\bar{v}\cdot h(x,z)\bar{a}_y)$ is replaced by the first two terms of its Taylor series expansion $1+ik_0h(\bar{n}^f-\bar{n}^i)\cdot\bar{a}_y$, the full wave solution reduces to the sum of the reflected field from the flat (unperturbed) surface and the diffuse scattered field derived by Rice (1951), based on the small perturbation approach. On hindsight it is therefore possible to trace the rather elusive relationship between the small perturbation (low frequency) solution (Rice 1951) and the high frequency physical optics solution (Beckmann and Spizzichino (1963) even though the two solutions are not in agreement for surfaces that simultaneously satisfy the small perturbation restrictions and the large radii of curvature criteria. To this end, one can start with Rice's polarization dependent small perturbation solution for the fields scattered by a surface element dA in the neighborhood of the point $\bar{r}$ on the rough surface. At the point $\bar{r}$, the normal to the surface element at height $y=h(x,z)$ is given by the unit vector $\bar{n}=\nabla f/|\nabla f|$. Thus, on subjecting Rice's solution to the principle of invariance under coordinate transformations and replacing the factor appearing in his solution $ik_0h(\bar{n}^f-\bar{n}^i)\cdot\bar{a}_y$ (associated with the surface height fluctuations) by the exponential function $\exp(i\bar{v}\cdot h(x,z)\bar{a}_y)$ (associated with the corresponding phase fluctuation), Rice's solution is transformed into the full wave solutions for the singly scattered radiation (far) fields. If at this point the value for the unit vector $\bar{n}$ normal to the rough surface is replaced by its value $\bar{n}_s$ at the stationary phase (specular) points, the low frequency small perturbation solution transforms directly into the high frequency physical optics solution. The inter-relationships between the small perturbation, the full wave, the physical/geometric optics and the hybrid solutions (based on two scale models of the rough surface) are shown schematically in Fig. 4. Thus, when one applies the full wave approach to evaluate the fields scattered by composite surfaces consisting of a very broad range of roughness scales, it is not necessary to adopt the two scale model of the rough surface.

4. The Scattering Cross Sections and the Stokes Matrix Elements for Rough Surfaces

For radar remote sensing application, the random rough surfaces are usually characterized by their normalized like and cross polarized diffuse scattering cross sections $\langle\sigma^{PQ}\rangle(P,Q=V,H)$. They are defined as follows in terms of the incident and scattered fields

$$\langle\sigma^{PQ}\rangle = \frac{\langle|E^{Pf}|^2\rangle - |\langle E^{Pf}\rangle|^2}{|E^{Qi}|^2}\,\frac{4\pi(r^f)^2}{A_y} \qquad (15)$$

in which A_y is the projection of the rough surface area on the reference ($y=0$) plane and r^f is the distance to the observation point. The symbol $\langle\cdot\rangle$ denotes the statistical average. The normalized scattering cross sections therefore relate the diffuse (total minus the coherent) scattered intensity (with polarization P=V,H) to the incident intensity (with polarization Q=V,H). Similar expressions for the scattering cross sections can be written in general for elliptical incident and scatter polarization. However, for the purposes of enhancing or suppressing special features of the radar returns, the optimal polarizations of the transmitter and receiver are not known a priori. In these cases, it is necessary to characterize the scattered fields by the 4x4 Stokes matrix elements which contain both magnitude and relative phase data. The Stokes matrix relates the scattered Stokes vector elements to the incident Stokes vector elements. A modified form of the Stokes vector is as follows:

$$[I] = \begin{bmatrix} I_1 \\ I_2 \\ U \\ V \end{bmatrix} = \begin{bmatrix} \langle E_1E_1^*\rangle \\ \langle E_2E_2^*\rangle \\ 2\mathrm{Re}\langle E_1E_2^*\rangle \\ 2\mathrm{Im}\langle E_1E_2^*\rangle \end{bmatrix}, \qquad (16)$$

in which E_1 and E_2 are the vertically and horizontally polarized components of the electric field while $\mathrm{Re}\langle\cdot\rangle$ and $\mathrm{Im}\langle\cdot\rangle$ denote the real and imaginary parts. Thus, on expressing the relationships between the incident and scattered vertically and horizontally polarized fields as follows:

$$\begin{bmatrix} E_1^f \\ E_2^f \end{bmatrix} = \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix}\begin{bmatrix} E_1^i \\ E_2^i \end{bmatrix}\,\frac{\exp(-ik_0r^f)}{r^f}\,, \qquad (17)$$

the Stokes matrix that relates the incident Stokes vector to the scattered Stokes vector is given by Bahar and Fitzwater (1987). (See Appendix A). The four elements $m_{ij}(i,j=1,2)$ of the Stokes matrix are therefore related to the like and cross polarized scattering cross sections.

Explicit full wave expressions for all sixteen elements of the Stokes matrix need to be evaluated in order to obtain a complete polarimetric description of the rough surface scattered field. Extensive work has been carried out to determine the optimum polarization of the transmitter/receiver to enhance special features of the radar signals. Multifrequency polarimetric radar systems provide improved capabilities to enhance or suppress these features (Zebker et al. 1990). The Stokes matrix is also useful for Synthetic Aperture Radar (SAR) applications.

5. Concluding Remarks

In this paper the principal elements of the full wave solutions have been summarized. The full wave solutions are invariant to coordinate transformations and they satisfy the reciprocity, realizability and duality relationships in electromagnetic theory. Limiting forms of the full wave solutions are shown to reduce to the low frequency small perturbation solutions (Rice 1951) and the high frequency physical optics solutions (Backmann and Spizzichino 1963). Thus, the full wave solution provides the only direct relationship between the low frequency small perturbation ($\beta=4k_0\langle h^2\rangle \ll 1$, $\bar{n}\approx\bar{a}_y$) and the high frequency physical optics ($\beta \gg 1$ $kr_c \gg 1$, r_c is the radius of curvature) solutions. Tracing this relationship has been a rather elusive endeavor since even for surfaces that simultaneously satisfy the small perturbation limitation and the large radii of curvature restriction, these two solutions are not in agreement.

Thus, there are no complications in applying the full wave solutions to scattering by composite surfaces with multiple scales of roughness and there is no need to adopt hybrid (perturbed-physical optics) solutions based on the artificial decomposition of the surface into two surfaces with different scales of roughness. These hybrid solutions critically depend on the choice of the spatial wavenumber k_d at which it is assumed the separation occurs between the large and small scale surface. The full wave-two scale approximation can be expressed as a weighted sum of two cross sections and the like polarized cross section is less sensitive to the choice of k_d. However, since the physical optics approximation for the backscattered

cross polarized return is zero (Brown 1978), the hybrid solutions based on the two scale modes cannot be used to derive the cross polarized cross sections for composite rough surfaces. The full wave solutions for the Stokes matrix elements can therefore, be used to analyze data obtained from multifrequency polarimetric radars at all angles of incidence (Zebkar et al. 1990). This provides the users with a broad choice of parameters (including wavelength, angle of incidence and polarization) for purposes of the selective suppression or enhancement of background clutter received by both real and synthetic aperture radars. Since the approximate impedance boundary conditions are not used in the full wave analysis, it could also be applied to scatterers with a broad range of electromagnetic parameters (ϵ,μ). The full wave approach has been applied to random distributions of scatterers of irregular shape (Bahar and Fitzwater 1987) and it can therefore be used to realistically model foliage as well as foliage-covered terrain.

6. Acknowledgments

This research was supported by U.S. Army Research Office contract DAAL0387-K-0085 and U.S. Office of Naval Research contract N00014-87-K-0177. The author also wishes to thank W. Flood and D. Johnsen for their encouragement. The manuscript was prepared by E. Everett.

7. References

Bahar, E., Depolarization of electromagnetic waves excited by distributions of electric and magnetic sources in inhomogeneous multilayered structures of arbitrarily varying thickness--Generalized field transforms, J. Math. Phys., 14(11), 1502-1509, 1973a.

Bahar, E., Depolarization of electromagnetic waves excited by distributions of electric and magnetic sources in inhomogeneous multilayered structures of arbitrarily varying thickness--Full wave solutions, J. Math. Phys., 14(11) 1510-1515, 1973b.

Bahar, E., Depolarization in nonuniform multilayered structures--Full wave solutions, J. Math. Phys., 15(2), 202-208, 1974.

Bahar, E., and M. A. Fitzwater, Depolarization and Backscatter Enhancement in Light Scattering from Random Rough Surfaces - Theory and Experiment, J. Opt. Soc. Am A, 6, 33-43, 1989.

Bahar, E., and M. A. Fitzwater, Co-polarized and Cross-polarized Incoherent Specific Intensities for Waves at Oblique Incidence upon Layers of Finitely Conducting Particles of Irregular Shape, J. Opt. Soc. Am A, 4, 41-56, 1987.

Bahar, E., and M. El-Shenawee, Full Wave Multiple Scattering from Rough Surfaces, Proceedings of the IEEE AP-S International Symposium and URSI Radio Science Meeting, IV, 1548-1551, 1990.

Beckmann, P., and A. Spizzichino, The Scattering of Electromagnetic Waves from Rough Surfaces, Macmillan, New York, 1963.

Brown, G. S., Backscattering from a Gaussian-distributed perfectly conducting rough surface, IEEE Trans. Antennas Propag., AP-26(3), 472-482, 1978.

Rice, S. O., Reflection of electromagnetic waves from a slightly rough surface, Commun. Pure Appl. Math., 4, 351-378, 1951.

Schelkunoff, S. A., Conversion of equations into generalized telegraphists' equations, Bell Syst. Tech. J., 34, 999-1043, 1955.

Valenzuela, G. R., Scattering of electromagnetic waves from a tilted slightly rough surface, Radio Sci., 3(11), 1051-1066, 1968.

Wright, J. W., A new model for sea clutter, IEEE Trans. Antennas Propag., AP-16(2), 217-223, 1968.

Zebkar, H. A., J. J. vanZyl, L. Norikane and F. Burnette, Calibrated imaging radar polarimetry: Technique examples and applications, Proceedings of the URSI Commission F Meeting, University of Maryland, College Park, Maryland, p. 100, May 1985.

8. Appendix A

The Stokes matrix [M] as expressed as follows in terms of the matrix elements f_{ij} (17)

$$[M] = \begin{bmatrix} \langle|f_{11}|^2\rangle & \langle|f_{12}|^2\rangle & \mathrm{Re}\langle f_{11}f_{12}^*\rangle & \mathrm{Im}\langle f_{11}f_{12}^*\rangle \\ \langle|f_{21}|^2\rangle & \langle|f_{22}|^2\rangle & \mathrm{Re}\langle f_{21}f_{22}^*\rangle & -\mathrm{Im}\langle f_{21}f_{22}^*\rangle \\ 2\mathrm{Re}\langle f_{11}f_{21}^*\rangle & 2\mathrm{Re}\langle f_{12}f_{22}^*\rangle & \mathrm{Re}\langle f_{11}f_{22}^* + f_{12}f_{21}^*\rangle & -\mathrm{Im}\langle f_{11}f_{22}^* - f_{12}f_{21}^*\rangle \\ 2\mathrm{Im}\langle f_{11}f_{21}^*\rangle & 2\mathrm{Im}\langle f_{12}f_{22}^*\rangle & \mathrm{Im}\langle f_{11}f_{22}^* + f_{12}f_{21}^*\rangle & \mathrm{Re}\langle f_{11}f_{22}^* - f_{12}f_{21}^*\rangle \end{bmatrix} \quad (A1)$$

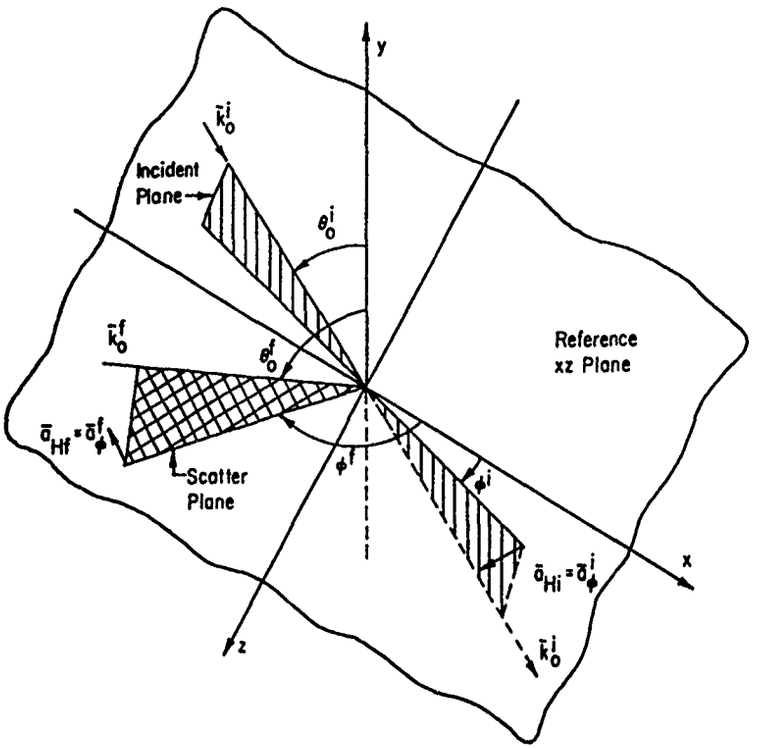


Fig. 1. Planes of incidence and scatter with respect to the reference coordinate system. (Mean reference plane for rough surface is y=o).

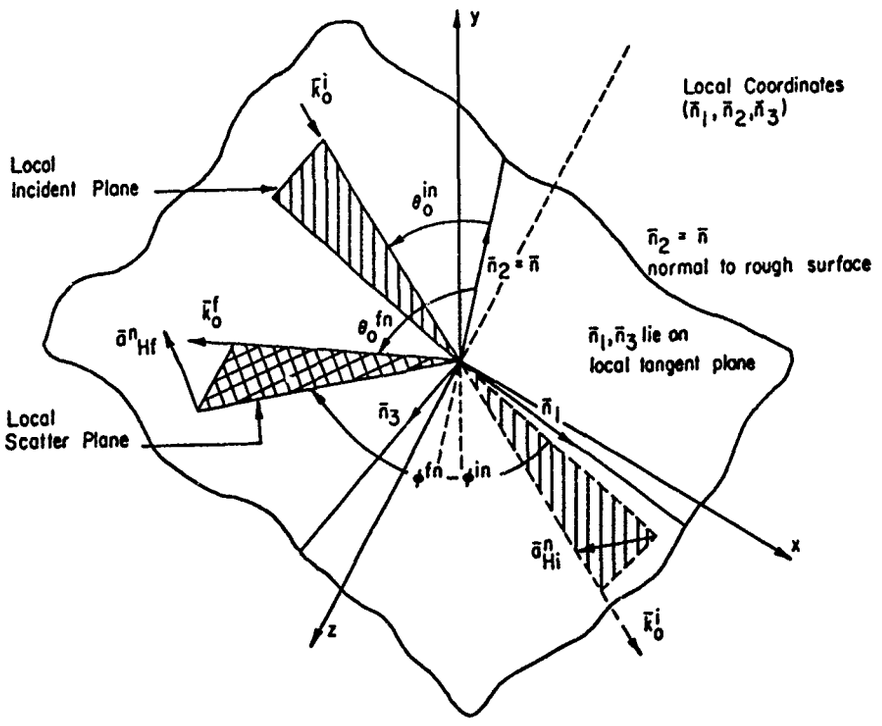


Fig. 2. Local planes of incidence and scatter and local coordinate system ($\bar{n}_1,\bar{n}_2,\bar{n}_3$).

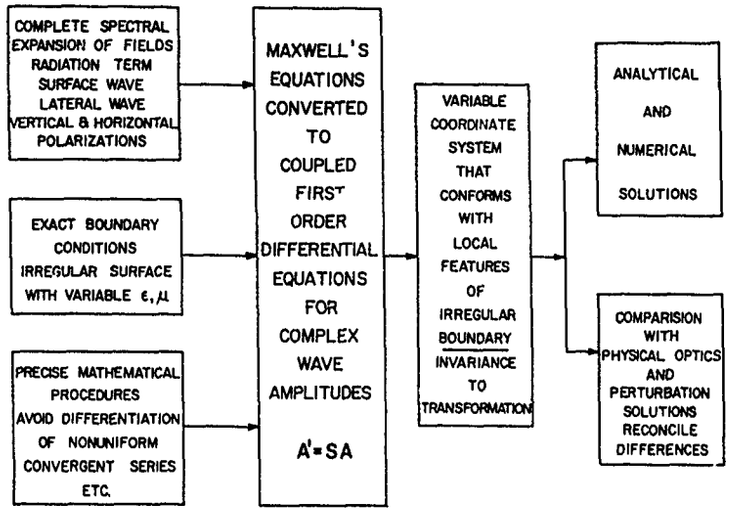


Fig. 3. Principal elements of the full wave approach.

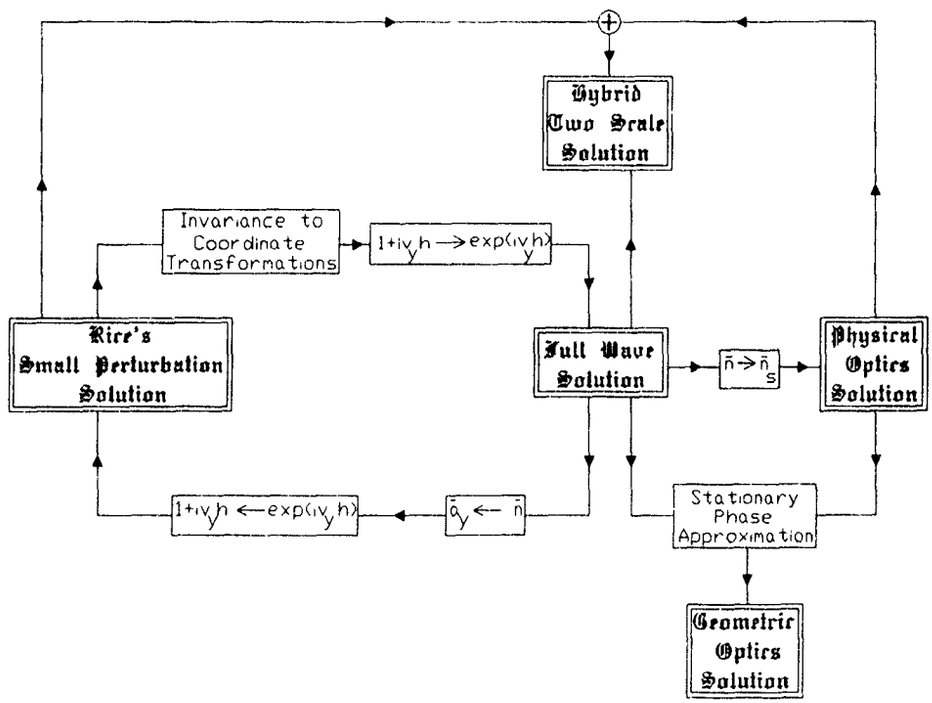


Fig. 4. Principal properties of the full wave approach.

EMISSIVITY AND TRANSMISSIVITY OF A RANDOMLY LAMINAR STRUCTURE

Cornel Eftimiu
McDonnell Douglas Research Laboratories
P.O.Box 516, St. Louis, MO 63166, USA

ABSTRACT

We consider a randomly laminar structure of constant width and investigate the scattering of scalar waves for active remote sensing and radiometric purposes.

The approach is based on analytic wave theory involving the expansion of the field in Wiener-Hermite functionals which are statistically orthogonal. A hierarchy of approximations is established and shown to represent in each order a partial summation of perturbation theory terms. The average field as well as the incoherently scattered field are determined, yielding the emissivity and transmissivity of the randomly laminar medium. The special case of small correlation length is examined in detail and proved to fulfill the energy conservation requirement.

1. INTRODUCTION

The study of scattering and propagation of waves in media exhibiting significantly complex and irregular variations in properties like density, index of refraction, dielectric constant, etc. is routinely performed by considering the medium a realization of a statistical ensemble with, more often than not, a Gaussian probability distribution.

The problem has been treated by a variety of methods, amply described in excellent reviews [1-5]. The purpose of the present work is to add to these methods an approach which is novel in this context, and which is based on the Wiener-Hermite functional expansion of the field [6], which can be also explained in terms of the familiar approach based on perturbation theory. The direct application of (renormalized) perturbation theory, in the form it evolved by natural extension from Quantum Field Theory, has drawbacks which severely limit its efficiency. A central role in such perturbative calculations is played by the Dyson equation for the average, coherent component of the field, and by the Bethe-Saltpeter equation for the incoherent component. Unfortunately, neither of these equations can be solved exactly, and their approximate solutions via perturbation theory expansions, because of inherently arbitrary assumptions, may lead to violations of fundamental physical principles, like the conservation of energy [7]. Particularly troublesome is the fact that the approximations introduced in solving these equations are not necessarily mutually consistent.

The approach based on the Wiener-Hermite functional expansion, which will be briefly described below, simultaneously yields both coherent and incoherent components of the field, rendering transparent their physical interdependence.

2. PROBLEM STATEMENT

The essential problem under consideration is sketched in Fig. 1. An incoming plane wave $\exp(-ik_0 z)$ propagates from $z = -\infty$ and is normally incident on the face $z = 0$ of a medium with an index of refraction randomly distributed in the z-direction only. The total field in the half-space $z \leq 0$ can thus be written as

$$\psi_-(z) = e^{-ik_0 z} + Re^{ik_0 z} \qquad (1)$$

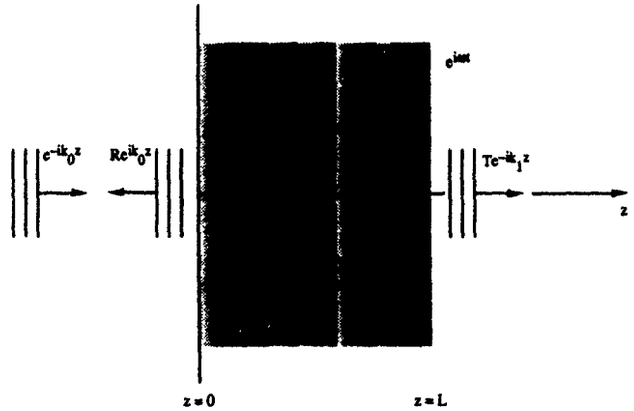


Figure 1. The wave propagation and scattering problem.

where R is the reflection coefficient, itself a random number. The field (1) can be separated into a coherent (average) component and an incoherent component:

$$\psi_-(z) = \psi_-^{coh}(z) + \psi_-^{incoh}(z) \tag{2}$$

with

$$\psi_-^{coh}(z) = e^{-ik_0 z} + \langle R \rangle\, e^{ik_0 z}$$

$$\psi_-^{incoh}(z) = \hat{R}\, e^{ik_0 z} \tag{3}$$

where $\hat{R} = R - \langle R \rangle$, the brackets indicating ensemble average.

The average energy carried by these components can be easily found by calculating the z-component of the Poynting vector:

$$S_-^{coh} = -\,\mathrm{Im}\left[\left(\psi_-^{coh}(z)\right)^* \frac{d}{dz}\,\psi_-^{coh}(z)\right] = k_0\left[1 - |\langle R \rangle|^2\right]$$

$$S_-^{incoh} = -\,\mathrm{Im}\left\langle\left(\psi_-^{incoh}(z)\right)^* \frac{d}{dz}\,\psi_-^{incoh}(z)\right\rangle = -\,k_0 \langle |\hat{R}|^2 \rangle\ . \tag{4}$$

If one defines the reflectivity of the medium as

$$r = \langle |R|^2 \rangle \tag{5}$$

it follows that the emissivity can be directly expressed in terms of the total energy flux:

$$\begin{aligned} e &= 1 - r \\ &= 1 - |\langle R \rangle|^2 - \langle |\hat{R}|^2 \rangle \\ &= \left[S_-^{coh} + S_-^{incoh}\right]/k_0 \\ &= S_-^{tot}/k_0\ . \end{aligned} \tag{6}$$

If the randomly distributed structure is of finite extent L, the field for $z \geq L$ is of the form

$$\psi_L(z) = T\, e^{-ik_1 z} \tag{7}$$

in which k_1 is the (constant and, for simplicity, real) wavenumber of the uniform half-space $z \geq 0$, and T is the transmission coefficient. Here as well, we can write

$$\psi_L(z) = \psi_L^{coh}(z) + \psi_L^{incoh}(z) \tag{8}$$

and

$$S_L^{coh} = k_1\, |\langle T \rangle|^2$$

$$S_L^{incoh} = k_1 \langle |\hat{T}|^2 \rangle \tag{9}$$

with $\hat{T} = T - \langle T \rangle$. The tranmissivity, defined as

$$t = \frac{k_1}{k_0} \langle |T|^2 \rangle \tag{10}$$

(note that the energy flux of the incident field is simply k_0), is thus

$$\begin{aligned} t &= \left[S_L^{coh} + S_L^{incoh}\right]/k_0 \\ &= S_L^{tot}/k_0 \end{aligned} \tag{11}$$

The quantities e and t are the main measurable parameters of the problem, and our objective is to determine them by a direct examination of the field propagating and scattering inside the laminar structure. In the randomly distributed medium, i.e. for $0 \leq z \leq L$, the field $\Psi(z)$ is a solution of the equation

$$\left(\frac{d^2}{dz^2} + k^2\left[1 + f(z)\right]\right)\psi(z) = 0 \tag{12}$$

subject to continuity conditions at $z = 0$ and $z = L$, where $f(z)$ is a real, stochastic, stationary, normal process with zero mean: $\langle f(z) \rangle = 0$. The latter assignment identifies k as the wavenumber in the averaged medium. The covariance of the fluctuating (square of the) wavenumber, i.e., the correlation function

$$c(z - z') = \langle f(z)\, f(z') \rangle, \qquad 0 \le z, z' \le L \tag{13}$$

is assumed given.

3. WIENER-HERMITE EXPANSION

To solve the stated problem, we first introduce the Wiener-Hermite set of functions:

$$\begin{aligned} H_0 &= 1 \\ H_1(z) &= f(z) \\ H_2(z) &= f(z_1)\, f(z_2) - c(z_1 - z_2) \\ H_3(z) &= f(z_1)\, f(z_2)\, f(z_3) - f(z_1)\, c(z_2 - z_3) \\ &\quad - f(z_2)\, c(z_1 - z_3) - f(z_3)\, c(z_1 - z_2) \\ &\;\;\vdots \end{aligned} \tag{14}$$

which are constructed such that its elements are statistically orthogonal, i.e.,

$$\langle H_m(z_1, z_2, \ldots, z_m)\, H_n(z_1, z_2, \ldots, z_n) \rangle = 0 \tag{15}$$

for each fixed $n > 0$ and all $m = 0, 1, \ldots, n - 1$. We now seek the solution of (12) in the form of the functional expansion

$$\begin{aligned} \psi(z) &= \psi_0(z) + \psi_1(z) + \psi_2(z) + \cdots \\ &= \phi_0(z) + \int_0^L dz_1\, \phi_1(z; z_1)\, H_1(z_1) \\ &\quad + \int_0^L \int_0^L dz_1 dz_2\, \phi_2(z; z_1, z_2)\, H_2(z_1, z_2) + \cdots \end{aligned} \tag{16}$$

in which the coefficient functions $\phi_n(z; z_1, \ldots, z_n)$ are deterministic and, for $n > 1$, symmetric in the integration variables. Because of (15), the first moment, i.e., the average field is simply

$$\langle \psi(z) \rangle = \phi_0(z)$$

and the second moment of the field is

$$\begin{aligned} \langle \psi^*(z)\, \psi(z') \rangle &= \phi_0^*(z)\, \phi_0(z') + \int_0^L \int_0^L dz_1 dz_1'\, \phi_1^*(z; z_1)\, \phi_1(z'; z_1')\, c(z_1 - z_1') \\ &\quad + \int_0^L \int_0^L \int_0^L \int_0^L dz_1 dz_2 dz_1' dz_2'\, \phi_2^*(z; z_1, z_2)\, \phi_2(z'; z_1', z_2') \\ &\quad \cdot \left[c(z_1 - z_1')\, c(z_2 - z_2') + c(z_1 - z_2')\, c(z_2 - z_1') \right] + \cdots \end{aligned} \tag{18}$$

so that the problem reduces to the determination of the coefficients ϕ_n. A hierarchical system of equations can be derived by substituting (16) into (12), using the recurrence relation

$$\begin{aligned} f(z)\, H_n(z_1, \ldots, z_n) &= H_{n+1}(z, z_1, \ldots z_n) \\ &\quad + \sum_{i=1}^{n} c(z - z_i)\, H_{n-1}(z_1, \ldots, z_{i-1}, z_{i+1}, \ldots, z_n) \end{aligned} \tag{19}$$

and then ensemble averaging after successively multiplying by H_n $(n = 0, 1, \ldots)$. One thus obtains an infinite system of equations, the first few of which are

$$\left(\frac{d^2}{dz^2} + k^2 \right) \phi_0(z) = - k^2 \int_0^L dz_1 \phi_1(z; z_1)\, c(z - z_1) \tag{20}$$

$$\left(\frac{d^2}{dz^2} + k^2 \right) \phi_1(z; z') = - k^2\, \phi_0(z)\, \delta(z - z_1) - 2k^2 \int_0^L dz_1 (z; z_1, z')\, c(z - z_1) \tag{21}$$

$$\left(\frac{d^2}{dz^2} + k^2\right) \phi_2(z;z',z'') = -\frac{1}{2} k^2\left[\phi_1(z;z'')\,\delta(z - z') + \phi_1(z;z')\,\delta(z - z')\right] - 6k^2 \int_0^L dz_1\, \phi_3(z;z',z'',z_1)\, c(z - z_1) \tag{22}$$

For all practical intents and purposes, one can conceptualize solving this system of equations only after closing it through truncation. We define a truncation of order N as the approximation in which only the first N equations of the system are retained, and one assumes that for all $n \geq N$, $\phi_n = 0$. The system of N equations, for the functions ϕ_n, with $n = 0,1,\ldots, N-1$, is thus closed and can be solved. For example, if one considers the order $N = 3$ of truncation, its equations are (20) and (21) plus the truncated equation (22):

$$\left(\frac{\partial^2}{\partial z^2} + k2\right) \phi_2(z;z',z'') = -\frac{1}{2} k^2\left[\phi_1(z;z')\,\delta(z - z'') + \phi_1(z;z'')\,\delta(z - z')\right] . \tag{23}$$

Introducing now the Green's function $G(z,z')$, a solution of the equation

$$\left(\frac{\partial^2}{\partial z^2} + k^2\right) G(z;z') = -\delta(z - z') . \tag{24}$$

the explicit expression of which will be given below, we can write

$$\phi_2(z;z',z'') = \frac{1}{2} k^2\left[G(z;z')\,\phi_1(z';z'') + G(z,z'')\,\phi_1(z'';z')\right] . \tag{25}$$

Substitution of (25) into (21), yields the equation

$$\left(\frac{\partial^2}{\partial z^2} + k^2\right) \phi_1(z;z') = -k^2\,\phi_0(z)\,\delta(z - z') + k^4 \int_0^L dz_1\, \left[G(z,z_1)\,\phi_1(z_1;z') + G(z,z')\,\phi_1(z';z_1)\right] c(z - z_1) . \tag{26}$$

which, together with (20) forms a closed system of equations for ϕ_0 and ϕ_1.

Clearly, the sequence of hierarchic approximations can be interrupted in principle at any order. However, we shall show below that even the low orders of truncation are applicable to and yield adequate solutions, corresponding to interesting physical circumstances.

4. GREEN'S FUNCTION

A detailed description of the Green's function must be provided before proceeding to an examination of the truncated system of equations derived above. For $0 \leq z,z' \leq L$, the solution of the equation (25) represents waves originating at the source point z', and propagating in both directions, towards $z = 0$ or $z = L$. A general representation is thus

$$G(z,z') = \frac{1}{4ik}\left[e^{-ik|z-z'|} - e^{ik|z-z'|} + a(z')\,e^{-ikz} + b(z')\,e^{ikz}\right] . \tag{27}$$

Because of continuity at $z = 0$ and $z = L$, the Green's function for $z < 0$ and $0 \leq z' \leq L$ reads

$$G_-(z,z') = G(0,z')\,e^{ik_0 z} \tag{28}$$

while for $z > L$ and $0 \leq z' \leq L$,

$$G_L(z,z') = G(L,z')\,e^{ik_1(L-z)} \tag{29}$$

Requiring also continuity at $z = 0$ and $z = L$ of the first order derivatives with respect to z allows us to determine the functions $a(z')$ and $b(z')$, resulting in the following explicit expression:

$$G(z,z') = \frac{1}{4ik}\left[e^{-ik|z-z'|} - e^{ik|z-z'|} + A\left(e^{-ik(z-z')} + e^{ik(z-z')}\right) + B\,e^{-ik(z+z')} + C\,e^{ik(z+z')}\right] \tag{30}$$

where

$$A = \frac{(k + k_0)(k + k_1)\,e^{ikL} + (k - k_0)(k - k_1)\,e^{-ikL}}{(k + k_0)(k + k_1)\,e^{ikL} - (k - k_0)(k - k_1)\,e^{-ikL}} \tag{31}$$

and

$$A - 1 = \frac{k - k_1}{k + k_1} e^{-2ikL} \quad B = \frac{k - k_0}{k + k_0} \quad C \ . \tag{32}$$

Note also that

$$A^2 - B\,C = 1 \ . \tag{33}$$

Alternatively, we may rewrite (30) as

$$G(z,z') = \frac{A + 1}{4ik} \begin{cases} U(z)\,V(z') & 0 \le z' \le z \le L \\ U(z')\,V(z) & 0 \le z \le z' \le L \end{cases} \tag{34}$$

where

$$U(z) = e^{-ikz} + \frac{k - k_1}{k + k_1} e^{-2ikL+ikz}$$

$$V(z) = e^{ikz'} + \frac{k - k_0}{k + k_0} e^{-ikz'} \tag{35}$$

These expressions remain valid in a number of special cases, for instance when $k = k_0$, or $k = k_1$, and even in the limit $L \to \infty$, provided one assumes that Im $k < 0$. In the special case when the plane $z = L$ is a perfect conductor the continuity of the first derivative at $z = L$ no longer applies and the only condition to be imposed at $z = L$ is $G(L,z') = 0$. One thus obtains in this special case

$$G(z,z') = \frac{1}{2ik\,\nu(L)} \begin{cases} u(z)v(z') & 0 \le z' \le z \le L \\ u(z')v(z) & 0 \le z \le z' \le L \end{cases} \tag{36}$$

where

$$u(z) = e^{ik(L-z)} - e^{ik(L-z)}$$

$$v(z) = e^{ikz} + \frac{k - k_0}{k + k_0} e^{-ikz} \tag{37}$$

5. SMALL CORRELATION LENGTH CASE

We are now ready to examine the solution of the equations developed in Section 3 for the field inside the randomly stratified medium. The correlation function (13) must be, of course, explicitly known and in much of the work previously done in this problem the most popular models adopted were the exponential $\sigma^2\exp(-|z-z'|/\ell)$ and the Gaussian $\sigma^2\exp(-\pi(z-z')^2/\ell^2)$. In either of these models, σ^2 stands for the (non-dimensional) covariance of the fluctuating wavenumber, and ℓ for the correlation length, generally defined as

$$\ell = \sigma^{-2} \int dz\ c(z) \tag{38}$$

Whatever the model, it can be shown that a solution of the system of equations (20) and (26) is equivalent to a partial summation of an infinite number of (renormalized) perturbation theory terms [6]. Evidently, orders of truncation higher than N = 3, have solutions equivalent to additional summations of infinite numbers of such terms. For an efficient use of the formulation presented here, it is therefore important to devise non-perturbative methods for solving the truncated equations. Attempts to solve such a system of equations have already been made, but we shall not pursue this approach here. We shall instead endeavor to present a basically model independent case, when the correlation length is so small that, whatever the model, one can approximate the correlation function by a Dirac function:

$$c(z - z') = \frac{2\eta}{k} \delta(z - z') \qquad (0 \le z,z' \le L) \tag{39}$$

where

$$\eta = \frac{1}{2} \sigma^2 k \ell \ . \tag{40}$$

In this case, equations (20) and (26) become

$$\left(\frac{d^2}{dz^2} + k^2\right) \phi_0(z) = - 2\eta k\, \phi_1(z;z) \tag{41}$$

and

$$\left(\frac{\partial^2}{\partial z^2} + k^2\right) \phi_1(z;z') = - k^2\, \phi_0(z')\, \delta(z - z')$$

$$- 2\eta k^3 \left[G(z,z)\, \phi_1(z;z') + G(z,z')\, \phi_1(z';z)\right] \tag{42}$$

respectively. Let us now note that, if $k\ell \ll 1$, as it should in order for the approximation (39) to be justified, the parameter η is also generally small, unless σ^2 is much larger than unity. If η is a small parameter, the solution of the system (41-42) by perturbation theory is entirely justified. It should be noticed though, that if ϕ_0 is sought as a series in powers of η, and this series is truncated and approximated by a polynomial (say, of degree 2 in η) then, because of (41), the corresponding series for ϕ_1 should be approximated by a polynomial of degree 1. Should powers of η higher than the second be considered non-negligible, consistency requires examination of orders of truncation higher than N = 3. Conversely, if powers of η higher than the first are negligible, than clearly the N = 3 order of truncation still does not reduce to the order N = 2, because, even if ϕ_1 is of zero order in η, ϕ_2 is still not negligible [as can be directly seen from (25)]. However, the contribution of ϕ_2 to the second moment of the field is of order η^2. To summarize, if one considers only terms of order η in ϕ_0, one may ignore ϕ_2, and further truncate equation (42) to read

$$\left(\frac{\partial^2}{\partial z^2} + k^2\right) \phi_1(z;z') \simeq - k^2 \phi_0(z') \delta(z - z') \tag{43}$$

Consequently,

$$\phi_1(z;z') \simeq k^2 G(z,z') \phi_0(z') \tag{44}$$

so that (41) becomes

$$\left(\frac{d^2}{dz^2} + k^2\right) \phi_0(z) = - 2\eta k^3 G(z,z) \phi_0(z) \tag{45}$$

or, explicitly,

$$\left(\frac{d^2}{dz^2} + k^2\right) \phi_0(z) = i\eta k^2 \left[A + \tfrac{1}{2}(Be^{-2ikz} + Ce^{2ikz})\right] \phi_0(z) \tag{46}$$

This equation, resulting from our assumptions, is a reduced variant of the integro-differential equation proposed [8,9] for the averaged field, which is itself a modification of the Dyson equation. We shall note first that if one attempts to solve (46) by using a straightforward perturbation theory there appear "mixed-secular terms", the presence of which restricts the values of L one might want to consider. For this reason, we shall use a two-scale perturbation theory [10], and accordingly introduce the new variable $Z = \eta z$, temporarily considered independent of z. If one then writes the solution in approximate form as

$$\phi_0(z,Z) = \phi_0^{(0)}(z,Z) + \eta \phi_0^{(1)}(z,Z) \tag{47}$$

and substitute in (46)

$$\frac{d^2}{dz^2} = \frac{\partial^2}{\partial z^2} + 2\eta \frac{\partial^2}{\partial z \partial Z} + \eta^2 \frac{\partial^2}{\partial Z^2} \quad , \tag{48}$$

it follows that $\phi_0^{(0)}(z,Z)$ and $\phi_1^{(1)}(z,Z)$ satisfy the equations

$$\left(\frac{\partial^2}{\partial z^2} + k^2\right) \phi_0^{(0)}(z,Z) = 0 \tag{49}$$

$$\left(\frac{\partial^2}{\partial z^2} + k^2\right) \phi_0^{(1)}(z,Z) = \left\{ik^2 \left[A + \tfrac{1}{2}(Be^{-2ikz} + Ce^{2ikz}) - 2 \frac{\partial^2}{\partial z \partial Z}\right\} \phi_0^{(0)}(z,Z) \right. . \tag{50}$$

Equation (49) is readily integrated, yielding

$$\phi_0^{(0)}(z,Z) = \Gamma_+(Z) e^{-ikz} + \Gamma_-(Z) e^{ikz} \quad . \tag{51}$$

Substitution of (51) into (50) shows that the mixed-secular terms disappear, provided

$$\begin{aligned} kA \Gamma_+(Z) + \tfrac{1}{2} kB\Gamma_-(Z) + 2 \frac{\partial}{\partial Z} \Gamma_+(Z) = 0 \\ kA \Gamma_-(Z) + \tfrac{1}{2} kC\Gamma_+(Z) - 2 \frac{\partial}{\partial Z} \Gamma_-(Z) = 0 \end{aligned} \tag{52}$$

Hence,

$$\begin{aligned} \Gamma_+(Z) &= a_+ e^{-\frac{1}{2}k\mu Z} + a_- e^{\frac{1}{2}k\mu Z} \\ \Gamma_-(Z) &= \frac{2}{B} a_+ (A - \mu) e^{-\frac{1}{2}k\mu Z} - \frac{2}{B} a_- (A + \mu) e^{\frac{1}{2}k\mu Z} \end{aligned} \tag{53}$$

where

$$\mu^2 = A^2 - BC/4 = (1 + 3\Lambda^2)/4 \ . \tag{54}$$

It follows then from (50) that

$$\phi_0^{(1)}(z,Z) = -\frac{i}{16}\left[B\,\Gamma_+(Z)\,e^{-3ikz} + C\Gamma_-(Z)\,e^{3ikz}\right] \tag{55}$$

so that ϕ_0 is completely determined to first order in η. Moreover, ϕ_1 is also determined from (44) or, more precisely, from

$$\phi_1(z;z') = G(z,z')\,\phi_0^{(0)}(z') \ . \tag{56}$$

Finally, the constants $a_\pm$ must be determined as well as the quantities $\langle R\rangle$ and $\langle T\rangle$, which characterize the average field. Continuity of the average field at $z = 0$ and $z = L$ requires that

$$\begin{aligned} 1 + \langle R\rangle &= a_+c(\mu) + a_-c(-\mu) \\ 1 - \langle R\rangle &= \frac{k}{k_0}\left[a_+d(\mu) + a_-d(-\mu)\right] \end{aligned} \tag{57}$$

and

$$\begin{aligned} a_+c(\mu,L) + a_-c(-\mu,L) &= \langle T\rangle\, e^{-ik_1L} \\ \frac{k}{k_1}\left[a_+d(\mu,L) + a_-d(-\mu,L)\right] &= \langle T\rangle\, e^{-ik_1L} \end{aligned} \tag{58}$$

where

$$\begin{aligned} c(\mu,L) = {} & e^{-ikL(1 - \frac{i}{2}\eta\mu)} - \frac{2}{B}(A - \mu)\, e^{ikL(1 + \frac{i}{2}\eta\mu)} \\ & - \frac{i\eta B}{16}\, e^{-3ikL(1 - \frac{i}{6}\eta\mu)} + \frac{i\eta C}{8B}(A - \mu)\, e^{3ikL(1 + \frac{i}{6}\eta\mu)} \end{aligned} \tag{59}$$

$$\begin{aligned} d(\mu,L) = {} & (1 - \tfrac{i}{2}\eta\mu)\, e^{-ikL(1 - \frac{i}{2}\eta\mu)} + \frac{2}{B}(A - \mu)(1 + \tfrac{i}{2}\eta\mu) e^{ikL(1 + \frac{i}{2}\eta\mu)} \\ & - \frac{3i\eta B}{16}\, e^{-3ikL(1 - \frac{i}{6}\eta\mu)} - \frac{3i\eta C}{8B}(A - \mu)\, e^{3ikL(1 - \frac{i}{6}\eta\mu)} \end{aligned} \tag{60}$$

and

$$\begin{aligned} c(\mu) &= c(\mu,0) \\ d(\mu) &= d(\mu,0) \end{aligned} \tag{61}$$

It follows then from (57) and (58) that

$$a_\pm = \pm\frac{2}{\Delta}\left[c(\mp\mu,L) - \frac{k}{k_1}\,d(\mp\mu,L)\right]$$

with

$$\begin{aligned} \Delta = {} & \left[c(\mu) + \frac{k}{k_0}\,d(\mu)\right]\left[c(-\mu,L) - \frac{k}{k_1}\,d(-\mu,L)\right] \\ & - \left[c(-\mu) + \frac{k}{k_0}\,d(-\mu)\right]\left[c(\mu,L) - \frac{k}{k_1}\,d(\mu,L)\right] \ . \end{aligned} \tag{63}$$

The energy fluxes inside the laminar structure can now be directly calculated and, naturally, they emerge as functions of z:

$$S^{coh}(z) = -\,\mathrm{Im}\left[\phi_0^*(z)\,\frac{d}{dz}\,\phi_0(z)\right] \tag{64}$$

and

$$\begin{aligned} S^{incoh}(z) &= -\,\mathrm{Im}\,\frac{2\eta}{k}\int_0^L dz'\,\phi_1^*(z;z')\,\frac{\partial}{\partial z}\,\phi_1(z;z') \\ &= -\,\mathrm{Im}\,\frac{2\eta}{k}\int_0^L dz'\,\left|\phi_1^{(0)}(z')\right|^2 G^*(z,z')\,\frac{\partial}{\partial z}\,G(z;z') \end{aligned} \tag{65}$$

By virtue of the continuity conditions at z = 0,L, imposed on both the average field and the Green's function,

$$S_-^{coh} = S^{coh}(0)$$

$$S_L^{coh} = S^{coh}(L) \ , \tag{66}$$

and likewise

$$S_-^{incoh} = S^{incoh}(0)$$

$$S_L^{incoh} = S^{incoh}(L) \ . \tag{67}$$

In the absence of absorption in the medium (i.e. when k is real),

$$S_-^{tot} = S_L^{tot} = S^{tot} \ , \tag{68}$$

which implies that S^{tot} should be independent of z, a direct test of the conservation of energy. In the presence of absorption, the difference $[S^{tot}(0) - S^{tot}(L)]/k_0$ clearly represents the absorptivity, the fraction of incoming energy absorbed in the medium.

6. NUMERICAL ILLUSTRATIONS

We present here some sample numerical calculations of the energy fluxes for a randomly stratified slab of width L = 20 cm, with a chosen correlation length of 2 mm. The half-space $z < 0$ is assumed to be vacuum, while the half-space at $z > L$ is assumed filled with a uniform dielectric with $\epsilon = 81\epsilon_0$. The average wave dielectric constant in the slab may be either complex, like $\epsilon = (3.2\text{-}i0.16)\epsilon_0$, or real: $\epsilon = 3.2\epsilon_0$. In the latter case, $k\ell = 0.075$ at a frequency of 1 GHz, such that $\eta = 0.037\sigma^2$ is a reasonably small quantity even for values of σ^2 somewhat larger than unity. Alternatively, if, say, $\sigma^2 < 0.5$, η remains small up to frequencies of order 10 GHz or so. A small imaginary part of ϵ does not appreciably affect these estimates.

In Fig. 2a we show the coherent, incoherent and total energy fluxes, in units of the incoming energy flux, k_0, for all values of z. As expected, the coherent energy flux decreases monotonically inside the slab, while the incoherent one increases from negative values (indicating transport of energy towards the interface z = 0), to positive values. In this case it has been assumed that Im k = 0, and the total energy flux results independent of z. In Fig. 2b it is seen that most of the energy lost through absorption is deducted from the coherent component.

In Fig. 3 we give the emissivity and transmissivity as functions of frequency, both when Im k = 0 and for Im k < 0. The characteristic oscillatory behavior is seen to taper off with increasing frequency in the presence of absorption, but to persist, and somewhat modify its aspect in the case of a pure dielectric.

Finally, in Fig. 4 we show the emissivity and transmissivity for two values of σ^2. As it can be seen, increasing σ^2 affects the emissivity more than the transmissivity, particularly at higher frequencies.

The emissivity curves in Figs. 2 and 3, for σ^2= 0.1 and 0.2, respectively, are directly comparable with those calculated by using a different approach in [11], which was also the source for the numerical parameters chosen here for purposes of illustration. As far as we can ascertain, our results are in good agreement with those given in [11].

7. CONCLUDING REMARKS

The case study examined in the preceding Section is, evidently, only one of a variety of cases to which the derived theoretical results may be applied. Moreover, except for additional but rather straightforward computational efforts, improved approximations can be developed to enlarge the domain of applicability.

Extension of the approach presented in this work can also be made in other directions, e.g., by attempting to solve the equations of Section 2 without assuming a necessarily small correlation length, or by considering the case of oblique incidence, three-dimensional randomness, polarization effects, etc. Work on some of these aspects of the problem is currently in progress.

ACKNOWLEDGMENT

This work was performed under the McDonnell Douglas Independent Research and Development program.

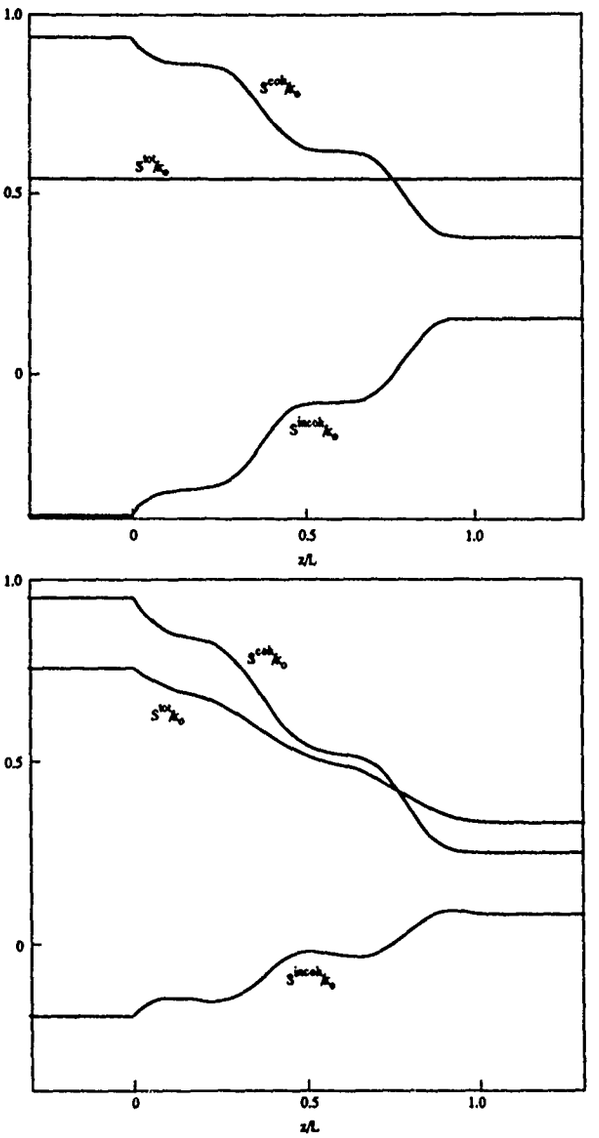


Figure 2. Coherent, incoherent and total energy fluxes for a medium with L = 20 cm, ℓ = 2 mm, σ^2 = 1, f = 1 GHz and (a) $\varepsilon = 3.2\varepsilon_0$; (b) $\varepsilon = (3.2 - i0.16)\varepsilon_0$.

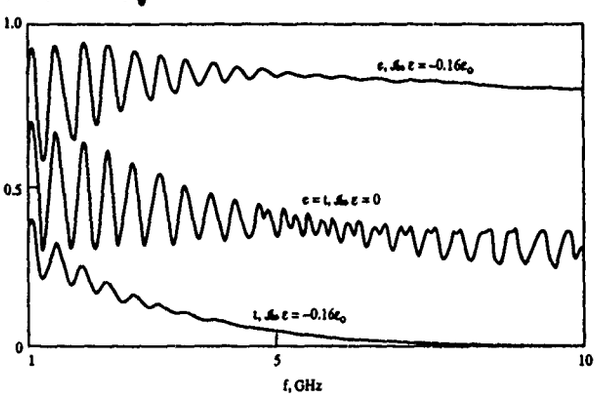


Figure 3. Emissivity and transmissivity for the same parameters as in Fig. 1, except that $\sigma^2 = 0.1$.

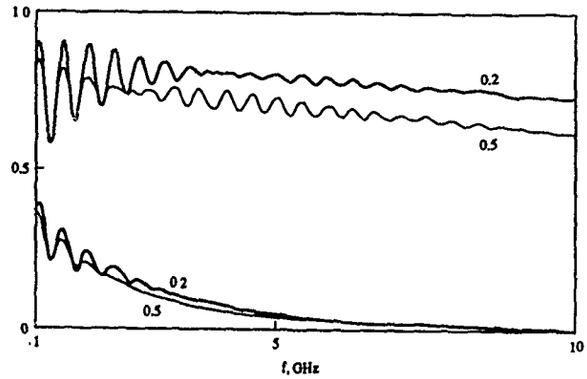


Figure 4. Emissivity and transmissivity for the same parameters as in Fig. 1, except that $\epsilon = (3.2 - i0.16)\epsilon_0$ for all curves and $\sigma^2 = 0.2$ and 0.5.

REFERENCES

1. Tatarskii, V. I., Wave Propagation in a Turbulent Medium, McGraw-Hill, 1961.
2. Ishimaru, A., Wave Propagation and Scattering in Random Media, Academic Press, 1978.
3. Frisch, U., "Wave propagation in random media," Probabilistic Methods in Applied Mathematics, A. T. Barucha-Reid, ed., 76-198, Academic Press, 1968.
4. Tsang, L., J. A. Kong, and R. T. Shin, Theory of Microwave Remote Sensing, Wiley, 1985.
5. Ulaby, F. T., R. K. Moore, and A. K. Fung, Microwave Remote Sensing, Artech House, 1986.
6. Eftimiu, C., "Wave propagation in random media: Wiener-Hermite expansion approach," J. EM Waves and Appl., 1990.
7. Rosenbaum, S., "On energy conserving formulations in a randomly fluctuating medium,"Proceedings of the Symposium on Turbulence of Fluids and Plasmas, Polytech. Inst. of Brooklyn, N.Y., 1968.
8. Keller, J. B., "Wave propagation in random media, "Proceedings of Symposium on Applied Mathematics, 13, 227-246, AMS Providence, R.I., 1960.
9. Meecham, W. C., "On radiation in a randomly inhomogeneous medium, "Space Technology Labs, Rept. BSD-TR-G1-36, 1961.
10. Kevorkian, J. and J. D. Cole, Perturbation Methods in Applied Mathematics, Springer-Verlag, 1981, (chapter 3).
11. Tsang, L. and A. U. Kong, "Microwave remote sensing of a two-layered random medium," IEEE Trans., vol AP-24, 283-288, 1976.

THE NASA RADIOWAVE PROPAGATION PROGRAM

FARAMAZ DAVARIAN

Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Drive
Pasadena, California 91109, USA

Introduction

The objectives of the NASA radiowave Propagation Program are to enable new satellite communication applications and to enhance existing satellite communication networks. These objectives are achieved by supporting radiowave propagation studies and disseminating the study results in a timely fashion.

Over the years, the Propagation Program has made sound contributions to the art of satellite communications. Studies initiated by this program in the 1980s enabled the infant concept of conducting mobile communications via satellite to reach a state of relative maturity in 1990. Also in the 1980s, the program supported the satellite communications community by publishing and revising two handbooks dealing with radiowave propagation effects for frequencies below and above 10 GHz, respectively. Furthermore, the Propagation Program has served the international community through its support of the International Telecommunications Union. It has also made great contributions to science by supporting state-of-the-art work at universities.

Currently, the Propagation Program is focusing on the Advanced Communications Technology Satellite (ACTS) and its propagation needs. The paper gives an overview of this programs's involvement in the ACTS project.

ACTS: A Switchboard in the Sky

The ACTS, now under construction and scheduled for launch in May 1992, is the main focus of NASA's Communications program. The prominent features of ACTS are its use of electronically hopping spotbeam antennas, onboard switching, and Ka-band transmission. The use of spot beam and switching technologies enables ACTS to provide multiple voice channels to very small aperture terminals (VSAT) and to do so in a single satellite hop. A baseband processor provides the switching. System access is based on time division multiple access (TDMA) and demand assignment by the network's master control station. An onboard microwave matrix switch is also provided, but is not discussed in this paper. Figure 1 shows the coverage map of the ACTS distributed network.

Two hopping beams with opposite linear polarizations provide coverage for most major cities. With a 1 ms TDMA frame time, a beam hops to many locations with variable dwell time to pick up the offered traffic. Each location is visited by its beam 1000 times a second. The uplink and downlink transmissions take place within two 1-GHz slots at about 30 and 20 GHz, respectively.

The baseband processor demodulates and stores the received signals. A baseband switch within the processor routes data to its destination, interconnecting the two uplink to the two downlink beams. Data is subsequently read out of output storage locations, modulated, and transmitted on a downlink beam. The modulation used is Minimum Shift Keying (MSK), and the transmission rate is either 110 or 27.5 Mbps on the uplink and 110 Mbps on the downlink.

Capacity allocation is made in 64 kbps increments; in every 1-ms frame period, a channel can transmit one 64-bit word occupying one time slot.

One of the most attractive properties of ACTS is its ability to combat fades induced by drizzle or rain. Via a clever scheme, ACTS is able to protect the entire distributed network against signal outages as large as 15 dB. The link is designed for a 5- dB clear weather margin, but terminals experiencing fade can be provided a further 10-dB fade protection.

To sense potential fades, ACTS is equipped with two propagation beacons at 20 and 27 GHz. These beacons are continually received and processed at each ground terminal for fade detection. Once the received beacon power falls below a certain threshold, the affected terminal sends a message to the network's master control station for further protection.

This protection is provided by means of burst reduction and data encoding. The burst rate is reduced four times resulting

THE NASA RADIOWAVE PROPAGATION PROGRAM

FARAMAZ DAVARIAN

Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Drive
Pasadena, California 91109, USA

Introduction

The objectives of the NASA radiowave Propagation Program are to enable new satellite communication applications and to enhance existing satellite communication networks. These objectives are achieved by supporting radiowave propagation studies and disseminating the study results in a timely fashion.

Over the years, the Propagation Program has made sound contributions to the art of satellite communications. Studies initiated by this program in the 1980s enabled the infant concept of conducting mobile communications via satellite to reach a state of relative maturity in 1990. Also in the 1980s, the program supported the satellite communications community by publishing and revising two handbooks dealing with radiowave propagation effects for frequencies below and above 10 GHz, respectively. Furthermore, the Propagation Program has served the international community through its support of the International Telecommunications Union. It has also made great contributions to science by supporting state-of-the-art work at universities.

Currently, the Propagation Program is focusing on the Advanced Communications Technology Satellite (ACTS) and its propagation needs. The paper gives an overview of this programs's involvement in the ACTS project.

ACTS: A Switchboard in the Sky

The ACTS, now under construction and scheduled for launch in May 1992, is the main focus of NASA's Communications program. The prominent features of ACTS are its use of electronically hopping spotbeam antennas, onboard switching, and Ka-band transmission. The use of spot beam and switching technologies enables ACTS to provide multiple voice channels to very small aperture terminals (VSAT) and to do so in a single satellite hop. A baseband processor provides the switching. System access is based on time division multiple access (TDMA) and demand assignment by the network's master control station. An onboard microwave matrix switch is also provided, but is not discussed in this paper. Figure 1 shows the coverage map of the ACTS distributed network.

Two hopping beams with opposite linear polarizations provide coverage for most major cities. With a 1 ms TDMA frame time, a beam hops to many locations with variable dwell time to pick up the offered traffic. Each location is visited by its beam 1000 times a second. The uplink and downlink transmissions take place within two 1-GHz slots at about 30 and 20 GHz, respectively.

The baseband processor demodulates and stores the received signals. A baseband switch within the processor routes data to its destination, interconnecting the two uplink to the two downlink beams. Data is subsequently read out of output storage locations, modulated, and transmitted on a downlink beam. The modulation used is Minimum Shift Keying (MSK), and the transmission rate is either 110 or 27.5 Mbps on the uplink and 110 Mbps on the downlink.

Capacity allocation is made in 64 kbps increments; in every 1-ms frame period, a channel can transmit one 64-bit word occupying one time slot.

One of the most attractive properties of ACTS is its ability to combat fades induced by drizzle or rain. Via a clever scheme, ACTS is able to protect the entire distributed network against signal outages as large as 15 dB. The link is designed for a 5- dB clear weather margin, but terminals experiencing fade can be provided a further 10-dB fade protection.

To sense potential fades, ACTS is equipped with two propagation beacons at 20 and 27 GHz. These beacons are continually received and processed at each ground terminal for fade detection. Once the received beacon power falls below a certain threshold, the affected terminal sends a message to the network's master control station for further protection.

This protection is provided by means of burst reduction and data encoding. The burst rate is reduced four times resulting

Note that these climate zones correspond to the Global rain model as originally introduced by Robert Crane of Dartmouth College.

Workshop participants took note of the short life time of the ACTS satellite (2 to 3 years) and its approaching launch date (mid 1992). The time-critical nature of the ACTS propagation studies was recognized, and it was strongly suggested that planning and terminal development must start as soon as possible. In response to the Workshop recommendations, a two-phase plan was put together. Phase I consisted of a terminal prototype development effort, and Phase II consisted of an experimental terminal construction effort.

Propagation Terminal Configuration

A block diagram of the ACTS propagation terminal is given in Figure 3. The terminal consists of a single antenna, the RF portion, the IF portion, the receiver, the Data Acquisition System, and the radiometer [2]. A single 1-meter antenna was selected to reduce the cost and size of the terminal. The antenna can accommodate dual-frequency and dual-polarization reception. The receiver and radiometer are dual-frequency. The data acquisition system collects and temporarily stores beacon and radiometric data as well as meteorological and other relevant information. The stored data will be periodically dumped on tape or optical disk for transportation.

The beacons at 20 and 27 GHz have EIRPs of 19 and 16 dBW, respectively. The 20-GHz beacon is also used for low bit rate telemetry data transmission. Therefore, a modulation loss is associated with the 20-GHz beacon receiver. Considering a 1-meter antenna and a receiver noise temperature of about 1750 K in clear weather, the receiver dynamic range (margin over threshold) for both frequencies is above 20 dB.

Data Analysis

As mentioned earlier, propagation data will be collected at eight to ten sites across the United States. The collected data will be provided to an investigator, who is likely to be from a university, for data reduction and analysis. It is expected that many centers will participate in the analysis of the reduced data. Therefore, efforts such as model development and testing, algorithm development, etc., will be performed by more than one center.

Financial Support

NASA is providing financial support for the development of the propagation terminals, and their installation and maintenance. NASA will also fund the data analysis effort performed by the principal investigator at a university. Other centers participating in the data analysis part of this effort will finance their own expenses.

Conclusion

The next Workshop on ACTS Propagation Studies will take place in the Los Angeles area on November 27 and 28, 1990. It is hoped that a complete design of the propagation terminal will be available for presentation. Workshop participants will have an opportunity to critique the terminal design and provide comments. The construction of a terminal prototype will proceed shortly after the Workshop.

The current plan for ACTS propagation studies will be revisited at the Workshop. Issues such as the site locations of the terminals, methodology for data processing and model development, fade mitigation algorithms, etc., will be discussed. Although the study plan is not expected to be finalized at this meeting, it is hoped that it will reach a degree of maturity that can allow clear guidelines for the campaign.

The experimental (data collection) phase of the ACTS propagation campaign is expected to begin in mid 1992 and continue for at least 2 years. The results and findings of this campaign will be disseminated in a timely fashion to the ACTS and other technical communities.

Acknowledgement

The research conducted in this publication was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

The author would like to acknowledge the contributions of Dr. D. Chakraborty of JPL and several colleagues at Virginia Polytechnic Institute.

References

[1] F. Davarian, editor, Presentations of the First ACTS Propagation Studies Workshop (APSW I), JPL D-6918, December 15, 1989.

[2] F. Davarian, editor, Proceedings of NAPEX XIV, JPL Publication 90-27, pp. 178-202, July 1, 1990.

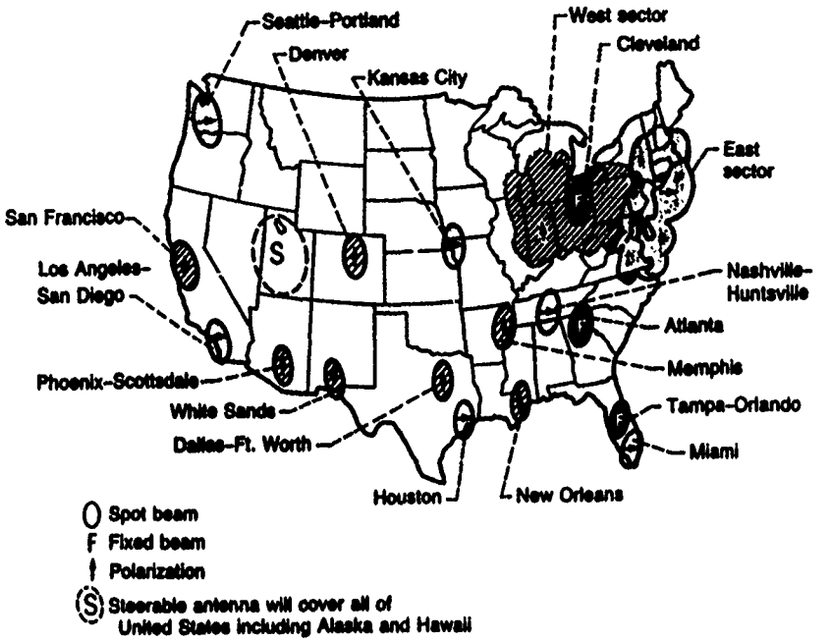


Figure 1. ACTS Multibeam Antenna Coverage

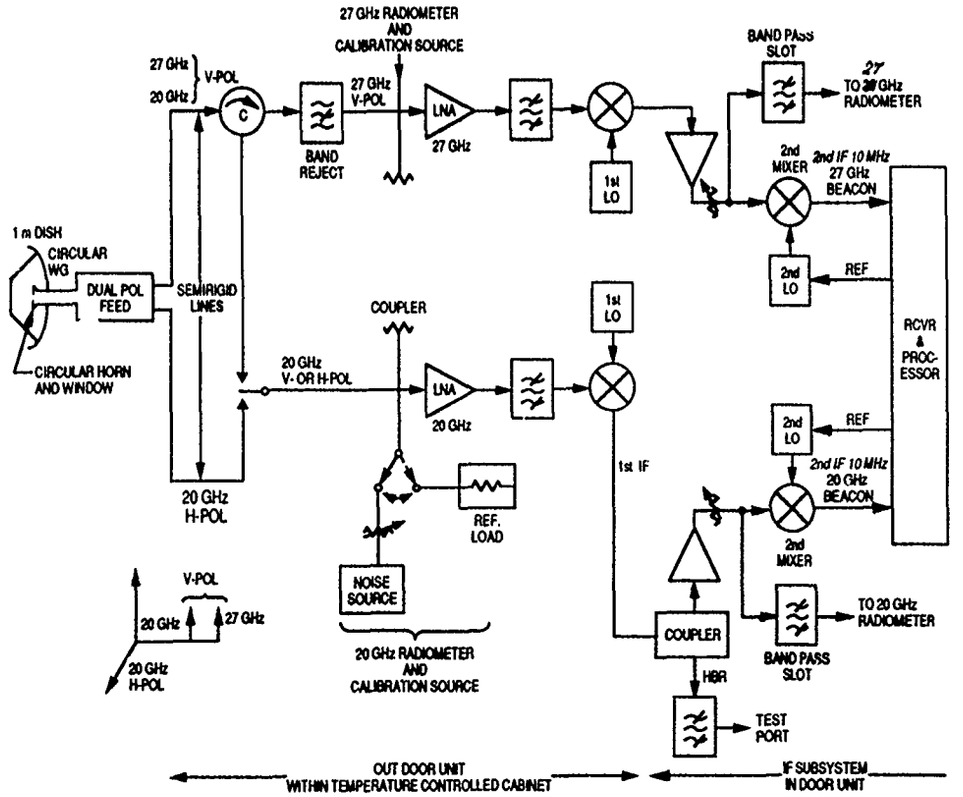


Figure 3. ACTS Prototype Terminal Initial Design

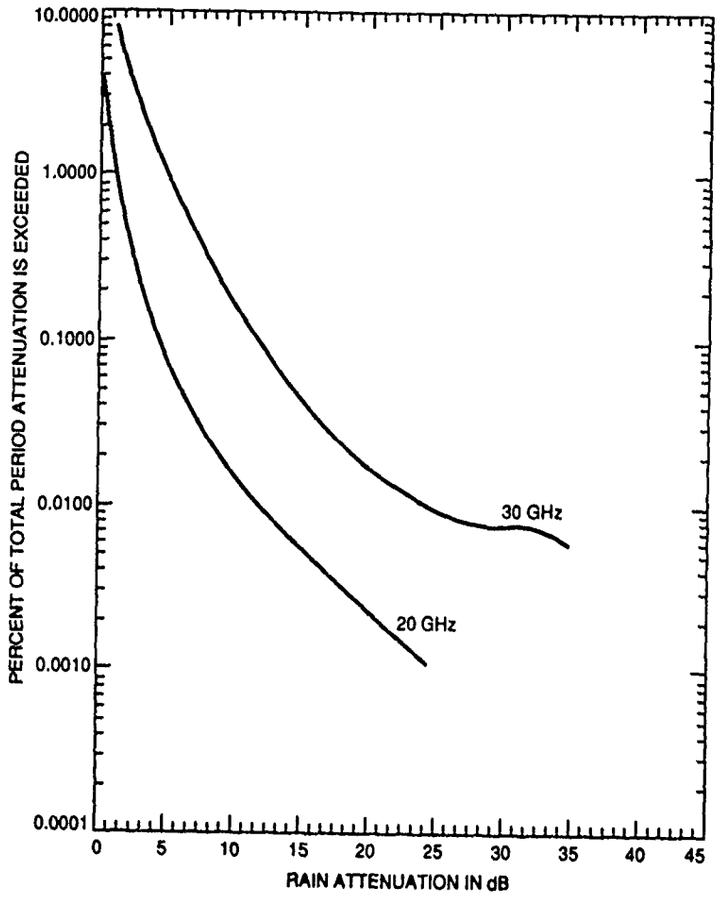


Figure 2. Cumulative Attenuation Statistics at Blacksburg, Virginia

DISCUSSION

J.H. RICHTER

Do you forsee use of the 90-100GHz band for satellite communications applications ?

AUTHOR'S REPLY

Yes. For systems with no real time data transmission requirements and in areas of the world in which it does not rain very much, such as Crete in Greece and southern California in the US.

G. TACCONI

What is the moments order relevant to the statistical significance of attenuation ?

What is the stability of the estimate of such statistical significance ?

AUTHOR'S REPLY

The longer an observation is made, the more stable are the parameter estimates derived from it. In the ACTS experiment, to ensure good results, the observation period is recommended to be at least 3 years.

LOGICIEL DE CARTOGRAPHIE RADIOELECTRIQUE INTERACTIF

par
M. GERVAISE, C. ARMAND, G. DEBERDT, A. POUCIN
THOMSON-CSF
46-47, quai A. Le Gallo
B.P. 407
92103 Boulogne
France

(English version available on request from the authors)

RESUME

La sophistication des techniques et produits de radiocommunications et la durée nécessaire à leur développement entraînent un besoin de démonstration et de validation de leurs architectures et performances.
Ce besoin est ressenti aussi bien par les concepteurs des systèmes que par les clients.

Dans ce cadre, un logiciel a été développé pour dimensionner le déploiement des réseaux hertziens.

Celui-ci fait l'objet du présent exposé où l'on montre la diversité de ses applications et l'intérêt de son emploi dans l'élaboration des systèmes complexes de radiocommunications.

I. INTRODUCTION

Dans l'élaboration des systèmes de radiocommunications développés par la branche B.C.C. de THOMSON-CSF, il est nécessaire de prendre en compte tous les paramètres qui interviennent dans l'éventail des situations opérationnelles prévisibles.

Ces paramètres comprennent non seulement les caractéristiques techniques des équipements radio, mais aussi leur implantation dans un environnement géographique déterminé, leurs concepts d'emploi sur le champ de bataille incluant une éventuelle dégradation des performances, l'effet plus ou moins favorable des conditions de propagation, l'existence possible de brouillage (E.C.M.).

La multiplicité de ces combinaisons aussi bien que la diversité des localisations à envisager pour l'implantation des matériels radio, compte tenu du relief et des obstacles présents sur le terrain, ont conduit B.C.C. à se doter d'un moyen de prévision puissant et adapté avec l'outil de cartographie radioélectrique interactif I.R.S. (abréviation pour Interactive Radiomapping Software).

II. PRESENTATION GENERALE ET APPLICATIONS

En développant I.R.S., THOMSON-CSF a voulu créer un outil efficace capable de déterminer les performances opérationnelles d'un réseau hertzien et de servir d'aide à la décision en vue du choix de la meilleure solution face à des contraintes multiples.

Cet objectif s'est trouvé atteint par la réalisation d'un ensemble informatique modulaire :

- orienté vers la gestion scientifique des données qui concourent à la définition du déploiement d'un réseau dans son environnement,
- caractérisé par une grande souplesse d'utilisation.

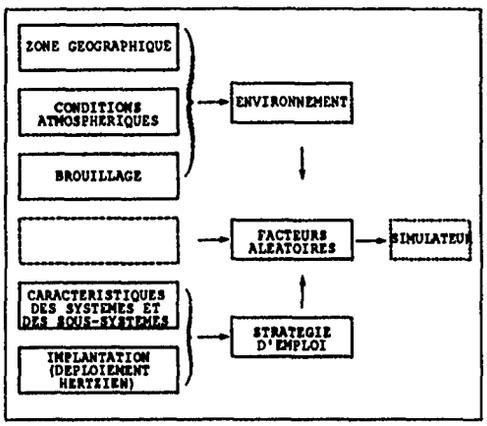
En effet les données, dans leur plus grande partie, peuvent être introduites ou modifiées au cours de l'exécution du programme selon des procédures de dialogue interactif.

L'environnement informatique nécessaire comprend une station de travail, ayant une puissance de 12,5 Mips et munie d'un accélérateur graphique, associée à des moyens d'édition des résultats tels que console graphique et imprimante couleur haute résolution.

La mise en œuvre du logiciel commence avec la définition du théâtre des opérations (figure 1) :

- zone géographique,
- conditions atmosphériques,
- environnement électromagnétique,

et l'implantation des centres nodaux dans l'avant-projet.



ARCHITECTURE GLOBALE DE LA SIMULATION
FIGURE 1

Le logiciel utilise en effet, pour les calculs de couverture radioélectrique et d'étude de connexité des centres nodaux, un terrain numérisé. Celui-ci donne, en chaque point d'un maillage disposé régulièrement en longitude et latitude, une altitude et une information servant à indiquer la présence d'un obstacle d'une hauteur supérieure à dix mètres et considéré comme nuisible à la propagation hertzienne.

La base de données exploitées peut être complétée ou étendue sur certaines zones sensibles, en utilisant des cartes régionales numérisées plus précises ou des relevés topographiques qui peuvent être faits localement.

Un point important est l'exactitude du terrain numérique utilisé sur lequel repose la validité des résultats ultérieurs.

Au cours de cette étape, une zone de terrain peut être visualisée sous forme de :

- courbes de niveau,
- relief vu à partir d'un point donné.

Ces représentations autorisent la correction des erreurs introduites éventuellement dans les données relatives au terrain et au déploiement. Elles permettent d'affiner le choix des sites initiaux.

Une fois défini le théâtre des opérations et retenus les emplacements envisagés pour les centres nodaux, une nouvelle étape consiste à prendre en compte, grâce à des modèles physiques appropriés, les caractéristiques des équipements et du canal de propagation.

Ceci permet de déduire les performances des matériels hertziens dans leur environnement. Elles sont établies à partir d'analyses faites :

- sur la connexité des différents centres nodaux,
- sur les couvertures radioélectriques calculées à partir d'un ou plusieurs sites.

Pour ce faire, on utilise :

- un gestionnaire de "Base de données" de terrain numérisé afin d'extraire des profils de liaisons,
- un algorithme d'évaluation de l'atténuation de propagation hertzienne I.T.S. 1.2.2., (cf. [1], [2], [3], [4]),
- la méthode BARTON qui donne la qualité (par exemple sous forme de taux d'erreurs binaires) d'une liaison hertzienne, (cf. [5], [6], [7], [8]),
- un ensemble de modules issus de la théorie des graphes, (cf. [9], [10], [11]),
- un logiciel graphique pour la restitution "2D", "3D", "4D".

Ces cinq composantes sont décrites dans la partie concernant la structure globale du simulateur.

L'intérêt de cette phase est :

- d'optimiser l'emplacement des sites,
- d'examiner si les caractéristiques globales des équipements sont satisfaisantes en termes de puissance émise, fréquence, hauteur d'aérien, etc.

Les performances systèmes sont traduites par des histogrammes, des cartes radioélectriques, des graphes ..., ce qui permet d'apprécier si la qualité du déploiement correspond à l'objectif de couverture globale.
Le cas échéant, un ou plusieurs paramètres peuvent être réajustés.

Une dernière étape dans l'élaboration du projet vise à étudier la vulnérabilité du réseau maillé et sa capacité à acheminer les flux d'informations.
L'outil de simulation permet d'observer la dégradation apportée, sur la couverture et le maillage du réseau, par :

- la suppression d'un ou plusieurs sites,
- la présence de conditions climatiques défavorables à la propagation,
- la présence de contre-mesures électroniques (E.C.M.).

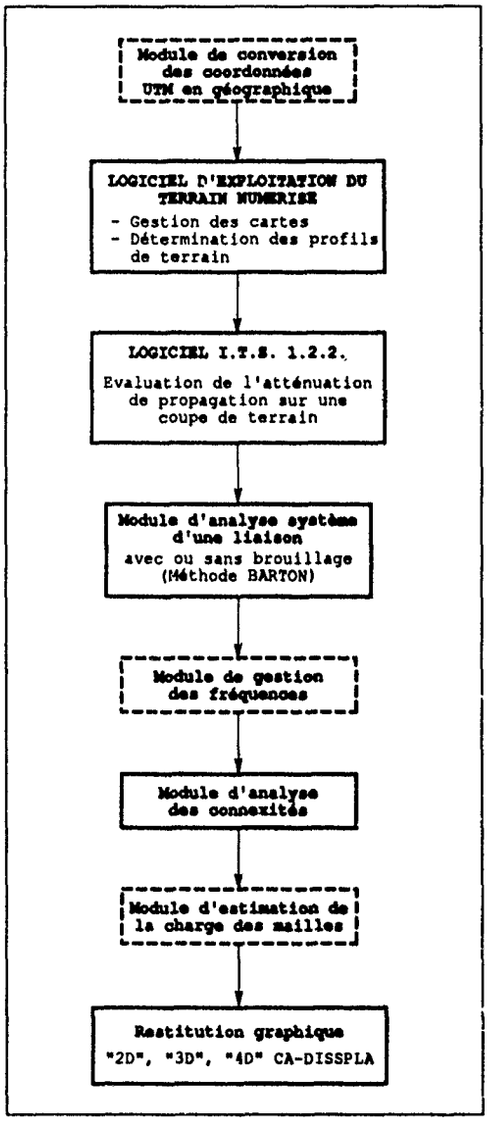
Il est également possible, par le biais du module basé sur la théorie des graphes, de calculer la charge des mailles, d'isoler les groupes d'abonnés indépendants, de déterminer d'autres possibilités de raccordement.

Le logiciel de cartographie radioélectrique a été développé initialement pour étudier et dimensionner le déploiement d'un réseau radio, mais il peut être également utilisé pour :

- le "design" des équipements :
  - étude de la fiabilité des communications,
  - comportement des équipements dans différents environnements,
- prédiction d'une portée radioélectrique,
- analyse d'une liaison point à point.

III. STRUCTURE ET MISE EN ŒUVRE

L'enchaînement des cinq composantes de I.R.S. est représenté sur la figure suivante.



STRUCTURE GLOBALE
FIGURE 2

III.1 Le gestionnaire de terrain numérisé

Celui-ci détermine une coupe de terrain échantillonnée au pas de 9" x 12" d'arc (250 m x 250 m).

Chaque point du terrain numérisé repéré par ses coordonnées (longitude, latitude) est caractérisé par :

- son altimétrie,
- son "sur-sol" (existence de végétation ou d'infrastructure).

Pour extraire les informations concernant un point numérisé, on utilise les trois types de fichiers suivants :

(1) Fichier catalogue des cartes,
(2) Fichiers "altimétries",
(3) Fichiers "sur-sol".

Le fichier catalogue permet de faire la correspondance entre les coordonnées d'un point et la carte qui le contient ; celle-ci est identifiée par un numéro.

Les données obtenues à partir de coupes successives sont utilisées pour établir des cartes géographiques et radioélectriques d'une région située au nord-est de la FRANCE.

III.2 Evaluation de l'affaiblissement de propagation

L'algorithme de propagation I.T.S. 1.2.2., résultant de travaux menés sur une période de vingt années aux U.S.A. par l'"Institute for Telecommunication Sciences", permet d'évaluer l'atténuation de propagation par ondes de sol ou diffusion troposphérique.

Celui-ci est basé sur la théorie électromagnétique et sur des analyses statistisques des caractéristiques du terrain et des mesures de propagation. Il prédit la valeur médiane de l'atténuation comme fonction de la distance, et estime les variations du signal dans le temps et l'espace.

Le modèle est considéré comme semi-empirique, puisqu'il combine des théories élémentaires avec des données expérimentales.

L'algorithme associé au modèle est conçu pour être utilisé entre 20 MHz et 20 GHz en traitant une grande variété de distances et de hauteurs d'antennes : il est parfaitement adapté pour analyser les situations où le terrain joue un rôle important.

Pour le calcul de l'atténuation, le modèle utilise des traitements théoriques relatifs à :

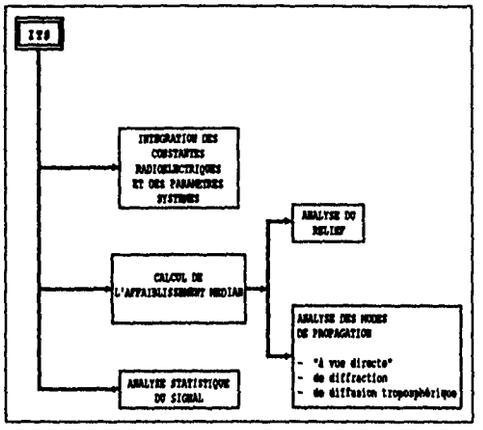
- la réflexion sur un terrain rugueux,
- la réfraction à travers une atmosphère standard.
- la diffraction autour de la terre et au-dessus d'obstacles à flancs raides,
- la diffusion troposphérique.

La mise en œuvre de l'algorithme est flexible selon deux modes généraux d'utilisation :

- calcul d'un affaiblissement médian pour une liaison donnée,
- analyse statistique de l'affaiblissement de propagation. Dans ce mode, l'algorithme donne une valeur d'affaiblissement en indiquant les degrés de confiance dans celle-ci (par exemple sous forme de : probabilité temporelle et probabilité de service).

Le modèle est "statistique", en effet le signal reçu fluctue à cause :

- des variations temporelles induites par des conditions atmosphériques changeantes,
- des variations spatiales dues aux différentes configurations de terrain.



ARCHITECTURE DE L'ALGORITHME I.T.S.
FIGURE 3

III.2.1 Principe général des calculs

Ce paragraphe décrit l'utilisation des différents paramètres globaux servant à calculer l'affaiblissement de propagation.

La valeur L_{cr} caractérisant l'affaiblissement de propagation est donc la somme de l'affaiblissement de propagation en espace libre et de l'affaiblissement de référence AREF entre antennes isotropes.

$$L_{cr} = L_{bf} + AREF \ [dB] \quad (1)$$

L'affaiblissement en espace libre est défini par :

$$L_{bf} = 32,5 + 20 \log_{10} f + 20 \log_{10} d \ [dB] \quad (2)$$

f : fréquence [MHz],
d : longueur de la liaison [km].

L'atténuation de référence AREF est calculée en utilisant des méthodes basées sur l'évaluation de différents modes de propagation relatifs à trois types de distance (figure 4):

- Liaison à vue directe

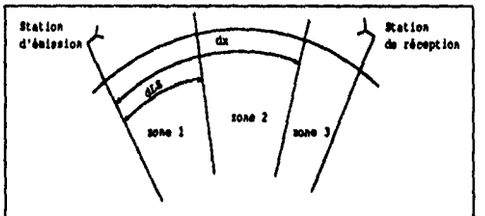
  Les formules relatives à l'optique géométrique (Théorie des 2 rayons) sont utilisées pour calculer l'atténuation.

- Propagation transhorizon

  • Juste au-delà de la zone de visibilité radioélectrique, la diffraction est le phénomène dominant intervenant dans la propagation.
  La méthode de prédiction calcule une valeur moyenne pondérée Ad résultant de l'estimation de la diffraction due à 2 obstacles à flancs raides et à un terrain rugueux (sphérique).

  • A des distances élevées au-delà de l'horizon radioélectrique, la diffusion troposphérique est le phénomène prépondérant. La méthode de prédiction pour ces distances est celle de Rice (1967) qui a été partiellement adoptée.

L'atténuation de référence AREF pour la propagation transhorizon est soit l'atténuation de diffraction Ad, soit l'atténuation de diffusion As. On choisit la plus petite des deux. La distance à laquelle la diffraction et la diffusion sont égales est définie par dX.



- zone 1 : Visibilité radioélectrique. Les lois de l'optique géométrique sont utilisées pour calculer l'affaiblissement de propagation.

- zone 2 : L'affaiblissement de propagation résultant est la somme des contributions relatives à la diffraction sur terre sphérique lisse et sur terre rugueuse comportant au maximum deux obstacles.
Dans cette zone il n'y a pas de visibilité radioélectrique entre l'émetteur et le récepteur.

- zone 3 : L'affaiblissement de propagation est relatif à la diffusion troposphérique, phénomène prépondérant à des distances élevées situées derrière l'horizon radioélectrique

MODES DE PROPAGATION
FIGURE 4

Pour fournir AREF comme fonction continue de la distance, les méthodes suivantes sont utilisées :

1) A "vue directe"

Quand les antennes d'émission et réception sont en visibilité

radioélectrique l'une de l'autre, la théorie des 2 rayons de l'optique géométrique s'applique pour calculer les atténuations A0 et A1 aux distances spécifiées d0 et d1 qui sont en deçà de la zone transhorizon.

La distance d0 est choisie pour approximer la plus grande distance à laquelle l'atténuation ne dépasse pas l'atténuation de propagation en espace libre.

La distance d1 est supérieure à d0, mais comprise dans des limites qui permettent d'appliquer la théorie de l'optique géométrique.

Les méthodes décrites aux alinéas suivants sont utilisées pour calculer l'atténuation de diffraction ALS à la distance dLS.

Les trois valeurs d'atténuation A0, A1 et ALS calculées aux distances d0, d1 et DLS respectivement sont utilisées pour déterminer les pentes k1 et k2 de la courbe "lissée" représentant AREF en fonction de la distance d (1 < d < dLS).

$$AREF = A0 + k1\ (d-d0) + k2\log_{10}\ (d/d0)\ [dB] \quad (3)$$

La distance à l'horizon sur une terre lisse dLS peut être plus grande que la distance à l'horizon dL sur une terre rugueuse.

La formule (3) peut être simplifiée par :

$$Ae = A0 - k1\ d0 - k2\ \log_{10}(d0)\ [dB] \quad (4)$$

$$AREF = Ae + k1\ d + k2\ \log_{10}(d)\ [dB] \quad (5)$$

Des méthodes détaillées sont données pour calculer k1 et k2.

2) Par diffraction

L'atténuation de diffraction est calculée en combinant respectivement des estimations Ak et Ar de la diffraction sur des obstacles à flancs raides (théorie de Fresnel-Kirchoff) et sur terre sphérique lisse (théorie de la diffraction développée par Vogler (1964)).

La théorie de la diffraction sur obstacles à flancs raides est utilisée pour estimer l'atténuation au-dessus d'une colline isolée ou d'une crête.

Dans l'application, l'atténuation de diffraction Ak est calculée comme si la liaison radio interceptait 2 crêtes isolées à flancs raides.

En général, pour un terrain irrégulier, l'atténuation de diffraction Ad est calculée comme pondération des deux estimations Ar et Ak.

$$Ad = (1 - W)\ Ak + W\ Ar\ [dB] \quad (6)$$

Le facteur de pondération, W, est déterminé empiriquement comme fonction de la fréquence radio et des paramètres relatifs au terrain.

L'atténuation de diffraction Ad est calculée aux distances d3 et d4 dans les zones de diffraction éloignées. Une ligne droite est tracée entre les points (A3, d3) et (A4, d4) et est définie par le point origine Aed (pour d = 0) et la pente md comme suit :

$$md = (A4 - A3)\ /\ (d4 - d3)\ [dB/km] \quad (7)$$

$$Aed = Af0 + A4 - (md \times d4)\ [dB] \quad (8)$$

Af0 est un facteur de "clutter".

L'atténuation de référence à toute distance d telle que dLS < d < dx est donnée par :

$$AREF = Ad = Aed + (md \times d)\ [dB] \quad (9)$$

3) Par diffusion troposphérique

Lorsque la distance d ou la distance angulaire θ est élevée, ($\theta = e + d/a$ radians),

d = longueur de la liaison,
a = rayon radio terrestre,
e = somme des angles de site émission et réception.

l'atténuation de diffusion As peut être inférieure à l'atténuation Ad relative au phénomène de diffraction.

Lorsque le produit distance d (km) par distance angulaire θ (rd) excède 0,5 , l'atténuation de diffusion As est calculée pour être comparée à Ad.

Pour certaines valeurs du produit θd, l'atténuation As est supposée avoir une dépendance linéaire sur la distance, cependant As est calculée à deux distances d5 et d6. Une ligne droite joint les points (A5, d5) et (A6, d6) et est définie par le point initial Aes (pour d = 0) et par la pente ms.

$$Aes = A5 - (ms \times d5) \ [dB] \qquad (10)$$

$$ms = (A6 - A5) \ / \ (d6 - d5) \ [dB] \qquad (11)$$

L'atténuation de référence AREF à toute distance d supérieure à dx, est donnée par :

$$AREF = As = Aes + (ms \times d) \ [dB] \qquad (12)$$

III.2.2 Analyse statistique du signal

Le modèle I.T.S. Longley-Rice dans sa version 1.2.2. comporte un module de calcul de la variabilité du signal, qui exclut a priori la variabilité à court terme, ou les petites variabilités.

Ce module restitue la valeur AVAR qui est la somme du terme AREF et de termes prenant en compte les variations statistiques du signal.

Il faut noter que l'affaiblissement global médian est donné par l'équation (1) ; avec statistique, il est déterminé par :

$$L_{cr} = L_{bf} + AVAR \qquad (13)$$

AVAR est déterminé à partir de trois paramètres qui définissent les **types de variabilité** :

- variabilité dans le temps : variation des médianes horaires locales sur un profil de liaison donné, au cours du temps,

- variabilité de lieu : variabilité dans les statistiques à long terme d'un profil de liaison à un autre,

- variabilité de situation : variation dans la variabilité de lieu, qui se produit de situation à situation.

L'ordre dans lequel les trois types de variabilité apparaissent, donne quatre façons de les combiner, qui ont toutes des utilisations légitimes pour tel ou tel type de service.
Nous appellerons ceci les **quatre modes de variabilité** :

- mode 0 : message seul
  Les variabilités de temps, de lieu et de situation sont combinées pour donner un niveau de confiance,

- mode 1 : mode individuel
  La fiabilité est donnée par la disponibilité au cours du temps, alors que la confiance est une combinaison de la variabilité de lieu et de situation,

- mode 2 : mode mobile
  La fiabilité est une combinaison de la variabilité de temps et de lieu, alors que la confiance est donnée par la variabilité de situation,

- mode 3 : mode "radiodiffusion"
  La fiabilité est donnée par l'assertion à deux paramètres indépendants d'une fraction au moins Q_t du temps, pour une fraction d'au moins Q_l des lieux (configuration de propagation), la confiance est donnée par la variabilité de situation.

De plus, il existe encore **deux options** possibles pour chaque mode de calcul :

- une première option, dans laquelle on élimine la variabilité de lieu, comme cela est naturel dans le cas où l'on traite un profil de liaison bien défini, dans le mode point à point,

- une seconde option, dans laquelle on élimine la variabilité de situation directe, comme cela est naturel lorsque l'on considère des problèmes d'interférence. Il est à noter qu'il peut encore rester une petite variabilité de situation résiduelle.

III.2.3 Paramètres globaux intégrés

Les principaux paramètres qui interviennent dans le calcul d'atténuation entre antennes isotropes, sont les suivants :

1) Paramètres systèmes

- Fréquence : 20 MHz < F < 20 GHz.
  La fréquence porteuse du signal est utilisée,
- Hauteurs d'antenne :
  0,5 m < H < 3000 m.
  représentent la hauteur entre les centres de radiation et le sol,
- Polarisation des antennes :
  verticale ou horizontale.

2) Paramètres d'environnement

- Géographique :
  - coupe de terrain entre les deux terminaux,
  - nature du terrain (sec, moyen, humide, verdoyant, mer ...),
    La nature du terrain est caractérisée par la permittivité relative et la conductivité électrique du sol.
- Climatique :
  - Nature du climat
    Sept types couvrant le globe peuvent être traités par le modèle (équatorial, continental subtropical, maritime subtropical, désertique, tempéré continental, maritime tempéré sur terre et sur mer),
  - Co-indice de réfraction
    Associé au type de climat, il sert à caractériser l'atmosphère et ses variations dans le temps.
- Electromagnétique :
  Les paramètres sont traités par la méthode BARTON.
- Statistique (exemple donné pour le mode individuel).
  Le niveau du signal reçu étant fluctuant, l'atténuation de propagation peut être estimée à l'aide d'une probabilité temporelle Qt et d'une probabilité de service Qc.
  - Probabilité temporelle

    Cela signifie que la valeur moyenne de l'affaiblissement prise en compte sur une heure sera inférieure à ce niveau pendant Qt % des heures de l'année.
  - Probabilité de service

    D'autre part, il est évident que les calculs d'affaiblissement sont effectués à partir de certains paramètres globaux (angles de site, distance équivalente, co-indice de réfraction atmosphérique ...) qui ne reflètent qu'imparfaitement l'ensemble des données réelles d'une liaison donnée.
    C'est à ce problème que correspond la notion de probabilité de service Qc.
    L'affaiblissement à Qc % est la valeur non dépassée (au pourcentage de temps considéré) par Qc % des liaisons ayant mêmes paramètres que ceux de la liaison étudiée.

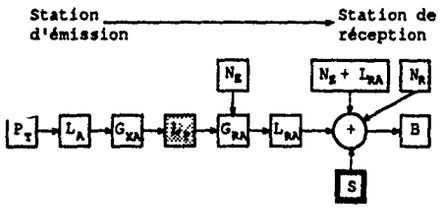
    Il faut noter que l'atténuation globale L_{cr}, calculée sans statistique, est une valeur médiane et correspond en fait à Qt = 50 % et Qc = 50 %.

III.3 Analyse système d'une liaison

La méthode BARTON permet d'estimer la qualité d'une liaison en quantifiant l'interaction entre :

- les caractéristiques des équipements radio,
- l'environnement électromagnétique (bruits galactique, atmosphérique, industriel, présence de contre-mesures),
- la propagation électromagnétique.

• **Sans brouillage**



P_T	Puissance transmise	(W)
L_A	Perte du coupleur antenne de la station émettrice	(dB)
G_{XA}	Gain d'antenne émission	(dB)
L_p	Atténuation de propagation	(dB)
G_{RA}	Gain d'antenne réception	(dB)
N_E	Facteur de bruit équivalent d'antenne	(dB)
L_{RA}	Perte du coupleur antenne de la station réceptrice	(dB)
N_R	Facteur de bruit intrinsèque au récepteur	(dB)
B	Bande passante du récepteur	(Hz)
S_E	Seuil de sensibilité du récepteur Type de modulation du signal	(dB)

• **Avec brouillage**

Prise en compte des paramètres caractérisant la menace

ANALYSE SYSTEME D'UNE LIAISON
FIGURE 5

III.3.1 Environnement électromagnétique

Certains paramètres décrivant cet environnement sont utilisés seulement lors de l'étude de vulnérabilité du réseau hertzien face à un brouilleur (de type S.O.J.) positionné à distance.

- La menace est caractérisée par :
  - la densité de puissance émise par le brouilleur dPBrE (en W/Hz)

    $$dPBrE = \frac{PBr}{NC \times B}$$

    PBr : puissance émise par le brouilleur (en W)
    B : largeur d'un canal (en Hz)
    NC : nombre de canaux susceptibles d'être brouillés simultanément.

  - la localisation du brouilleur :

    - DBr distance entre la maille du réseau considérée (côté réception) et le brouilleur (en km),
    - ALTIBr altitude à laquelle est situé le brouilleur (en m).

- Bruit atmosphérique et galactique

  En dehors du brouillage d'origine humaine, l'environnement se manifeste par des émissions radio-électriques. Le terme prépondérant jusqu'à 20 MHz est le bruit atmosphérique, essentiellement dû aux éclairs lors des orages.
  De 20 MHz à 200 MHz, le bruit galactique devient prépondérant.

  Le facteur de bruit galactique vaut un peu plus de 20 dB à 20 MHz, il décroît de 20 dB par décade (niveau 0 dB à 200 MHz environ). Au-delà de 300 MHz, le bruit galactique peut être considéré comme négligeable [Référence de bruit : kT_0B avec $T_0 = 300$ K].
  B : largeur de la bande passante (en Hz)
  T_0 : température de référence (K)
  k : constante de Boltzman : $1.38\ 10^{-23}$ J/K

- Parasites industriels et urbains

  Les parasites industriels sont créés par un grand nombre d'équipements divers présentant des rayonnements parasites. Certains de ces équipements sont répertoriés au titre des I.S.M. (Appareils industriels, scientifiques et médicaux ...).
  Si le réseau étudié est déployé en zone rurale, les parasites sont le plus souvent inférieurs au bruit galactique. Celui-ci est alors le paramètre à prendre en compte pour établir la portée d'un émetteur en ce qui concerne sa zone de service et sa zone de brouillage.

in a 6- dB additional margin. Coding gain yields an additional 4 dB of margin. The allowable burst rates in the fade-protection mode are 55 and 13.75 Msps. Burst rate reduction does not affect the information transmission rate of a terminal; however, it results in an overhead for the system. That is, it takes four time slots to transmit a telephone circuit, whereas, under the clear-weather condition, it would take one time slot. Because of the resulting system overhead, fade protection must be used prudently to prevent unnecessary network loading.

Propagation Effects and the ACTS Network

As stated earlier, ACTS is able to combat atmospheric-induced impairments up to 15 dB of fade. Because of its Ka-band trans p62 mission, fades caused by rain and, to a lesser degree, by clouds will be frequently observed in the ACTS distributed network. As an example, Figure 2 shows the cumulative attenuation statistics at 20 and 30 GHz for Blacksburg, Virginia, with an elevation angle of 45° and vertical polarization. It can be noted from this figure that fades of 5 dB or higher, and 15 dB or higher, occur with percentages of 1.0 and 0.05, respectively, at 30 GHz. Therefore, ACTS provides an availability factor of 99% for a terminal at Blacksburg without dynamic fade protection. The availability factor is elevated to 99.95% for the same terminal by introducing the dynamic fade protection feature, i.e., burst rate reduction and coding. This example clearly demonstrates the importance of the atmospheric effects on the performance of a link and the ACTS ability to improve link availability.

Note that ACTS is not an operational satellite (it is an experimental system) and, therefore, the validity of a fade protection scheme and its associated threshold values and parameters should be verified experimentally as well as analytically. Robust fade detection schemes need to be developed to ensure high probability of correct fade detection and low probability of a miss. Furthermore, to determine link availability for different climate regions, propagation data is needed.

To address the ACTS needs in the area of propagation, a workshop was convened by the invitation of the Propagation Program [1]. The workshop agenda included planning for the ACTS propagation studies. The objectives of the workshop were to identify general and ACTS-related propagation needs, and to prepare recommendations for a study plan incorporating scientific and systems requirements related to deployment of 8-10 propagation terminals in the USA in support of ACTS experimental activities. The Workshop took place in the Los Angeles area on November 27 and 28, 1989, and was attended by about 40 participants. At the conclusion of the Workshop, many recommendations regarding propagation studies for ACTS were made.

Results of Workshop Deliberations

A brief summary of recommendations resulting from the workshop deliberations is given below. The recommendations are to:

- Complete models for the prediction of attenuation statistics in climate regions of the USA that have not been studied.
- Obtain a statistical description of attenuation and individual propagation events, such as fade slope, fade durations, and interfade intervals, for annual time percentages of 1 to 10% needed for the design of low-margin communications systems.
- Develop algorithms for the dynamic allocation of communication resources.
- Provide a statistical description of the physical processes that give rise to attenuation in the next higher window (about 90 GHz) needed for the design of higher frequency communication systems.

The Workshop also considered the desirability of recording various propagation parameters and certain environmental data, and recommended that the terminal be configured to record the following propagation and meteorological parameters:

- Beacon receiver levels at 20 and 30 GHz.
- Radiometric sky noise (brightness) temperature at 20 and 30 GHz.
- Point rain rate near the terminal.
- Atmospheric temperature and humidity at the earth's surface.
- Ambient temperature of sensitive components affecting the measurements of receiver signal level and sky noise temperature.

Seven regions for making propagation observations were recommended as follows:

Region Climate	Zone
Colorado	B2
Western Washington	C
Northern Michigan	D1
New Hampshire	D1
North Carolina	D3
Florida	E
Southern California	F

Note that these climate zones correspond to the Global rain model as originally introduced by Robert Crane of Dartmouth College.

Workshop participants took note of the short life time of the ACTS satellite (2 to 3 years) and its approaching launch date (mid 1992). The time-critical nature of the ACTS propagation studies was recognized, and it was strongly suggested that planning and terminal development must start as soon as possible. In response to the Workshop recommendations, a two-phase plan was put together. Phase I consisted of a terminal prototype development effort, and Phase II consisted of an experimental terminal construction effort.

Propagation Terminal Configuration

A block diagram of the ACTS propagation terminal is given in Figure 3. The terminal consists of a single antenna, the RF portion, the IF portion, the receiver, the Data Acquisition System, and the radiometer [2]. A single 1-meter antenna was selected to reduce the cost and size of the terminal. The antenna can accommodate dual-frequency and dual-polarization reception. The receiver and radiometer are dual-frequency. The data acquisition system collects and temporarily stores beacon and radiometric data as well as meteorological and other relevant information. The stored data will be periodically dumped on tape or optical disk for transportation.

The beacons at 20 and 27 GHz have EIRPs of 19 and 16 dBW, respectively. The 20-GHz beacon is also used for low bit rate telemetry data transmission. Therefore, a modulation loss is associated with the 20-GHz beacon receiver. Considering a 1-meter antenna and a receiver noise temperature of about 1750 K in clear weather, the receiver dynamic range (margin over threshold) for both frequencies is above 20 dB.

Data Analysis

As mentioned earlier, propagation data will be collected at eight to ten sites across the United States. The collected data will be provided to an investigator, who is likely to be from a university, for data reduction and analysis. It is expected that many centers will participate in the analysis of the reduced data. Therefore, efforts such as model development and testing, algorithm development, etc., will be performed by more than one center.

Financial Support

NASA is providing financial support for the development of the propagation terminals, and their installation and maintenance. NASA will also fund the data analysis effort performed by the principal investigator at a university. Other centers participating in the data analysis part of this effort will finance their own expenses.

Conclusion

The next Workshop on ACTS Propagation Studies will take place in the Los Angeles area on November 27 and 28, 1990. It is hoped that a complete design of the propagation terminal will be available for presentation. Workshop participants will have an opportunity to critique the terminal design and provide comments. The construction of a terminal prototype will proceed shortly after the Workshop.

The current plan for ACTS propagation studies will be revisited at the Workshop. Issues such as the site locations of the terminals, methodology for data processing and model development, fade mitigation algorithms, etc., will be discussed. Although the study plan is not expected to be finalized at this meeting, it is hoped that it will reach a degree of maturity that can allow clear guidelines for the campaign.

The experimental (data collection) phase of the ACTS propagation campaign is expected to begin in mid 1992 and continue for at least 2 years. The results and findings of this campaign will be disseminated in a timely fashion to the ACTS and other technical communities.

Acknowledgement

The research conducted in this publication was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

The author would like to acknowledge the contributions of Dr. D. Chakraborty of JPL and several colleagues at Virginia Polytechnic Institute.

References

[1] F. Davarian, editor, Presentations of the First ACTS Propagation Studies Workshop (APSW I), JPL D-6918, December 15, 1989.

[2] F. Davarian, editor, Proceedings of NAPEX XIV, JPL Publication 90-27, pp. 178-202, July 1, 1990.

V CONCLUSION

L'étude prévisionnelle des performances d'un ensemble complexe de liaisons hertziennes dans des conditions d'environnement réelles est devenue réalisable grâce aux puissances de calcul et à la qualité des interfaces graphiques disponibles sur les ordinateurs.

Les modèles de simulation (type I.R.S.) donnent aux concepteurs un moyen de gérer la complexité du déploiement des systèmes de radiocommunications et d'en améliorer la qualité.

BIBLIOGRAPHIE

[1] P.L. RICE - A.G. LONGLEY - K.A. NORTON - A.P. BARSIS National Bureau of Standards "Transmission Loss predictions for tropospheric communications circuits" Technical Note 101 Vol. 1 and 2 (1966)

[2] G.A. HUFFORD - A.G. LONGLEY - W.A. KISSICK National Telecommunications and Information Administration "Irregular terrain model in the area prediction mode" Rapport 82-100 (1982)

[3] G.A. HUFFORD N.T.I.A. MEMORANDUM "Irregular terrain model" (1985)

[4] Military Handbook "Facility Design for troposcatter Transhorizon Microwave system design" MIL-HDBK-417, (1977)

[5] M. BARTON "Radar System Analysis" Englewood Cliffs (1964)

[6] CCIR "Niveau minimal de bruit externe" Rapport R670 (1978)

[7] CCIR "Caractéristiques du bruit atmosphérique radioélectrique et applications" Rapport R322-3 (1986)

[8] J. CUEUGNIET, J.G. REMY "Systèmes de radiocommunications avec les mobiles" EYROLLES (1988)

[9] H.T. LAU "Algorithms on graphs" TPR - TAB Books Inc. - Blue Ridge Summit, PA (1989)

[10] M. GONDRAN, M. MINOUX "Graphes et algorithmes" Collection de la Direction des Etudes et Recherches d'Electricité de France - EYROLLES Paris (1985)

[11] A. AHO - J. HOPCROFT - J. ULLMAN "Structures de données et algorithmes" IIA Inter Editions Paris (1987)

[12] G.D. GIERHART AND M.E. JOHNSON "The IF-77 Electromagnetic Wave Propagation Model" DOT/FAA/ES-83/3-NTIS ; ADA - 134504 (1983)

DISCUSSION

C. GOUTELARD, FR

Vous avez fait allusion à l'utilisation des cartes relevées par le satellite français SPOT qui donne actuellement les images ayant les meilleures définitions avec des cellules de 20m x 20m.

La diminution des cellules d'analyse entraîne la remontée d'autres phénomènes dans les paramètres du modèle, comme les caractéristiques électromagnétiques du sous-sol où les constructions comme les lignes de chemin de fer, les routes, ...

Pouvez-vous donner votre avis sur cette limitation ?

AUTHOR'S REPLY

En zone urbaine, si on a plusieurs lois paramétriques (résultant d'une approche empirique) pour estimer les différentes contributions à l'affaiblissement, je pense qu'un pas 40m x 40m est bien dimensionné, vu que les futurs réseaux de radiocommunications type ESM (900MHz) PCN (1,8GHz) auront des tailles de cellules réduites à quelques centaines de mètres (350m).

En zone rurale, pour des régions accidentées, un pas de 250 m est trop grand (même si l'on prend des constantes radioélectriques uniformes sur tout le profil de liaison) car la diffraction variant en $K\log_{10}$ h/r (h engagement obstacle dans la première zone de Fresnel, r rayon de l'ellipse de Fresnel) des contributions importantes en première approximation peuvent être sous estimées surtout aux fréquences supérieures à 900MHz.

D. YAVUZ

It may be of interest to you that we have an almost identical system of STC that we call CIPAS (Communications Interference and Propagation Analysis System) using TIREM and DMA digitized terrain data. The system has been operating since about a year and half on an IRIS Silicone Graphics Workstsation. We have been using the system for Hareywick, JMDS coverage characterisation, including a hopping averaging module. The system is highly user friendly and totally menu based. It has and is expected to continue to support a lot of our radio coverage activities.

AUTHOR'S REPLY

I will be interested to have a contact with you about your activities, as we are working for the same subject.
You will find my coordinates at the beginning of AGARD publication 28th.

G. TACCONI

Can the models you presented be considered "robust" in relation with the operative final target of your work ?
In the follow up of such models can be envisaged, at model concept level, an up grading in terms of robustness ?

AUTHOR'S REPLY

For system performances evaluation the accuracy of the used propagation model (ITS 122) is supposed to be sufficient for rural zone. However, for certain path profiles ITS 122 can be enhanced. So the overall simulation that was the purpose of the 28th AGARD report can be upgrading by the means of complementary modules (propagation in urban areas, signal processing, protocols), and the upgrade of the modular software can be easily made. To achieve reliable outputs for radiocommunication system design.

COMBINING OF SIGNALS IN A GEOGRAPHICAL DIVERSITY SYSTEM

by

T.J.Speight
Hull-Warwick Communications Research Group
University of Hull
Hull HU6 7RX
United Kingdom

Summary

This paper will describe a technique which will improve performance of macroscopic diversity processing while minimising additional computational and communications channel overheads. It will be aimed specifically at the use of such diversity processing in the HF environment.

1) Introduction

Mobile stations suffer from a number of disadvantages over fixed terminals especially in the H.F. band. A few of these disadvantages are :

1) Communication depends upon a propagation path whose characteristics may change with both time and position.
2) Only relatively low transmitter powers are available.
3) Mobile antennas in the HF band tend to be simple, inefficient and largely non-directive.
4) Reception may be impaired by local noise/interference from co-sited EM systems.

One of a number of possible ways to mitigate these problems is the use of diversity combining. Diversity combining schemes can be classified into two types ; 'macroscopic' and 'microscopic' diversity[1]. Microscopic diversity is a commonly used technique and is used essentially for combining short-term Rayleigh fading signals. An example of where microscopic diversity would be effective is the case of mobile station reception in an urban environment where the signal received by the base station is the sum of signals transmitted from the mobile station which have travelled on different paths. To receive two uncorrelated signals in the microscopic diversity situation receiving antennas need to be separated by as little as 0.5λ. Macroscopic diversity is used to combat long-term log-normal fading. In this case a number of totally separate independently fading paths are received from two or more different base stations. This technique will overcome shadowing effects in a line-of-sight communications system based in hilly terrain and also propagation effects in a communications system in the HF band.

In the HF band macroscopic diversity can be described as 'geographical diversity' [2] because of the large distances between base station sites. Figure 1 shows a geographical diversity system which employs 5 base station sites. To perform combining the 5 versions of the signal must be brought to a central site (probably one of the base stations); therefore interconnects are required to this site. Because of the distances between base stations moving base-band data between sites would require interconnects of very high capacity. For this reason previous implementations of macroscopic diversity have almost exclusively used a selection combining scheme ie select the base station which will produce demodulated data with the lowest error rate. Selection combining is an inefficient combining method. This paper will describe a method of combining in a geographical diversity system which approximates to maximal ratio combining but using interconnects which can be implemented in the low data rates available to typical HF channels.

2.1) Development of Optimum Weight Combining

The reduction of overhead on the interconnects is achieved by the use of post-demodulation combining. Therefore each base station demodulates the data sent by the mobile into a version of the original source data. Combining is then carried out on this binary data. In addition to this demodulated data the combining site also requires channel information. This would ideally be Real Time Channel Evaluation (RTCE) [3] data but could also be off-line prediction data. This data can readily be turned into a value of probability of bit error and it is this value together with demodulated binary data which will be used in the new combining method.

The 2 ground rules which the combining scheme must use are as follows:

1) If any of the combiner inputs have a P_e of 0.5 (or higher), ie they carry no information, then such an input should have no effect on the output of the combining system.

2) If any input has a P_e of 0.0 (always correct) then this input should have total control of the combiner output however close the other inputs are to 0.0.

From these two rules two points can be made about the graph of P_e versus weight (this graph is called the weighting function):

1) When the P_e of any input is 0.5 then its weight should be zero, ie the curve must pass through the point (0.5,0).

2) Weight must tend to infinity as P_e tends to 1.0.

Also it would be logical to assume that the graph has no maxima or minima (ie always increasing). From these postulates a family of curves with the general shape of that in figure 2 can be drawn. An empirical expression which has these properties is as follows:

$$\text{Weight} = 1/(P_e^n) - 1/(0.5^n) \qquad (1)$$

$$= 1/(P_e^n) - 2^n$$

where 'n' is any positive real number.

After calculation of a weight for each combiner input a voting system is used to make a decision on the binary data. An example of the scheme follows:

A five branch combiner has P_e's of:

$P_{e1} = 0.07$ $P_{e2} = 0.11$ $P_{e3} = 0.2$ $P_{e4} = 0.25$
$P_{e5} = 0.25$

Using equation (1) with n = 0.5 the following weights are calculated:

$W_1 = 2.37$ $W_2 = 1.60$ $W_3 = 0.82$ $W_4 = 0.59$
$W_5 = 0.59$

Suppose that the outputs of demodulators at these branches have the following binary outputs:

$O_1 =$ '1' $O_2 =$ '0' $O_3 =$ '0' $O_4 =$ '1'
$O_5 =$ '1'

Now the total weight for binary output '1' is
2.37 + 0.59 + 0.59 = 3.55.
The total weight for binary output '0' is
1.60 + 0.82 = 2.42.
Therefore the output of the combiner for the above parameters is: '1'.

A notation for error patterns in the combining system will now be introduced. C is used to denote a correct input and I an incorrect input branch. The branches are then ordered so that the first C/I refers to the branch with the highest SNR and the last C/I with the lowest SNR.

By examining all possible error patterns the total output P_e can be calculated. This is shown in Appendix 1. The results obtained are as follows:

$P_{e(\text{weight \& vote})} = 0.0322$
$P_{e(\text{majority vote})} = 0.03721$
$P_{e(\text{selection})} = 0.07$

The weight and vote algorithm compares favourably with other traditional combining methods, as can be seen from the example above. The reason for the error rate advantage over majority vote combining is that the error pattern CCIII gives a correct output using the weight and vote algorithm. This error pattern has a higher probability of occurring than its inverse IICCC which would have given a correct output using majority vote combining. However the output error rate is not the lowest possible for the above P_e's. As mentioned earlier the weighting function of equation (1) is one of a family of curves so it does not necessarily give the optimum weights. If Appendix 1 is studied it can be seen that the error pattern CICII is combined to give a correct output but its inverse ICICC which has a higher probability of occurring gives an incorrect output. An iterative method can be used to obtain optimum weights. Figure 3 shows a flow chart of this process.

Using the values of input P_e already stated the optimum weights are:

$W_{1'} = 2.24$ $W_{2'} = 1.73$ $W_{3'} = 0.69$ $W_{4'} = 0.65$
$W_{5'} = 0.65$

Using these weights the following result is obtained:

$P_{e(\text{optimum weight})} = 0.030325$

The above error rate is the lowest possible if all that is available are binary demodulator outputs and SNR data for each branch receiver.

2.2 Optimum Weight Combining in a Fading Channel

If RTCE data is not available all the time for the branches of the combining system then the SNR data for each branch will be out of date. Thus in turn the weight values assigned to the branches of the combiner may be wrong and this will cause a decrease in the performance of the combiner. For example, suppose a channel has a high SNR when it is tested and, therefore, it is assigned a high weighting value: if fading occurs then this branch of the combiner will have a large effect on the combiner output but it could have a relatively high P_e. Intuitively, it is seen that when fading occurs then the weights of the inputs to the combiner should be made more equal. Therefore it may be that a majority vote scheme ,where all weights are equal irrespective of SNR, gives optimum performance under fading conditions.

A typical case where no RTCE data is available is when an offline prediction program is used to makes a first estimate of the channel. Over short periods (3-7minutes) the field strength follows a Rayleigh distribution where the median and mean values are 1.6dB and 1.1dB respectively below the RMS value [4].

Using the above mean and median values and the following P_e's an error distribution for various combining schemes was determined.

$P_{e1} = 0.95$ $P_{e2} = 0.90$ $P_{e3} = 0.85$ $P_{e4} = 0.75$
$P_{e5} = 0.75$

Figure 4 shows a graph of probability of occurrence versus total output error rate for various combining schemes. The selection combining scheme uses a single branch of the combiner inputs and therefore the output from a combiner using this method will have a broad probability distribution. Clearly the selection combining scheme is strongly affected by fading and should not be used unless real-time channel data is available.

Referring again to figure 2.3, the majority vote algorithm has the highest peak and narrowest probability distribution which suggests it is least affected by fading. This is as expected because the inputs are not weighted, whatever the SNR of the channel.

The optimum weight combining algorithm has a distribution which is between the majority vote and selection combining schemes. The optimum weight method uses all combiner inputs but the inputs which have a higher SNR have more control of the output than the ones with a low SNR.

The average P_e for the three combining schemes under Rayleigh fading are :

$P_{e(selection)} = 0.0935$
$P_{e(majority\ vote)} = 0.553$
$P_{e(optimum\ weight)} = 0.0470$

It should be stressed that these error rates are for particular input values and do not give an absolute result. Nevertheless the input values are relatively similar which is the case where majority vote gives an output which is near the optimum. However, the more severe the fading the more likely is the majority vote combining of being the best combining scheme. It would seem logical then to 'move' towards majority vote combining when fading occurs. Figure 5 shows a probability of occurrence versus error rate graph (with the same inputs as figure 4) showing optimum weight combing and a combining method 'between' majority vote and optimum weight. This new method is achieved by making the weights of the inputs more equal, specifically :

$$W_{new} = W_{old} + (W_{1(old)})/3$$

As expected the new weighting scheme has a higher peak; however the average error rate is reduced :

$P_{e(optimum\ weight)} = 0.0470$
$P_{e(new\ weighting\ method)} = 0.0457$

Therefore there are different optimum weights for fading or non-fading conditions.

To further illustrate how these new weights are obtained it is necessary to look more closely at how the error patterns are combined from demodulated data.

If the selection and majority vote algorithms are compared some error patterns give the same output for both of these schemes. For these error patterns it can be assumed that whatever weighting system is used then this too will give the same output as both selection and majority vote. Hence it is possible to say that the following error patterns (EPx, where x is decimal version of the binary number created assuming I as a binary '1' and C as a binary '0'.) are always correct for any combining scheme :

EP0, EP1, EP2, EP3, EP4, EP5, EP6, EP8, EP9, EP10

Similarly the following always give an incorrect output:

EP21, EP22, EP23, EP25, EP26, EP27, EP28, EP29, EP30, EP31

Note that EP12 is not counted in the group of 'always right' error patterns even though both the majority vote and selection algorithms give a correct output for this pattern. EP12 is one of the group of error patterns, $C^n[n+1]\{.....\}$, (where the number of I's is less than half the number of branches), which can give a correct or incorrect output for an optimum weight combining scheme, even though this error pattern always gives a correct output for selection and majority vote schemes. Note also that the inverse of EP12, EP19 is not included in the always incorrect group for a similar reason.

Any combining scheme can now be characterised by its behaviour to the error patterns which give different outputs for majority vote and selection and also the CnIn+1{...} type error patterns. Therefore critical error patterns can be calculated for various numbers of branches. A selection is shown below:

Branches	Critical Error Patterns
4	6,7
5	7,11,12,13,14,15
6	14,15,22,23,24,25,26,27,28,29,30,31
7	15,23,27,28,29,30,31,43.....63
8	30,31,46,47,54.....63,78,79,86.....127

The use of these critical error patterns can substantially speed up the calculation of optimum weights.

2.3 Use of Optimum Weights for a First Prediction in an HF Message System

If a mobile station first logs onto the paging system the base stations can estimate SNR data on the channel to the mobile station for various different frequencies by the use of an offline prediction program. By looking at the SNR data for different frequencies it should be possible to determine which frequency will give the lowest overall error rate using an optimum weight combining scheme. It would seem logical that the frequency with the highest total weight will be the channel with the lowest error rate. Because the prediction program can only give a monthly mean of the channel conditions it is not critical that the weight calculation is accurate so the iterative process for calculating weights, which is described earlier, was not used. The weighting function used is :

$$\text{Weight} = 1/(P_e^{0.5}) - \sqrt{2}$$

This function was used in conjunction with the propagation prediction program SUP252[4]. This program was written in FORTRAN IV and originally designed to run on a VAX or similar minicomputer; however this program was converted to run on an IBM PC clone. The original SUP252 program gave a lot of data about one channel, ie from one transmitting station to one receiving station. This program was modified and a typical output of the new program SUP252A3 is shown in Appendix 2.

Conclusions

This paper has oulined a scheme which will substantially improve performance especially in communications environments which experience substantial slow fading such as in the HF band.

A first prediction algorithm which utilizes the concepts has also been described and it is anticipated that this method will be used in an HF message system being developed at Hull.

It was anticipated that off-air results of a geographical diversity system would be presented with this paper. However this was not possible but it is hoped that these will be carried out in the near future. The mobile is to be located at Hull and base stations will be at Cobbett Hill (Hampshire,England), Coventry (West Midlands,England) ,and Wick (Caithness,Scotland). This gives a good spread of base stations sites in a North South axis.

References

[1] Lee, W.C.Y : 'Mobile Communications Engineering', McGraw-Hill 1982, ISBN 0-07-037039-7.

[2] Darnell, M. : 'Probelms of mobile communications and techniques for performance improvement', Proc. IEE(part F), Vol 132,No.5, Aug 1985, pp 433-440.

[3] Darnell, M. : 'Real-time channel evaluation', AGARD Lecture Series No.127 (Modern HF communications), May/June 1983.

[4] International Radio Consultative Commitee (CCIR) : Supplememnt to Report 252-2.

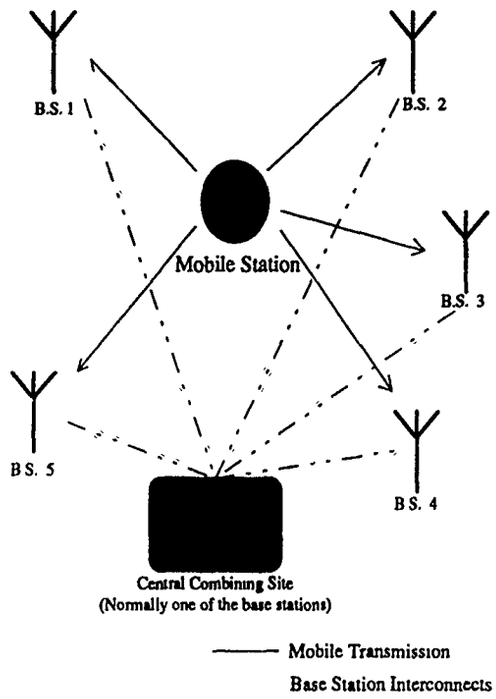


Figure 1. A geographical diversity system

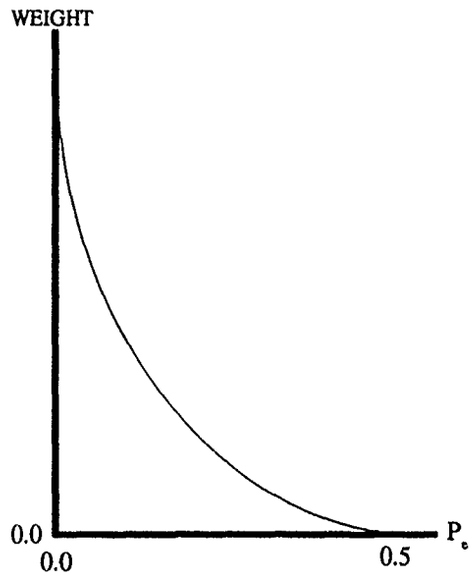


Figure 2. Generalized graph of P_e versus weight.

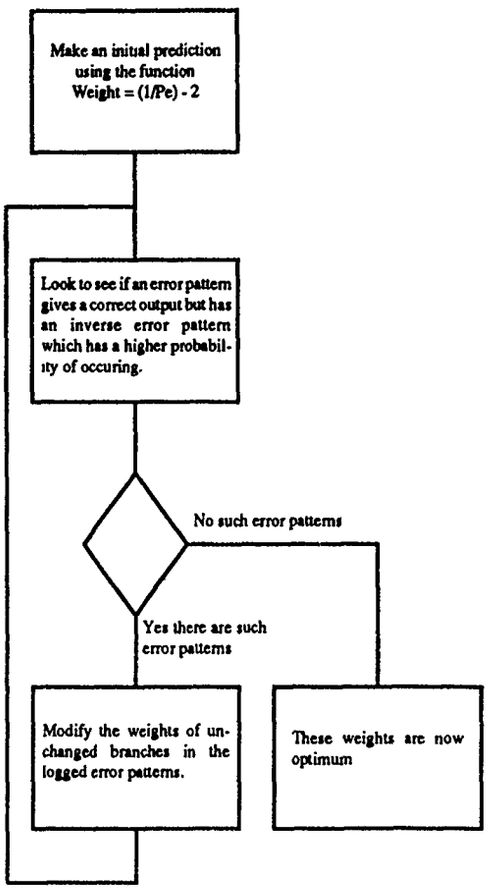


Figure 3. Iterative Method for calculation of optimum weights.

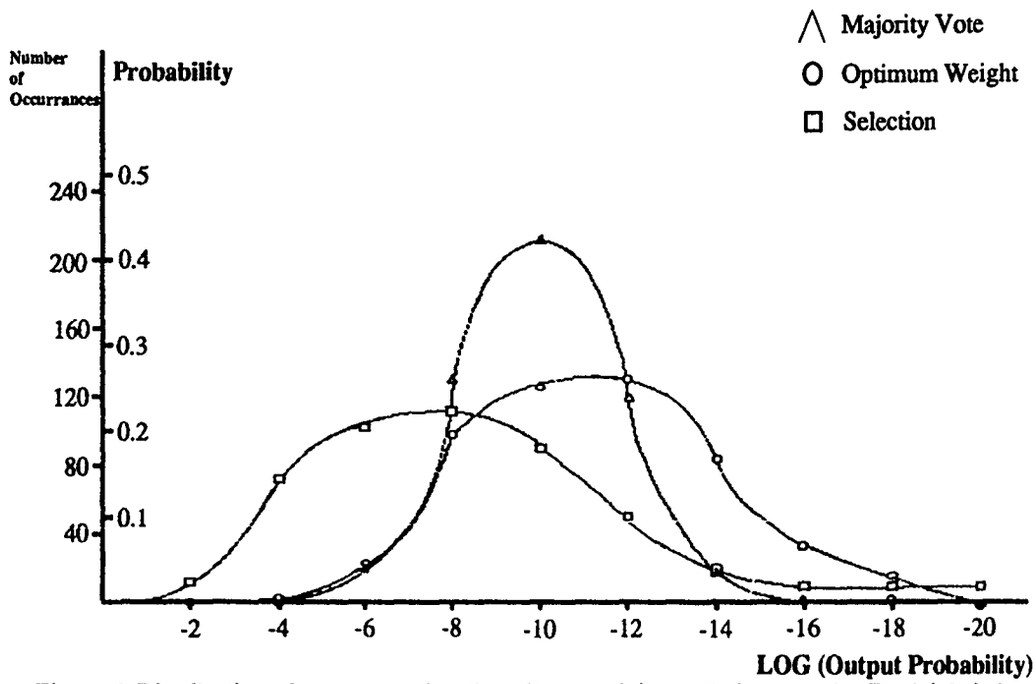


Figure 4. Distribution of error rates for diversity combining techniques under Rayleigh fading conditions

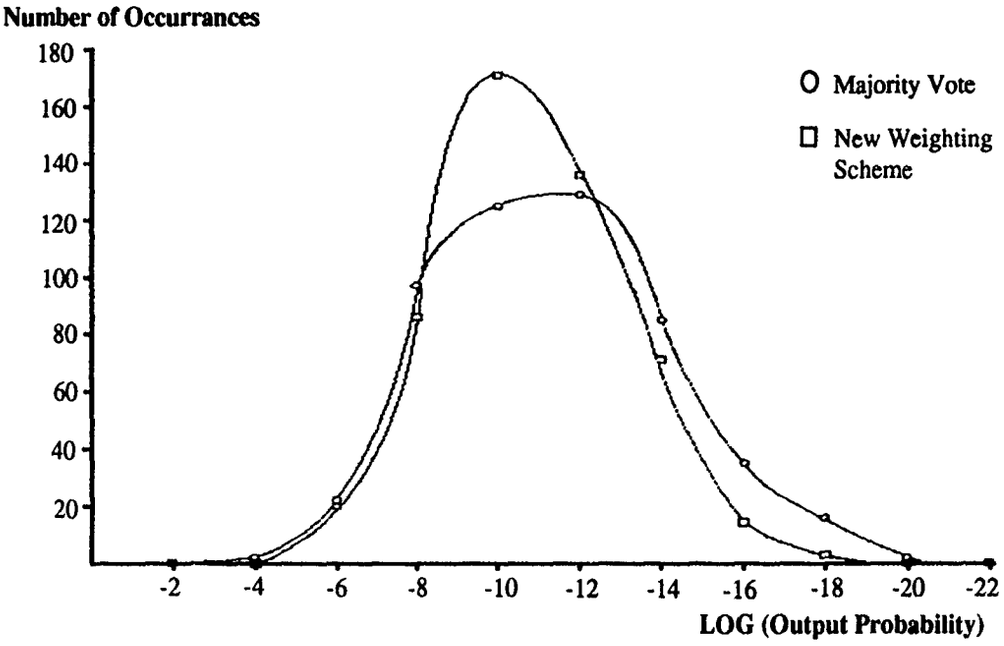


Figure 5. Comparison of optimum weight combining and a new combing scheme

APPENDIX 1

Full Study of Combiner Outputs using Weight = 1/(Pe0.5) - √2

$P_{e1} = 0.07$ $P_{e2} = 0.11$ $P_{e3} = 0.2$ $P_{e4} = 0.25$ $P_{e5} = 0.25$
$W_1 = 2.37$ $W_2 = 1.60$ $W_3 = 0.82$ $W_4 = 0.59$ $W_5 = 0.59$

Error Pattern	Probability	Voting	Output				
CCCCC	0.3725	5.97 - 0.00	C	ICCCC	0.0280	3.60 - 2.37	C
CCCCI	0.1242	5.58 - 0.59	C	ICCCI	0.0093	3.01 - 2.96	C
CCCII	0.0414	4.79 - 1.18	C	ICCIC	0.0093	3.01 - 2.96	C
CCICC	0.0931	5.15 - 0.82	C	ICCII	0.0031	2.42 - 3.55	I
CCICI	0.0310	4.56 - 1.41	C	ICICC	0.0070	2.78 - 3.19	I
CCIIC	0.0310	4.56 - 1.41	C	ICICI	0.0023	2.19 - 3.78	I
CCIII	0.0103	3.97 - 2.00	C	ICIIC	0.0023	2.78 - 3.19	I
CICCC	0.0460	4.37 - 1.60	C	ICIII	0.0008	2.19 - 3.78	I
CICCI	0.0153	3.78 - 2.19	C	IICCC	0.0035	2.00 - 3.97	I
CICIC	0.0153	3.78 - 2.19	C	IICCI	0.0016	1.41 - 4.56	I
CICII	0.0051	3.19 - 2.78	C	IICIC	0.0016	1.41 - 4.56	I
CIICC	0.0151	3.35 - 2.42	C	IICII	0.0004	0.82 - 5.15	I
CIICI	0.0038	2.96 - 3.01	I	IIICC	0.0008	1.18 - 4.79	I
CIIIC	0.0038	2.96 - 3.01	I	IIICI	0.0003	0.59 - 5.38	I
CIIII	0.0013	2.37 - 3.60	I	IIIIC	0.0003	0.59 - 5.38	I
				IIIII	0.0001	0.00 - 5.97	I

APPENDIX 2

Typical output of SUP252A3
Mobile at OSLO
Base Stations at : BERLIN, PARIS, ROME, FRANKFURT, LONDON

FREQ	SNR	P_e	WEIGHT	
		BERLIN		
2.00	-48.33	0.50	0.00	
2.50	-33.65	0.49	0.01	
3.00	-26.07	0.47	0.03	
3.50	-25.22	0.47	0.03	
4.00	-18.25	0.43	0.08	
5.00	-10.75	0.34	0.22	
6.00	-6.26	0.25	0.43	
7.00	-2.88	0.16	0.79	
8.00	+0.12	0.08	1.47	
10.00	-2.20	0.14	0.90	
14.00	-49.93	0.50	0.00	
		PARIS		
2.00	-36.84	0.49	0.01	
2.50	-28.47	0.48	0.02	
3.00	-23.72	0.46	0.04	
3.50	-17.23	0.42	0.09	
4.00	-13.67	0.38	0.15	
5.00	-7.40	0.27	0.36	
6.00	-3.16	0.16	0.75	
7.00	+0.13	0.08	1.47	
8.00	+2.08	0.04	2.36	
10.00	-36.66	0.49	0.01	
14.00	-59.03	0.50	0.00	
		ROME		
2.00	-44.88	0.50	0.00	
2.50	-37.52	0.49	0.01	
3.00	-32.51	0.49	0.01	
3.50	-26.78	0.47	0.03	
4.00	-18.31	0.43	0.08	
5.00	-12.17	0.36	0.18	
6.00	-17.36	0.42	0.09	
7.00	-13.98	0.39	0.14	
8.00	-11.36	0.35	0.20	
10.00	-6.12	0.24	0.44	
14.00	-77.97	0.50	0.00	
		FRANKFURT		
2.00	-40.25	0.49	0.01	
2.50	-30.04	0.48	0.02	
3.00	-23.76	0.46	0.04	
3.50	-21.41	0.45	0.05	
4.00	-15.95	0.41	0.11	
5.00	-9.44	0.32	0.26	
6.00	-5.19	0.22	0.52	
7.00	-1.91	0.13	0.95	
8.00	+1.24	0.05	1.91	
10.00	-3.07	0.16	0.76	
14.00	-55.54	0.50	0.00	
		LONDON		
2.00	-29.21	0.48	0.02	
2.50	-21.47	0.45	0.05	
3.00	-14.39	0.39	0.13	
3.50	-9.67	0.32	0.26	
4.00	-3.46	0.17	0.71	
5.00	+0.70	0.06	1.68	
6.00	+2.26	0.04	2.48	
7.00	-2.42	0.14	0.86	
8.00	-36.10	0.49	0.01	
10.00	-66.97	0.50	0.00	
14.00	-70.41	0.50	0.00	

TOTAL WEIGHTS FOR EACH FREQUENCY

2.00	0.03
2.50	0.08
3.00	0.18
3.50	0.34
4.00	0.67
5.00	1.73
6.00	3.47
7.00	5.82
8.00	6.80
10.00	2.12
14.00	0.00

8.00 MHz SELECTED

DISCUSSION

F. DAVARIAN

What are the system dynamics ? Is there enough time to calculate the coefficients in a rapidly changing environment ?

AUTHOR'S REPLY

It is unlikely that more than 5 base stations would be used in a geographical diversity system. In this case there are only 6 critical error patterns. The iterative algorithm used to calculate optimum weights needs a maximum of 3 iterations. This is a relatively low computational overhead. Changing environments do no as such put any pressure on the speed of calculating weights it is only the rate at which new channel data is available. This is likely to be at the packet rate. Since packets are not likely to be very short no problem is anticipated.

H. DE PEDRO

Do you account in your algorithm for the fact that there is a second link, probably also a fading path, that the message takes in reaching the combining site ?

AUTHOR'S REPLY

No, at present the algorithm does not take into account errors which occur due to the interconnect links between base stations. It was assumed that the critical path in the system is that from mobile to base station, and the algorithm has been developed with this in mind. No consideration has been given to errors which occur in the interconnect paths. However the point raised here is an interesting one. Channel information should be available for the base-base transmissions. If this is the case then the Pe for data from each base to base path can be calculated. Now a new total Pe can be calculated over the two links :

Pe(total) = Pe(mobile-base) x Pe - (base-base)

The weights used in the combining algorithm now uses Pe(total) instead of just Pe(mobile-base).

D. YAVUZ

The diversity effects your approach utilizes is based on the existence very large separations, perfect between base stations and central station. "Optimisation" is also based on the independant additive gaussian noise channels combined with a maximal radio combining (not really applicable to the complex HF channel).

Would it not be for more efficient/cost effective to use such dispersed assets in an internetworked store and forward packet radio system ?

AUTHOR'S REPLY

The diversity system described in the presentation does assume perfect communication between base stations. However it would be easy to use channel information from base-base station interconnects and calculate a new Pe specifically

Pe(total) = Pe(mobile-base) x Pe(base-base)

This would then be used to calculate weights. Although I conceed more work is required in this field.

To address the second point to your question I would like to say that it is not the combining method which determines whether the system is applicable to the HF band but the topological architecture of the diversity system, if Pe is used as the input of the combining scheme. I may have given the impression that this value of Pe is calculated directly from SNR. The value of Pe would normally be calculated by a real time channel evaluation technique such as counting of attempted corrections in an FEC coding scheme.

The design ethic of the system in which this technique is to be used is the maximization of reliable comunication from mobile to base station. Hence the use of a number of receiving sites which, no doubt, will improve the performance of this link.

C. GOUTELARD, FR

Votre proposition est intéressante car elle propose une combinaison sur les décisions et non sur l'optimisation du rapport S/N comme dans les méthodes classiques, ce qui est plus pratique dans la structure que vous proposez. Je partage totalement les remarques formulées par le Dr YAVUZ sur la nécessité d'éloigner les stations. En effet, pour des distances inférieures à 500 km, le Pr GOTT et moi-même avons montré la forte corrélation entre deux stations. Il faut donc éloigner les stations et dans une publication que nous avons faite au congrès AGARD de Münich en 1988 nous avons montré

qu'avec une base de trois stations séparées de 500, 700, 1000 km, cette décorrélation était juste atteinte.

AUTHOR'S REPLY

I acknowledge Prof Goutelard's comments. At the moment I have no off-air data for the diversity combining scheme. I hope to carry out trials of the geographical diversity system in the near future. This trial will include short paths (< 500 km) and long paths (> 1000 km), and this should show whether base stations should be separated by distances greater than I anticipated.

To the second point I would like to say that I have not read the article which you mention but would like to study it.

DISTRIBUTED LONG-RANGE RADIO SYSTEMS EMPLOYING MULTIPLE PROPAGATION MECHANISMS AND AN EXTENDED FREQUENCY RANGE

by
M. Darnell
Hull-Warwick Communications Research Group
Department of Electronic Engineering
University of Hull
Hull HU6 7RX
UK

The paper is essentially a discussion paper which introduces, with the aid of data from real radio paths of different types, the concept of multiple propagation mechanism radio communication in the HF and low-VHF bands up to about 100 MHz. The radio system architecture necessary to exploit such a multiple-mechanism environment is then considered, with particular reference to the channel encoding and channel evaluation procedures required. Digital signal processing is assumed.

INTRODUCTION: MULTIPLE MECHANISM PATHS & DIGITAL SIGNAL PROCESSING

The great majority of modern radio communication systems are designed to exploit a given propagation mechanism in order to meet their operational requirements. This means that the parameters of such systems are, to a large extent, dictated by the characteristics of the specified propagation mechanisms. Currently, most practical radio systems make use of fixed signal formats and control procedures. If a system is adaptive in response to channel conditions, as sensed say by real-time channel evaluation (RTCE), then the nature of that adaptation will be determined by the range of variation expected of the specified mechanism.

The design philosophy outlined above is somewhat artificial in that there are many types of radio path which do not exhibit a single dominant propagation mechanism; rather, a number of distinct mechanisms with very different characteristics may exist - possibly simultaneously. Consequently, these additional mechanisms will tend to be regarded effectively as noise/interference by a communication system designed for a specific mechanism. Examples of this include microwave line-of-sight (LOS) paths over which tropospheric propagation can also occur, HF surface wave paths with skywave interference, etc. In the part of the radio spectrum up to say 100 MHz, the situation may be extremely complex, as will be demonstrated in this paper. Nevertheless, from an operational viewpoint, this HF / low-VHF region is still of great importance; consequently, every available method should be sought to improve the availability and reliability of LOS and beyond-LOS (BLOS) communication systems using this range of radio frequencies. An obvious starting point is to take a realistic view of the nature of the radio path to be employed; the following section of the paper attempts to do this.

In parallel with this consideration of the radio path, an equally realistic view needs to be taken of the technology now available to implement radio systems. Over recent years, there has been a major change in the signal generation and processing techniques and devices available to the communication system designer; as yet, however, this change has not become apparent in operational systems. The reasons for this are simple to identify:

(i) the time taken for new designs to become operational is relatively long - a minimum of several years;

(ii) there is a tendency for designers to still think in "conventional" terms, ie implementing functions in an essentially analogue manner;

(iii) the fundamental nature and potential of the propagation medium, particularly at the lower frequencies below about 100 MHz, it not yet appreciated fully.

There is a need for design thinking to become digitally oriented; signal generation and processing algorithms and techniques implemented in digital form must exploit the unique capabilities of digital signal processing (DSP) devices, rather than simply replicating functions previously implemented in analogue form. The most obvious implication of this philosophy is that, although processing and control structures implemented in digital form are in part limited by the capabilities and structure of DSP hardware, the corresponding software can be altered at will in order to create an overall architecture that is both capable of performing a variety of functions and being responsive to user requirements and channel state. Coupled with flexible DSP architectures for control and channel encoding, there is now a widespread availability of relatively cheap and frequency-agile RF equipments.

Given the potential of system architectures of the type indicated above, the practical problem is how best to employ them to meet specified and evolving operational

requirements. This paper will suggest ways in which this might be achieved in the HF / low-VHF band below about 100 MHz, where multiple simultaneous propagation mechanisms can exist. A wrong view of the path, or a "traditional" approach to signal processing and control, may lead to the implementation of a radio system which does not perform optimally. In the following sections, three important aspects of the problem will be discussed, ie the nature of the propagation path, appropriate channel encoding and RTCE techniques, and radio system terminal architectures based upon DSP.

THE NATURE OF PROPAGATION IN THE BAND UP TO 100 MHZ

The individual propagation mechanisms that can occur in the band up to 100 MHz will now be reviewed briefly by means of the Table 1.

To this list of propagation mechanisms may be added the phenomenon of "ducting", whereby atmospheric refractive index variations close to the earth's surface can trap radio waves in ducts, analogous to waveguides. Since waveguide propagation tends to be low-loss, this form of energy transfer can allow radio waves to propagate at relatively high levels over distances considerably in excess of LOS. Ducting is comparatively unpredictable, its incidence being correlated to some extent with weather conditions. References 1 to 7 describe the characteristics of these propagation mechanisms in more detail.

All the entries in the preceeding Table correspond to well known and reasonably well characterised mechanisms in the band up to 100 MHz. The only exception to this is tropospheric scatter which is normally associated with high-capacity transmission systems at considerably higher frequencies. However, recent work [8] [9] has shown that troposcatter is a viable means of information transfer at much lower frequencies, down to 30 MHz.

To illustrate this mechanism, some experimental data from a troposcatter link operating at a frequency of 47 MHz over a mid-latitude path length of approximately 300 km will now be presented. The transmitter power was nominally 500 W, with the transmitting antenna taking the form of a 5-element Yagi. At the receiver, a 3-element Yagi was used at a height of 20 m above the immediate ground. Figs. 1 and 2 show the temporal variations of received signal level over an interval of a few seconds; the transmitted waveform takes the form of a regularly on/off keyed signal, with "on" and "off" periods both being 40 ms. Figs. 3 and 4 are plots of the diurnal variation of average signal level. These plots are typical of the range of variation monitored over a period of several months.

It is interesting to note that the 40 ms "on" periods are modulated with phase-coded information at about 4 kbits/s; the receiver terminal unit (nominally designed for meteor-burst reception) was able to demodulate this data successfully for a large proportion of the time, thus showing that the mechanism has significant communication potential at this lower frequency over distances of up to several hundreds of kilometres.

Fig. 5 shows two representative profiles recorded from a 1000 km meteor-burst (MB) link. It is seen that the received signal level varies widely as the burst progresses. In the tail of the burst, there is frequently a fading regime which, if it can be exploited (say by diversity processing), may extend the effective length of the burst substantially [10]. A common phenomenon encountered on this same link is the presence of a continuous low level component of the received signal which persists between bursts. This is attributable to ionospheric scatter and, if recognised in the system design, can be employed for low-rate data transmission, say for engineering order wire (EOW) purposes.

As indicated previously, the propagation environment below about 100 MHz is extremely complex; radio systems optimised for one mechanism will clearly make only inefficient use at best of any others which may be present. In the worst case, any other mechanisms which the system is not designed to exploit may be detrimental to overall performance and effectively appear as noise or interference to that system.

In the following two sections of this paper, techniques which appear to have promise in making efficient use of multiple-mechanism propagation paths (MMPPs) will be considered.

CHANNEL ENCODING & RTCE FOR MULTIPLE-MECHANISM PROPAGATION PATHS

The combination of the mechanisms in an MMPP will give rise to received signals with a wide range of variability in their characteristics. This variability will manifest itself to the communicator as a channel capacity which changes with time. If a radio system is to make effective use of this variable capacity, its parameters must be adapted in response to the channel state at any time. This section will now examine channel encoding/decoding techniques which are designed for such an environment. To initiate the necessary adaptation, the channel state must be quantified by means of appropriate RTCE procedures, preferably integrated with the channel encoding procedures.

In the context of the type of radio system being discussed, channel encoding will be taken to encompass the following functions:

(i) synchronisation;

(ii) error control;

(iii) modulation;

(iv) multi-user coding;

(v) security.

A unifying concept which relates all the above functions is termed "multi-functional coding" [11]. Here, it is assumed that the receiver will possess a certain fixed processing power; in an adaptive system, the major practical problem is how best to distribute this power between the various channel coding functions in order to optimise the system performance with respect to the user requirements. Taking a simplified example: as the state of the channel deteriorates, it may be desirable to shed some of the low-priority users in a multi-user system and apply the processing power thus released to more robust error control.

Perhaps the most basic single channel encoding function in a digital transmission system is that of synchronisation; without effective synchronisation, the system cannot operate at all. In an MMPP, one important effect that can occur is a rapid change in propagation delay due to a change in dominant mechanism. If this change does not exceed one symbol interval, then conventional synchronisation tracking arrangements such as delay-lock loops can cope successfully. Should the delay change be greater than one symbol interval, however, the system may need to reacquire symbol synchronisation. For MMPPs, therefore, it would appear necessary to incorporate an on-going "flywheel" synchronisation mode, together with an integral acquisition mode embedded at intervals within the transmitted data. Also , it is important that both symbol and block synchronisation are maintained.

In reference [12], two techniques, known as modulation-derived synchronisation (MDS) and code-derived synchronisation (CDS), are described; these allow the extraction of synchronisation information from the normal operating signals of a radio system by means of specific DSP algorithms. In the case of MDS, the output of a quadrature digital matched filter is monitored at each sampling instant; it is found that the output of this matched filter provides a robust timing waveform at the symbol rate. With CDS, the received signal is applied with all possible codeword time shifts to the error control decoder where it is comapred with all possible codewords; at the condition corresponding to the correct codeword and shift, the decoder output will be a minimum - assuming that the error control power of the code has not been exceeded; this is information that can be used for block synchronisation.

Recent research carried out by the Hull-Warwick Communications Research Group (HWCRG) has demonstrated that the conventional philosophy applied to the design of synchronisation preambles requires modification. The accepted requirement for an acquisition preamble is that it should have an aperiodic autocorrelation function (ACF) which is as impulsive as possible. It has been shown that preambles with significant negative ACF sidelobes perform better in the presence of noise. To illustrate this result, Fig. 6 shows the ACFs of 7 distinct 22-bit binary preamble sequences; Figs. 7 and 8 respectively show the practical and theoretical ranking orders for these seven sequences in terms of their synchronisation failure rates (threshold crossing) as a function of average bit error rate (BER). It is seen that the best performance is achieved by sequence (b) which has large negative ACF sidelobes. An additional benefit conferred by the presence of large positive and negative sidelobes with a specified time separation is that the preamble can simultaneously provide estimates of both block and symbol timing. Such preambles could be embedded in the transmitted signal at intervals dictated by the expected variations in propagation mechanisms/characteristics. Their length could be adapted in response to changes in received signal-to-noise ratio (SNR). This would allow the system to quicky reacquire both block and symbol synchronisation after significant changes in propagation delay.

In a variable channel capacity environment, it is important that both error control and modulation procedures can adapt to the channel state. Embedded coding [13] and embedded modulation [14] have the potential to do this. In both cases, the signal format simultaneously contains elements that have different levels of resilience to noise and distortion. When the SNR is high, all elements can be received successfully; as the SNR degrades, only the more robust elements can be received successfully. In this way, the effective information transmission rate is matched to the prevailing channel state. Typically, an ARQ protocol would be used in conjunction with these embedded schemes. These embedded schemes have been designed primarily with a multiple frequency-shift-keyed (MFSK) modulation format in mind. Such a modulation format is fundamentally robust in the HF environment, as demonstrated by the PICCOLO system. It also has the advantage that it can be made adaptive by varying the number of tones and symbol duration in response to channel conditions.

Complementary sequences [15] provide a vehicle which can implement all the channel encoding elements listed previously in a multi-functional manner. They are self-synchronising and possess distributed error control properties; in addition, they can be synthesised in uncorrelated sets thus giving a multi-user capability.

In order to control the parameters of adaptive channel

encoding schemes of the type described previously, it is necessary for data on current channel state, ie RTCE information, to be available in a convenient form. The trend in RTCE systems is now to attempt to extract as much information as possible on channel state from the normal operating signals and elements of the radio communication system, eg from demodulators, decoders, synchronisers, protocol activity, etc; these techniques are collectively termed "embedded" RTCE procedures [16]. The advantage of embedded RTCE is that the information required can be extracted simply and naturally from systems implemented using DSP devices.

In a MMPP environment, the RTCE procedures can be made more effective if more time is provided for the analysis to be carried out. One way of achieving this simply is illustrated in Fig. 9; here, the received signal is replicated and one version is fed directly to an RTCE analyser whilst the other passes through a delay unit. By the time the received signal is applied to the channel decoder, that unit is already in possession of information from the RTCE analyser which can be used to optimise the channel decoding procedures in advance based upon an accurate knowledge of the channel state which is just about to occur. The longer the delay that can be tolerated, the greater the potential effectiveness of the RTCE process.

Attention will now be turned in the next section to architectural and control considerations for systems implemented using DSP devices and algorithms.

IMPLICATIONS OF DSP-BASED ARCHITECTURES

For economic reasons, it is probable that the digital processing power available within the architecture of a modern radio system terminal will be limited. It is therefore important that this finite processing power is employed in the most effective manner possible. The processing power is likely to be of two types:

(i) a PC-based component, normally better suited to control functions;

(ii) a real-time DSP component, normally used for signal generation and processing.

A typical schematic architecture for a radio system terminal comprising these two components is shown in Fig. 10.

Functions which might be performed by the PC component include:

(a) protocol generation;

(b) signal routing;

(c) TX/RX mode and parameter control;

(d) running embedded propagation analysis models [17];

(e) control of channel encoding/decoding functions in a multi-functional framework;

(f) RTCE and trend analysis.

The real-time DSP component would tend to be dedicated to (e) above, and would have its power distributed optimally between the functions of synchronisation, error control coding/decoding, modulation/demodulation, multi-user coding/decoding, etc in response to channel state and operational requirements, as indicated in the previous section. The DSP component could also be used for active and passive RTCE when not in use for other purposes, eg passive monitoring.

There is thus an overall optimisation problem to be solved in a systematic manner. How should the available processing power be employed to best meet the user's requirements at any time? For example, what proportion of time and processing power should be spent in actually generating and processing signals, and what proportion spent in say passive monitoring and RTCE so that the transmission and reception process can be more efficient when it does occur? Similarly, more sophisticated and adaptive transmission protocols, whilst consuming more processing power, may actually reduce the transmission time required and hence improve the security of the system. These are problems which are more difficult to solve, and yet have greater significance, with highly variable MMPPs.

CONCLUDING REMARKS

This paper has sought to discuss some of the problems inherent in the design of radio communication systems to operate efficiently over MMPPs in the frequency range up to about 100 MHz. The problems are associated primarily with:

(i) path characterisation and modelling, including RTCE;

(ii) the nature and combination of adaptive channel encoding/decoding techniques appropriate for such paths;

(iii) the optimum use of digital processing and specialised DSP devices within the overall architecture of the radio system terminal.

Of particular interest is the apparent usefulness of the low-frequency troposcatter mode of propagation in the MMPP context.

Detailed research is currently being pursued by the HWCRG into these problems and will be reported, as appropriate, at future AGARD meetings.

REFERENCES

1. CCIR: "Ground-wave propagation curves for frequencies between 10 kHz and 30 MHz", Rec. 368-3, ITU, Geneva, 1978.

2. Davies, K.: "Ionospheric radio propagation", National Bureau of Standards (USA) Monograph No. 80, April 1965.

3. Whitehead, J.D.: "Survey of sporadic-E processes", Space Res. VII, North-Holland Publ. Co., Amsterdam, 1967.

4. Oetting, J.D.: "An analysis of meteor-burst communication for military applications", IEEE Trans., Vol. COM-28, No. 9, September 1980.

5. Bartholome, P.J. and Vogt, I.M.: "Ionoscatter communications: new design concepts and experimental results", SHAPE Technical Centre Report TR-53, April 1965.

6. Roda, G.: "Troposcatter radio links", Artech House, 1988.

7. Rotheram, S.: "Radiowave propagation in the evaporation duct", Marconi Review, Vol. 37, No. 192, 1984.

8. Forsse'n, K-G.: "Troposcatter data transmission tests at 30 MHz", Proc. Nordic Shortwave Conf. "HF-89", Faro, Sweden, August 1989.

9. Darnell, M., Riley, N.G. and Melton, D.: "Beyond line-of-sight radio systems in the low-VHF band", Proc. 2nd Bangor Symp. on "Communications", Bangor, Wales, May 1990.

10. Bartholome, P.J. and Vogt, I.M.: "COMET - a new meteor-burst system incorporating ARQ and diversity reception", IEEE Trans., Vol. COM-16, No. 2, April 1968.

11. Darnell, M. and Honary, B.: "Multi-functional coding schemes applicable to secure communication", Proc. IEE Int. Conf. on "Secure communication systems", London, October 1986.

12. Darnell, M. and Honary, B.: "New synchronisation techniques applicable to multi-tone transmission systems", AGARD CP-442 on "Propagation effects and circuit performance of modern military radio systems with particular emphasis on those employing band-spreading", Paris, 1988.

13. Darnell, M., Honary, B. and Zolghadr, F.: "Embedded coding technique: principles and theoretical studies", Proc. IEE-F, Vol. 135, No. 1, 1988.

14. Honary, B. and Darnell, M.: "Embedded modulation and coding for HF channels", IEEE Int. Symp. on "Information theory", San Diego, USA, January 1990.

15. Tseng, C.C. and Liu, C.L.: "Complementary sets of sequences", IEEE Trans., Vol. IT-18, No. 5, 1972.

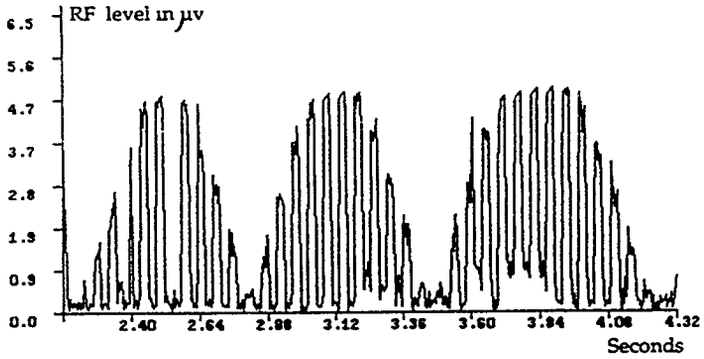
16. Darnell, M.: "Embedded real-time channel evaluation techniques", AGARD LS-145 on "Propagation impact on modern HF communication system design", Belgium/ France/Denmark, April 1986.

17. Uffelman, D.R. et al: "Real-time update of two HF channel evaluation models by oblique sounding", Report No. 5246, Naval Research Lab., Washington D.C., 1984.

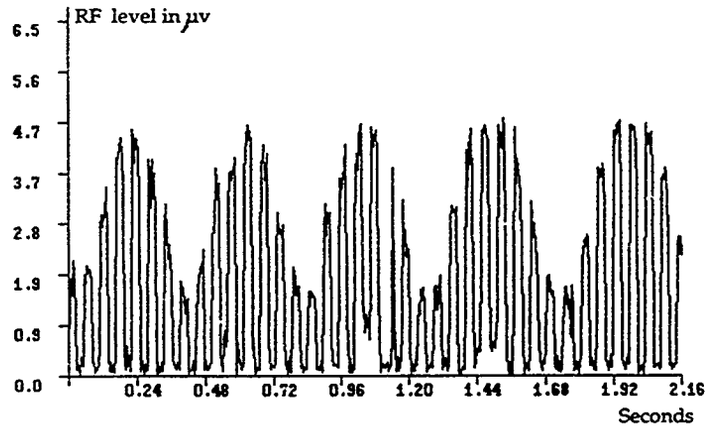
TABLE 1

Mechanism	Range	Variability	Availability
Surface wave (more efficient at lower frequencies)	500 km (over sea) (depends on ground constants)	amplitude: small phase: some	high
Normal skywave (via ionospheric refraction)	out to world-wide (depends on ionospheric properties)	amplitude: high phase: high frequency: small	variable (depends on system and path parameters)
Sporadic-E layer (sporadic ionisation at E-layer height)	out to about 2000 km	amplitude: high phase: high frequency: small	variable (depends on location, season, etc)
Meteor-burst (ionised meteor trails at about 100 km)	from about 250 to 2000 km	amplitude: high phase: variable frequency: variable (fading in tail)	high - but intermittent (1% duty cycle)
Ionospheric scatter (ionospheric irregularities at about 100 km)	from about 250 to 2000 km	amplitude: high phase: high frequency: small (low level due to scatter loss)	high - if ERP* high
Tropospheric scatter (tropospheric irregularities at a few km)	out to about 400 km	amplitude: high phase: high frequency: small (low-level due scatter loss)	high - if ERP* high

*Note: in the above Table, the abbreviation ERP indicates "Effective Radiated Power".

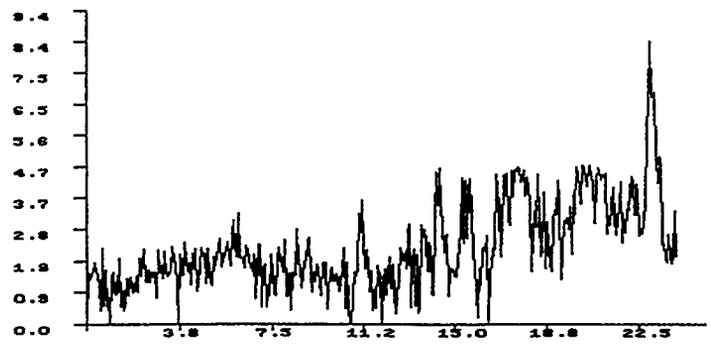


28-9-90 Newbury to Hull Short term fading characteristics

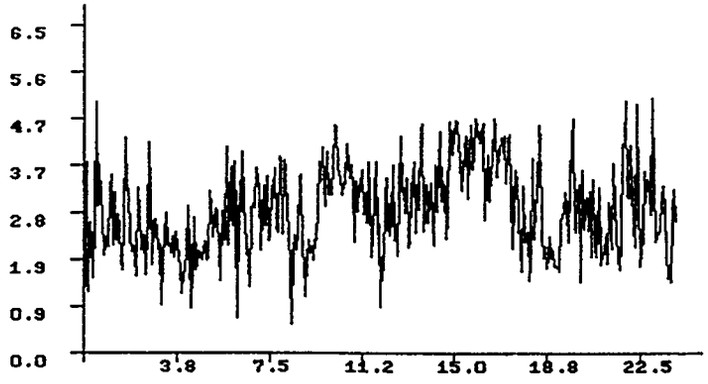


3-10-90 Newbury to Hull Short term fading characteristics

Figs. 1 & 2 Typical short term signal level variability for troposcatter signal



30-9-90 Newbury to Hull RF level in µv



1-10-90 Newbury to Hull RF level in µv

Figs. 3 & 4 Typical diurnal variability of signal level for troposcatter signal

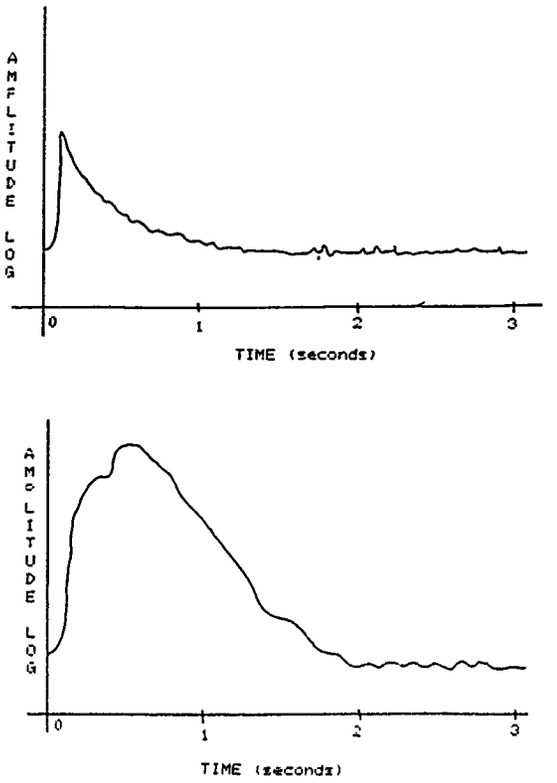


Fig 5 Typical meteor-burst profiles

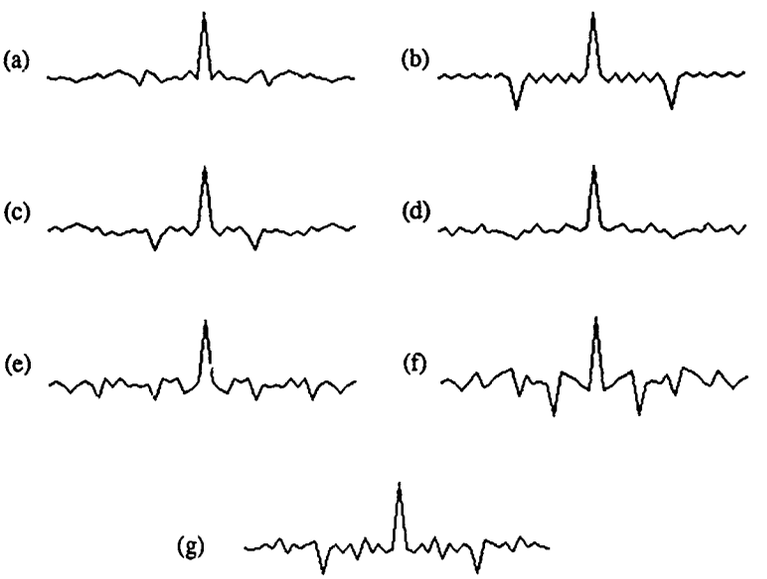


Fig. 6 Autocorrelation functions of the seven test sequences.

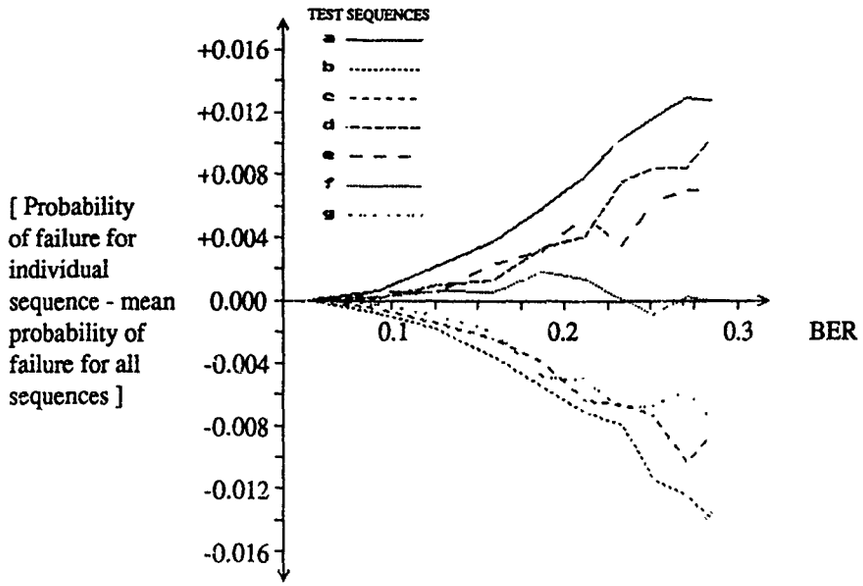


Fig. 7 Practical results showing the probability of failure for each sequence compared with the mean probability of failure.

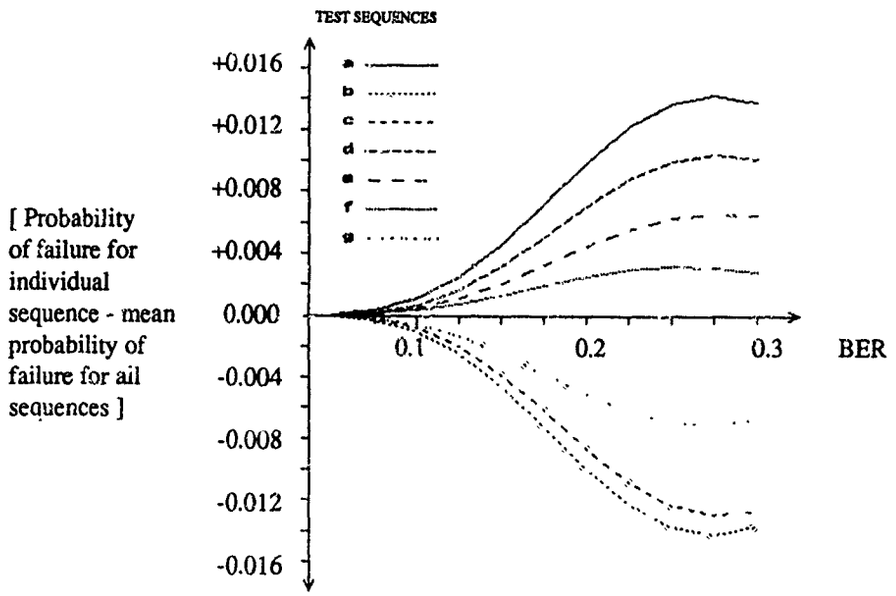


Fig. 8 Theoretical results showing the probability of failure for each sequence compared with the mean probability of failure.

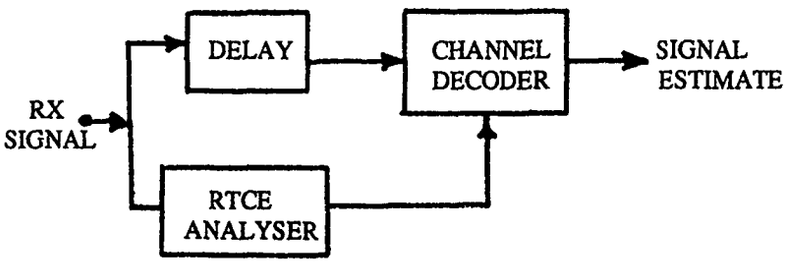


Fig. 9 RTCE with delay

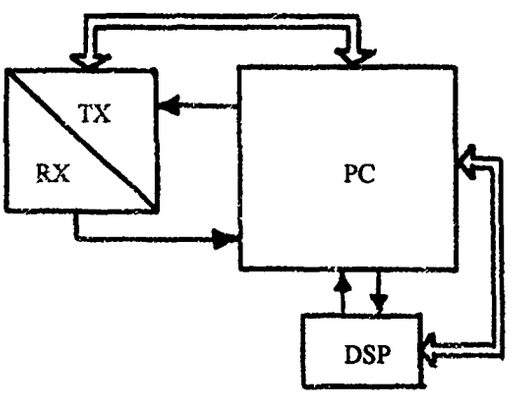


Fig. 10 Typical DSP-based architecture for radio system terminal

DISCUSSION

E. ANAGNOSTAKIS

Prof. DARNELL, could you, please, comment on the specific type of path metric which should be made use of in order properly to define the distance with respect to the optimal path, for a certain propagation channel ?

AUTHOR'S REPLY

The path metric described was simply the "distance" between the most likely path and the next most likely path. This was used as a simple illustrative example only.

I agree with the implication of the question. It may well be possible to use different metrics for different types of channel. As yet, we have not studied how this could be achieved in an adaptive system.

M. DARNELL

Could you please elaborate on how channel status can be estimated from measuring the average number of signal zero crossings ?

AUTHOR'S REPLY

At the demodulation output of, say, an MFSK transmission system, there will be present a number of tones at different frequencies, perturbed by noise. This noise will cause the position and number of zero-crossings per unit time interval to vary. The average number of zero-crossings per second has been analytically related to signal-to-noise ratio and can therefore be used as a channel evaluation measure when, in turn, can be used to adapt the transmission format.

H. DE PEDRO

MB and troposcatter channels require different protocols, such as continuous soundings and burst transmissions during bursts as opposed of continuous transmissions for troposcatter. How do you take advantage of randomly occuring bursts in XXX system ?

AUTHOR'S REPLY

I believe that some compromises would have to be made and that the signal format and protocols would not necessarily be optimum for all mechanisms. The most promising format would appear a relatively short block format with

frequent resynchronisation to cater for the different propagation delays associated with the different mechanisms. A single radio system terminal for all mechanisms is envisaged.

C. GOUTELARD, FR

Vous mentionnez l'utilisation de codes convolutionnels poinçonnés. Cette technique est utilisée dans les systèmes adaptatifs car ils permettent d'adapter la redondance du code.

Pouvez-vous nous dire les caractéristiques des codes que vous utilisez ?

Dans le canal HF, avec-vous fait des expérimentations et pouvez-vous nous dire comment vous avez choisi la longueur de contrainte ?

AUTHOR'S REPLY

The convolutional codes discussed briefly is the presentation were primarily to illustrate the concept of "embedded coding". They have only been used under simulated channel conditions.

We have, however, implemented a 3-level embedded block code and tested it over a meteor-burst channel. The main reason for this choice was one of implementational convenience.

In future, we intend to test and compare block and convolutional embedded codes over real HF channel : as yet, we have no performance results for such codes.

A Reliable Multi-User Distributive HF Communications System Using Narrowband CDMA.

T.E.Miller
Hull Warwick Communications Research Group
Department of Electronic Engineering
The University of Hull
Hull
HU6 7RX
UK

SUMMARY.

This paper describes the principles, design and application of a new type of code-division multiple-access (CDMA) communications system. In such a multi-user communications system, the two main constraints placed upon the codes used are that they should have inverse repeat time domain properties and that their periods should be in the ratio $1:2:4:....2^n$. This ensures that the line spectral components of the transmitted sequences do not coincide, and thus the codes are uncorrelated in the time domain. It is also important for these codes to have an autocorrelation function (ACF), which approximates to an impulse, so they can be conveiniently detected using matched filters. An HF multi-user system has been developed which makes use of multi-level pseudorandom sequences, particularly p-level m sequences. The clock rate of these sequences is adjusted so that the major spectral energy is transmitted within a nominal HF 3kHz bandwidth. Using relatively long duration (tens of milliseconds) sequences provides a high inherent resistance to noise, particularly white Gaussian noise, and fading. The system is designed around a Personal Computer based Digital Signal Processor (TMS320C25), giving both the receiver and transmitter considerable flexibility.

The system has been tested initially with an HF simulator using sinusoidally transformed p-level m sequences (these have an impulsive acf); other multi-level signals, such as integrated multi-level m-sequences or recorded noise, will also be tested. Initially, data has been transmitted by inverting sub-blocks within each sequence, the sub-block length being linked to the channel quality via RTCE. These sequences are stored at the receiver and the sub-blocks are 'intelligently' reinverted once all the data has been received. Since any encoded data must not effect the inverse repeat nature of the sequence, the sub-block inversion pattern must also be repeated in the second half of every sequence, thus providing the data retrieval algorithms with a degree of time diversity in the decoding process. In the paper, the techniques involved in developing this system are be discussed in detail and some preliminary trial results presented.

It is anticipated that the communication system design methodolegy described in the paper will provide the basis of a robust, distributed radio system with a capacity for channel sharing between a number of independant users.

1. INTRODUCTION: SYSTEM CONCEPT.

In a code division multiple access (CDMA) system, different codes are assigned to different users, each having a bandwidth comparable to the channel. Figure 1 shows a block diagram of a conventional CDMA system architecture. An ideal CDMA system uses perfectly uncorrelated sequences with ideal impulsive autocorrelation functions (ACFs). Inpractical CDMA systems, the sequences used are generaly periodic with very long cycle times. Data is typically transmitted by inverting/not inverting segments of these long sequences; the cross-correlation function (CCF) and ACF properties of the sequences over one data bit period determine the effectiveness of the approach. Integrate and dump processing is typically employed to recover the data at the receiver.

This paper will be concerned with the generation of completely uncorrelated sequence sets with impulsive, or near impulsive, ACFs for use in an in-band CDMA system. A novel method of data encoding for these bearer sequences will also be discussed.

2. CONDITIONS FOR UNCORRELATEDNESS.

In synthesising a set of completely uncorrelated sequences, two basic conditions must be met: for an K-sequence set (a) the format of the sequences transmitted by all users (except possibly the shortest) should be inverse repeat, and (b) the periods of the sequences should be in the ratios $1:2:4:...:2^{K-1}$. These two conditions ensure that each sequence occupies its own *slots* in the frequency spectrum, first by ensuring that only odd frequency components are present for each inverse repeat sequence and secondly by setting the periods so that these frequency components are interleaved in the spectrum. In this way, no spectral components for any sequences in the set coincide.

3. AVAILIABLE SEQUENCE TYPES.

A number of different types of periodic sequence structure have been evaluated for use in CDMA sets. These are described briefly below.

3.1 Multi-Level m Sequences

Any m-sequence has length N, with

$$N = p^n - 1 \qquad [1]$$

where p is a prime number and n is the number of stages in the eqivalent feedback shift-register generator.

A p=11, n=2, N=120 sequence was chosen to investigate the properties of p-level m-sequences. This sequence is, considerably shorter than those which may finally be used, but was chosen for simplicity of analysis.

To produce a bi-polar output, the sequence must be transformed so that the output values are evenly distributed around the zero level. The CCF between two or more transformed sequences, satisfying the conditions above, is required to be zero. The most commonly used transfomation is the linear

transformation below:

Sequence levels	Transformed levels
0	0
1	+(p-1)/2
:	:
:	:
(p-1)	-(p-1)/2

Figures 2 (a) and (b) show the p=11, N=120 sequence with the above transformation applied, together with its ACF. Any function which yields the inverse repeat structure can be used as a level shifting function for p-level m-sequences. The transform in equation [2] is the *sine* transform; this has similar level shifting properties to the linear transform, but suppresses the sidelobes in the ACF - see Figure 2 (b) and (c).

$$\text{Transformed level} = \sin((2\pi q)/p) \qquad [2]$$

Where:

$$0 <= q <= (p-1)$$

It can be seen that the levels in the sine transformed sequence tend to be compressed at the extreme values of the output. This is expected as the probability density (pdf) function of a sinusoid also exhibits this property. This is advantageous in a communications system because it increases the overall signal-to-noise ratio of the received signal by increasing the average output power of the transmitter. Obviously, this transform causes the transmitted signal to be less noise-like as its pdf is no longer approximately Gaussian.

3.2 Integrated p-level m-sequences.

Integrating a p-level m-sequence transforms it to a ramp-type sequence [2]. This can be seen in Figure 3 where (a) shows a ternary 80 digit sequence (p=3, n=3, N=80), (b) shows the ACF of this ternary sequence, and (c) shows the integrated form of this sequence. This integrated form has the ACF shown in Figure 3(d); it can be seen that this is not impulsive and thus would have little use as the 'ideal' output response of a detector in a CDMA communications system. However these multi-level integrated sequences can be processed to cause their effective ACF to become impulsive, and thus be of use in such a communications system.

Let a 'clock-multiplied' signal be defined as the resultant of the multiplication of a given p-level m-sequence by its own clock waveform. See Figure 3(e). If Δt is the clock period of the original signal, then the effective clock period of the clock-multiplied (CM) signal is $\Delta t/2$. The ACF of the CM signal, $A_{\infty}(r)$, may now be expressed in terms of the ACF of the original signal, $A_{11}(r)$, and that of the clock waveform, $A_{22}(r)$; thus:

$$A_{\infty}(r) = A_{11}(r) \,.\, A_{22}(r) \qquad [3]$$

where r is a discrete shift, variable in unit steps of $\Delta t/2$.

The single period ACF of the N p-level m-sequence can be expressed as follows.

$$A_{11}(r) = \begin{cases} +P & ; r = 0 \\ +P/2 & ; r = \Delta t/2 \text{ or } (2N-1)\Delta t/2 \\ -P & ; r = N\Delta t/2 \\ -P/2 & ; r=(N-1)\Delta t/2 \text{ or } (N+1)\Delta t/2 \\ 0 & ; \text{otherwise} \end{cases} \qquad [4]$$

where P is the magnitude of the ACF peak.

For the clock waveform, the normalised single period ACF is:

$$A_{22}(r) = \begin{cases} +1 & ; r = 0 \\ -1 & ; r = Dt/2 \end{cases} \qquad [5]$$

Therefore, substituting expressions [4] and [5] in [3]:

$$A_{\infty}(r) = \begin{cases} +P & ; r = 0 \\ -P/2 & ; r = \Delta t/2 \text{ or } (2N-1)\Delta t/2 \\ -P & ; r = N\Delta t/2 \\ +P/2 & ; r=(N-1)\Delta t/2 \text{ or } (N+1)\Delta t/2 \\ 0 & ; \text{otherwise} \end{cases} \qquad [6]$$

This function for a ternary m-sequence (N=80) can be seen in Figure 3(f). It should be noted that the CM signal has 2N digits in its period.

Now considering the CCF between the CM signal and the original p-level m-sequence: when the relative phase shift between the two signals is zero, or an even multiple of $\Delta t/2$, the CCF will be zero, since each digit of the original signal is multiplied by two digits of the CM signal having equal magnitude and opposite polarity. When the relative phase shift is an odd multiple of $\Delta t/2$, the CCF is:

$$A_{\alpha}(r) = 1/2 \,[\, A_{11}(r-1) - A_{11}(r+1) \,] \qquad [7]$$

where r is an odd multiple of $\Delta t/2$.

For sequences where p>2, $A_{\alpha}(r)$ will have the form shown in Figure 3(g). This function is a close approximation to the first derivative of $A_{11}(r)$. Such a function can be obtained by cross-correlating the original signal with its own time derivative. Hence the CM signal can be considered as being equivalent to the first time derivative of the signal from which it was obtained.

From the above, it is seen that the CCF is effectively being performed between integrated and differentiated versions of the same m-sequence. This is implemented by matched filtering the integrated signal with its CM form; Figure 3 (h) shows this result for the ternary sequence above. Note that the filter response now has the desired impulsive characteristics.

3.3 Binary m-sequences.

For a binary m-sequence, the sequence length N is always an odd number; this implies that the sequence can not have the inverse repeat format required for sets of sequences to be uncorrelated in an ideal CDMA system. When a binary m-sequence is followed by its inverse, the resulting sequence, of length 2N digits, has the desired inverse repeat format. The ACF for such a sequence (for N=255) is shown in Figure 4. The ACF is 'impulse like', although the sidelobes are numerous and of a relatively high amplitude. This signal could be used in the multi-user system, although its detection will be somewhat more complex than for p-level (p>2) m-sequences and integrated p-level m-sequences.

3.4 Noise signals.

A large set of samples of Gaussian white noise will, by definition, have an impulsive ACF and also, will be uncorrelated with any other set of samples of Gaussian white noise. Thus a set of samples recorded from a Gaussian source, would be ideal for a CDMA system, since the noise

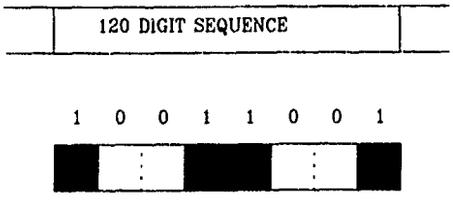
sets could be followed by their inverses and thus fulfill the inverse-repeat criterion. It should be noted that the characteristics of any sampled set of Gaussian white noise will be changed due to the sampling process, as the effective spectral bandwidth will be determined by the sampling rate.

A receiver was used to scan the HF (2-30MHz) radio frequency spectrum for an unused band; this proved impossible as every frequency channel examined had an uneven energy distribution, thus showing that there was always a small amount of signal in any channel. Suitable approximations to 'noise like' signals were recorded into computer memory and analysed. Figure 5 (a) shows a recorded noise signal, and (b) and (c) show the ACFs for 200 samples and 2500 samples of this signal respectively. These plots clearly show that the longer sample set has a more impulsive ACF. Figure 6 shows the signal and ACF for 200 samples from a wideband noise generator; this signal as expected has a more impulsive ACF than that of the recorded noise. Finally Figure 7 shows the signal and ACF for a music signal; the ACF has substantial sidelobes, although these have very distinct characteristics and could thus be searched for at the receiver, making the sequence again usable for data communications.

4. SYSTEM DESIGN.

4.1 Data Encoding.

Data can be imposed upon a sequence by inverting/not inverting blocks within the sequence, each block would thus represent a binary data bit. To retain the inverse repeat properties of the sequence (essential to CDMA) the second half of the sequence must be block inverted in the same way as the first half. For example the 11 level 120 digit sequence (transformed using either of the above mentioned methods) could be divided up into eight data blocks, each fifteen digits in length, and could thus convey four bits of information, the word 1001 being encoded as 10011001 as shown below where a shaded block represents a block inversion.

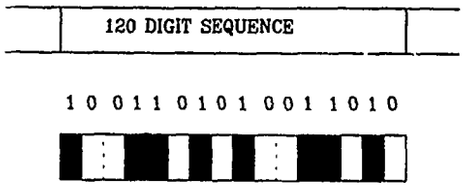


The cross correlation between the sixteen posible combinations of block inversion for this sequence and the original 120 digit sequence are shown in Figure 8. It can be seen that only the pure sequence or its inverse have an impulse result.

Using this method of data encoding means that redundancy, in the form of time diversity is inherent in the system. This could be used at the receiver to correct errors which only occur in one half of the sequence.

4.2 'Fractional-digit' Data Encoding.

It is possible using this technique of data encoding that the length of each data block, which may or may not be inverted, could be a non integer number of sequence digits in length. Using the 120 digit 11 level sequence as an example eight bits each of block length 7.5 samples could be conveyed by each sequence. The word 10011010 would thus be encode as 1001101010011010 as shown below.

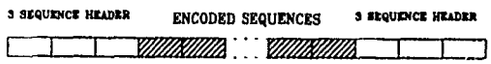


4.3 Data Decoding.

Data can be simply decoded by systematicly reinverting blocks within the received sequence and then cross correlating the result with the refference sequence. The data will be the combination of blocks which are reinverted that give the largest in phase correlation responce. It should be noted that the accuracy of the decoder is dependant on the rate at which the receiver samples the data, particularly if fractional-digit encoding is used since the nyquist sampling rate would have to be calculated from the fractional rate rather than the digit rate.

4.4 Data Format.

In the current system data is transmitted in the format shown below. The pure header sequences are used by the receiver to detect the begining and end of transmission, three are required since the sequences have periodic correlation properties.

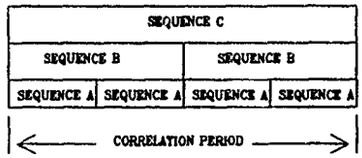


4.5 Transmitter Design.

The transmitter consists of an IBM compatible personal computer containing a TMS320C25 digital signal processor card and an HF tranceiver. Data taken from the keyboard is encoded ready for transmission and stored as a data file. This file is then transmitted, in the format shown above, by the signal processor through the transceiver.

4.6 Receiver Design.

For each sequence to be uncorrelated at the receive the periods of all the correlation filters must be equal to the period of the logest sequence, thus the inverse repeat properties of the sequences cause a zero responce from any user other than that beig filtered for. This is shown below.



The receiver uses the same equipment as the transmitter. The signal processor continuously samples the input from the tranceiver, when the correct header sequence is received the data is stored onto disk untill the header sequence is again received as an end of data signal. The flow diagram for this decision routine is shown in Figure 9. The stored data is then decoded by the signal proccesor and displayed for the user.

5. INITIAL TRIALS.

During the initial field trials the header signals were transmitted from Wick (Scotland) and recorded at Cobett Hill (London) and then later played into the receiver program to test its performance. Figure 10 shows the response of the header detection algorithm to the transmitted header sequence and the ideal response of the correlation filter used as part of the detector. The detector inverts all negative spikes and then looks for four spikes each with a separation of half a sequence period. A novel method of threshold detection is used to decide if a spike is pressent or not. Firstly the side lobes from the filter response (created by noise and interference) are squared and averaged over just under half a sequence period. This value is then scaled and subtracted from the filter response, this causes the result to be offset by a negative level, if there is any noise or interference present. The short time when this algorthm is zeroed can be seen as a 'bump' just before the spike on the plot, this is timed to occur when there is no spike present and thus does not effect the result. The actual threshold is obtained by an envelope detection routine and thus allows the receiver to have a dynamic range of approximately 40dB.

5.1 Initial Single User Tests.

To test the complete communications system, particularly the decoding algorithms, the configuration in Figure 11 was set up.

In the first test it was found that the decoder lost sequence synchronisation with the block inverted sequences stored on disk, this is due to a slight frequency drift on the clocks on the different TMS boards. This problem was easly overcome by designing the decoder to step a variable number of stored samples either way of the file location where synchronisation is expected and then calculating at which point the decoded result gives the highest confidence level (i.e. The file pointer is moved and the result is recalculated, then the file pointer is set to the point where the highest confidence is obtained). This technique thus makes the decoding software self-synchronising for any length of file.

The following results were obtained from a small data file (the lower case alphabet) and were primarily used to detect any short comings in the decoding algorithms although they give a good indication of how the system will perform in more thorough tests, which will be carried out over the next few weeks.

Resistance to Noise : With these short data files transmitted at 100 bits per second the decoding software reliably decoded all data that was detected by the header detection routine. Thus the header detection routine limited the amount of noise the system could operate in although it performed reliably in signal to noise ratios below -20dB.

Rayleigh Fading : Using the HF simulator the system was found to work reliably in fade rates in excess of 100 fades per minute.

Frequency Off-set : Figure 12 shows that the system requires the receiver to be tuned in to an accuracy of approximately +/- 1Hz and the number of characters in error that could be expected if this was not acheived.

6. CONCLUSION.

It has been shown that the sequences and signals described previously can essentially provide the 'bearer' signals in a CDMA communications system. A technique for modulating these bearer signals with information has also been discussed. The initial tests with this system show that it exhibits ideal properties for a reliable multi-user HF communications system. Simulated and field trials will be carried out to further demonstrate the effectiveness of these techniques over shared 3kHz-bandwidth HF single-side band channels.

7. REFERENCES AND BIBLIOGRAPHY.

1. Darnell, M: "The Theory and Generation of Sets of Uncorrelated Digital Sequences", The Institute of Mathematics and ITS Applications Conference Series: "Cryptography and Coding", Clarendon Press 1989.

2. Darnell, M: "Multi-level Pseudorandom Signals For System Evaluation." A Dissertation Submitted for the Degree of Doctor of Philosophy of the University of Cambridge 1968.

3. Golomb, S.W.: "Shift Register Sequences", Holden-Day 1967.

4. Golomb, Baumert, Easterling, Stiffler, Viterbi: "Digital Communications with Space Applications", Prentice-Hall 1964.

5. Shaar, A. A., Davies, P. A.: "A survey of one-coincidence sequences for frequency-hopped spread-spectrum systems.", IEE Proceedings, Vol. 131, Pt. F, No. 7, December 1984.

8. ACKNOWLEDGEMENT.

The authors gratefully acknowledge the assistance and financial support of RAE(Farnborough).

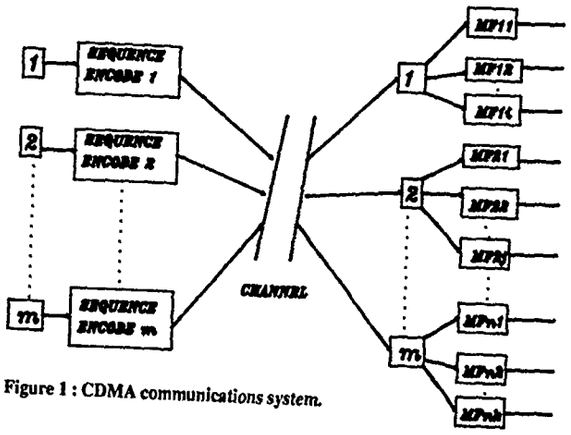


Figure 1 : CDMA communications system.

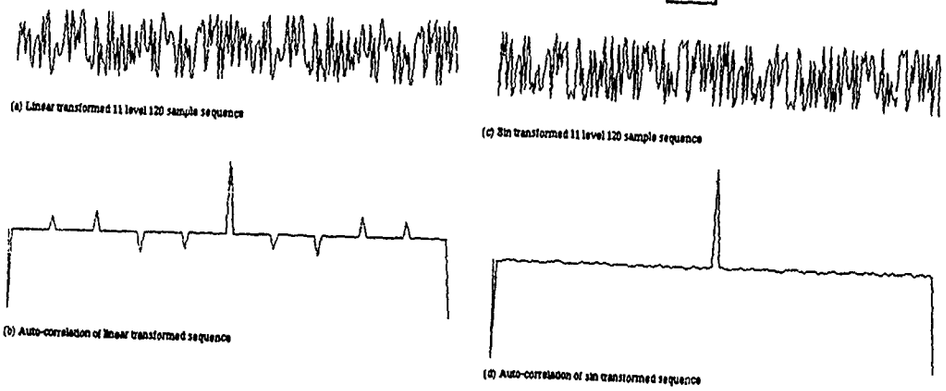


Figure 2 : Transformed p-level m-sequences (p=11, m=120) and their ACFs.

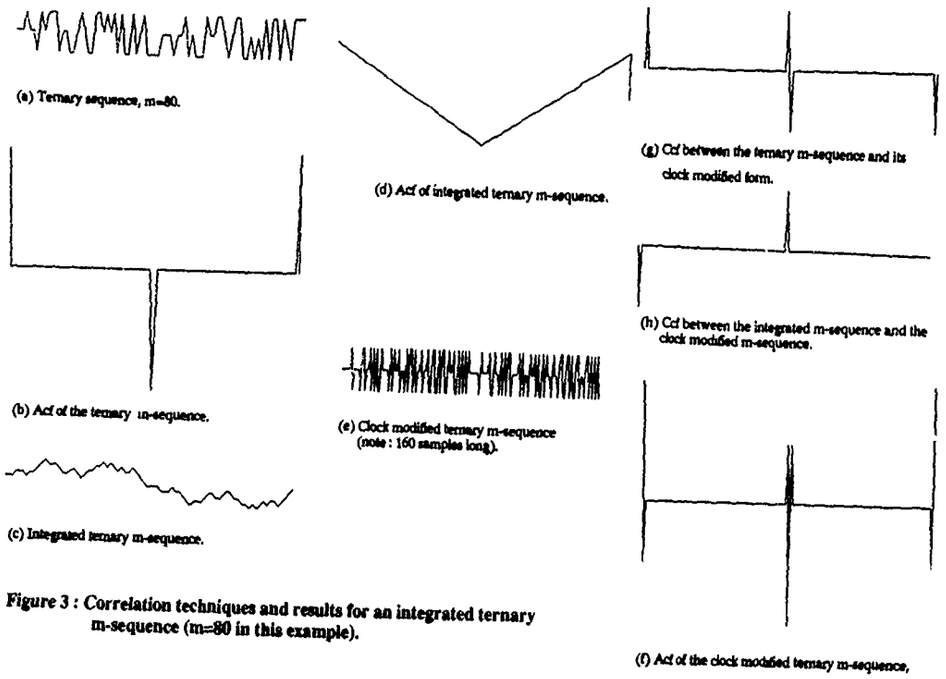


Figure 3 : Correlation techniques and results for an integrated ternary m-sequence (m=80 in this example).

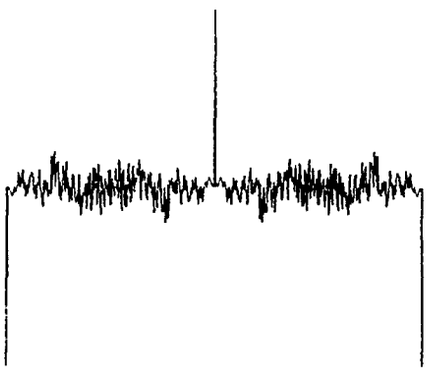
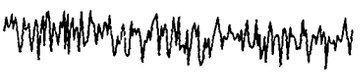


Figure 4 : ACF of binary m-sequence.



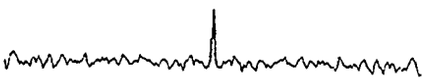
(a) Noise signal.

(b) Acf of 200 samples.

(c) Acf of 2500 samples.

Figure 5 : Results for recorded noise.

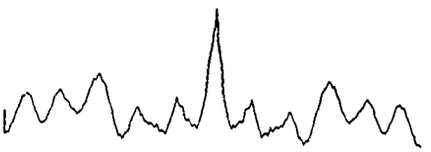
(a) Generated noise signal.



(b) Acf of 200 samples.

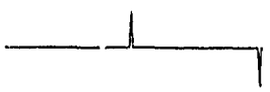
Figure 6 : Results for recorded, generated noise.

(a) Music signal.

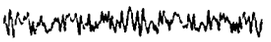


(b) Acf of 200 samples.

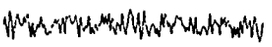
Figure 7 : Results for recorded music.



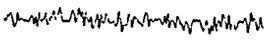
(a) 00000000



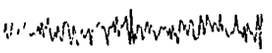
(b) 00010001



(c) 00100010

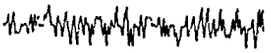


(d) 00110011

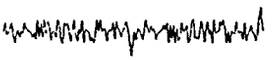


(e) 01000100

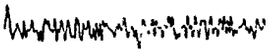
(f) 01010101



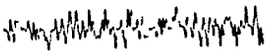
(g) 01100110



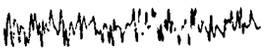
(h) 01110111



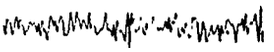
(i) 10001000



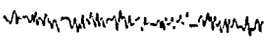
(j) 10011001



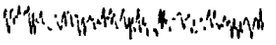
(k) 10101010



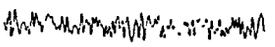
(l) 10111011



(m) 11001100



(n) 11011101



(o) 11101110

(p) 11111111

Figure 8 : Correlation results for the sixteen block inversion combinations of the integrated ternary m-sequence.

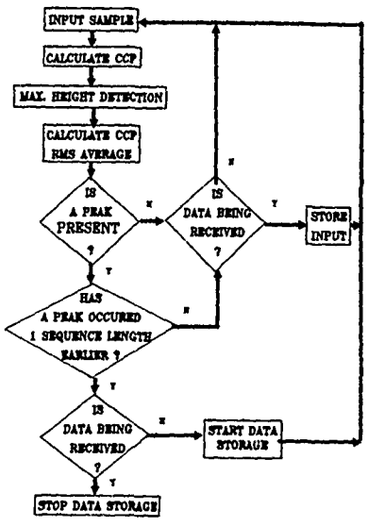


Figure 9 : Receiver flow diagram

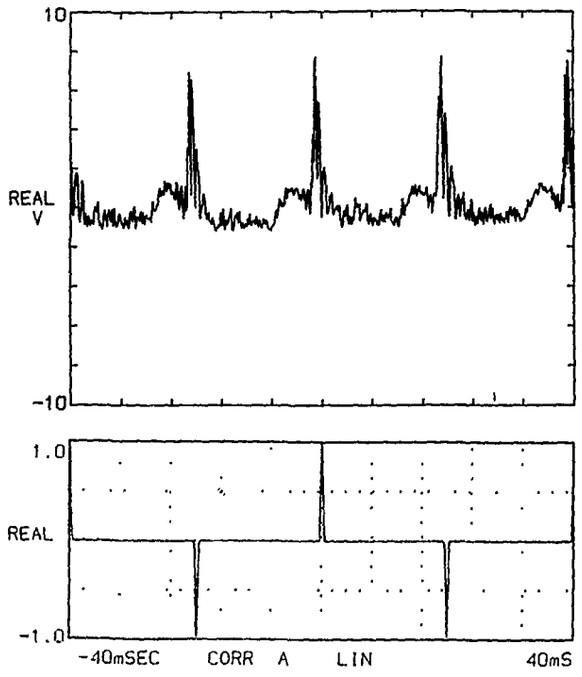


Figure 10 : Correlator response to trial results and theoretical correlation result.

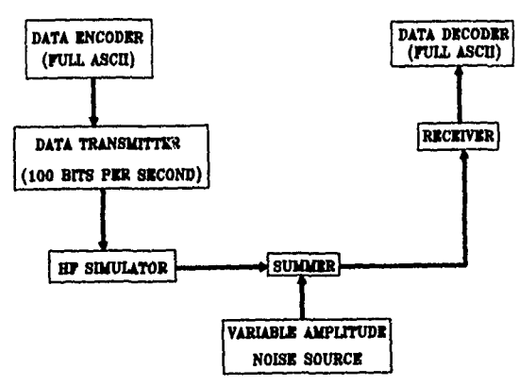


Figure 11 : Initial single user test configuration.

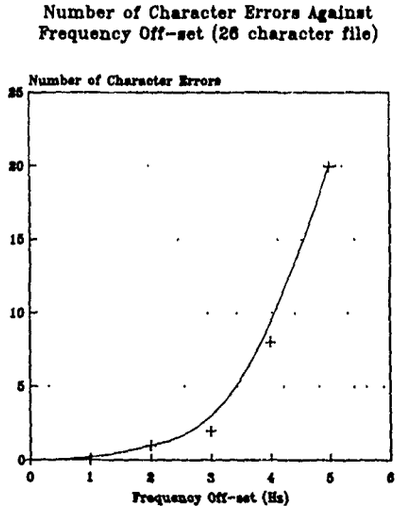


Figure 12 : Receiver sensitivity to frequency off-set.

DISCUSSION

D. YAVUZ

How many users do you anticipate accomodating within the baseband for "multiple access" as implied in the title of your paper (CDMA) ?

AUTHOR'S REPLY

I currently anticipate that the system will support 5 users since a larger number of users could cause the spectral lines to coincide due to frequency shifts within the HF channel. It is likely that there frequency shifts would not be uniform for all paths, and thus a larger number of users may be possible.

C. GOUTELARD, FR

Votre exposé m'a beaucoup intéressé. Vous présentez des séquences dont les longueurs varient selon une loi exponentielle en fonction du nombre d'utilisateurs, ce qui en particulier resserre les raies spectrales et rend le système sensible à l'effet doppler.

Quel est le nombre d'utilisateurs que vous jugez possible et avez-vous envisagé une méthode pour augmenter ce nombre ?

AUTHOR'S REPLY

I currently envisage the system to have five users. It is possible that this number of users could be increased. The number of users is limited by the distance between the spectral lines of different users. If this distance (in Hz) is less than the possible channel frequency shift, errors will occur. It is possible that a greater number of users could be used since the frequency shifts on the different paths are unrelated.

A DESIGN TESTBED FOR DISTRIBUTED V/UHF NETWORKS WITH MOBILE TERMINALS

Andrew Carnegie
Hull-Warwick Communication Research Group
Department of Electronic Engineering
University of Hull
Hull HU6 7RX
UK

Summary

A flexible design tool for mobile radio networks has been implemented. Both low level areas, such as modulation, coding and synchronisation, and high level techniques involving protocol details can be analysed simultaneously. The simulation software is based on a Rayleigh fading channel model with a network layer protocol implementation. This paper describes how a Rayleigh distribution can be generated and how it is used to implement the channel simulation by a conversion to probability of error based on the modulation technique required. The network model is discussed with reference to the OSI Open Systems Interconnection reference model and the variable configuration is described. To assess the success of the system an investigation into the use of variable, optimum length packets is illustrated.

1. Introduction

Since radio networks were developed in the early seventies the vast majority have had a terminal hierarchy in which a few base station terminals have the responsibility of controlling the whole network. This system has persisted from the first Aloha network through the development of packet radio by ARPA [1] to the present day. However it has the disadvantage that both reliability and security depend heavily on the network stations. If these fail then network collapse can follow. In addition, due to the power and complexity required of a computer capable of controlling a whole network, the portability of such a system is limited. Less hierarchical systems overcome these problems and are becoming important for both military and civilian applications. With the advent of cheap processing power completely non-hierarchical networks comprising identical terminals utilising a distributed control approach are now possible. This approach is particularly suitable for relatively small networks where the flow of control information is low.

The control protocols, access methods and transmission schemes developed for networks with a terminal hierarchy are not necessarily the most efficient and effective in a truly mobile completely non-hierarchical network. In the design of such a distributed mobile radio network many factors have to be considered simultaneously. These include modulation and synchronisation schemes, error control strategies and control protocols. Hence the need for a flexible integrated simulation facility to allow assessment of these and other problems was foreseen.

Such a facility has been designed. It allows many different techniques to be tested under completely repeatable conditions and their effects on either the whole network, or a specific part of it, to be determined. Emphasis has been placed on simulating a distributed mobile network with line of sight links in the V/UHF region utilising a packet switched protocol. However the flexibility of the system allows any of the above conditions to be relaxed.

The simulation has two definite layers. The lower level is a channel simulation which allows different modulation and channel synchronisation schemes, error control codes, low level diversity strategies and error models to be imposed on given channel conditions. Built on this is the network simulation. Different control protocols and access schemes can be tested in a network with an arbitrary configuration.

2. Channel Simulation

One of the worst problems the mobile radio network designer has to face is the fading of received signal strength in urban areas. There is very rarely a direct line of sight path between two terminals, the signal is received after having been widely scattered by intervening buildings and other obstacles. This results in the received signal being the sum of waves arriving from many angles. The small differences in propagation time of each of the components cause severe fading at the receiver.

Tests have been carried out in urban areas world-wide [2,3,4,5] and all show that the short term amplitude variations of the received power obey a Rayleigh probability density function. In addition this result has been predicted by theory [6]. Since this is the worst type of fading encountered in mobile radio systems the channel simulation will be based on such a Rayleigh fading channel. Although techniques perfected under these conditions may not be the most efficient they will work in almost any other channel conditions.

To achieve the flexibility and repeatability required from the simulation the envelope of a Rayleigh faded wave was saved to disk. This has the advantage that the wave can be completely characterised along its entire length in detail before any analysis takes place. In addition it will shorten simulation time as the Rayleigh fading wave will not need to be generated continuously as a test progresses. The only drawback is that the length of the signal is finite and during long tests the fading may start to repeat. This has been overcome by saving a long length of the envelope and making sure that any channel simulations do not use the same area of signal data twice.

2.1 Generating a Raleigh Distribution

One of the definitions of the Rayleigh distribution is that it is the sum of two independent random functions with Gaussian probability density functions. This is useful to bear in mind when considering how Rayleigh fades are generated. The method used to generate a fading envelope is similar to the method of fading a signal described by Ralphs and Sladen among others [7,8,9,10]. To achieve fading the input signal is quadrature filtered to obtain equal amplitude quadrature phase signals Then each of these is amplitude modulated by independent random functions with Gaussian distributions. The resultant signals are then uniformly summed and the output is a version of the input signal which has been Rayleigh faded.

However to generate only the envelope of a faded signal is simpler than this, as seen in figure 2.1. Assuming that a sine wave

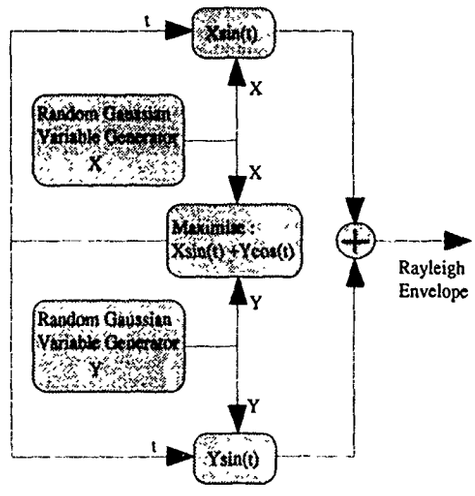


Figure 2.1 - Block diagram of method of generating Rayleigh fading envelope

is the input then a cosine wave is the quadrature phase component and the sum of the modulator outputs is :

$$y(t) = A\cos(t) + B\sin(t) \quad (1)$$

where A and B are random Gaussian variables. The envelope is the maximum value that this can have for every different value of A and B. Differentiating equation 1 gives :

$$\frac{dy(t)}{dt} = A\sin(t) - B\cos(t) \quad (2)$$

The value of t which maximises y(t) is hence given by:

$$t = \tan^{-1}(B/A) \quad (3)$$

Substituting this back into equation 1 will give the maximum value of y(t) and hence the envelope of the signal. This method requires only two independent Gaussian variables to be calculated and is hence relatively fast and simple to generate. An example section of the type of signal generated is shown in figure 2.2.

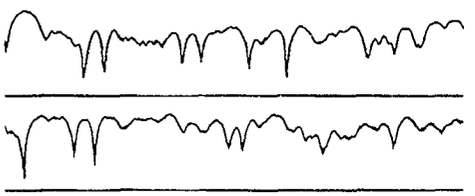


Figure 2.2 - Envelope of Rayleigh faded signal (log scale)

It is important that the faded data saved on disk does not repeat. The duration of the Rayleigh sequence generated before it repeats depends on the Gaussian sources used. If the lengths of the Gaussian sequences have no common divisor then the length of the Rayleigh envelope is their product. The sequences are generated using feedback shift registers to produce maximum length sequences. The shift registers chosen were 25 and 28 bits long. The sequences generated by these are 33,554,431 bits and 268,435,455 bits long respectively. The total length of the Rayleigh envelope is then the product of these two - 9,007,198,952,751,105 points. If the shift registers were to be clocked say 100,000 times per second then the sequence would take 9×10^{10} seconds or 2,817 years to end! Therefore an envelope can be generated which is as long as is required and will effectively never repeat.

2.2 Fade Rate

To conduct meaningful trials it is essential to be able to relate the data stored on disk to the fade rate which would be expected in an actual mobile channel. Bodtmann and Arnold [11] implement a similar type of fading generator and produce the figures that, assuming a mobile velocity of 30 mph and a carrier frequency of 900MHz, 2^{19}-1 points are equivalent to 131 seconds fading time. Of course this information has to be treated with caution and is rather rough but it is a good guide. Based on this one second is approximately equivalent to 4000 points. This information makes it possible to calculate the length of time the envelope data on disk represents.

Jakes [12] derives a useful expression for the mean rate of occurrence of fades below the amplitude R :

$$N(R) = \frac{\sqrt{2\pi}\, v\, p\, e^{-p^2}}{\lambda} \quad (4)$$

where p = R/Rr, Rr is the root mean square of the distribution, v is the mobile velocity and λ is the carrier wavelength. This shows that the fade rate is proportional to both velocity and carrier frequency. Together with knowledge of the number of points per second at a particular mobile velocity and carrier frequency these two parameters can be varied.

Equation 4 allows a rough comparison of the theoretical fade rate and the results obtained using 4000 points per second. If a fade is defined to be when the amplitude drops below the RMS value of the signal then the fade rate is found when p = 1. In addition 30 mph is 48.28 km/h or 13.4 m/s and the wavelength of a 900 MHz carrier is 0.3333 metres. Substituting these values into equation 4 gives a mean fade rate of 37 fades per second. Comparison with the number of fades produced by the generator is favourable, it gives between 35 and 40 fades per second.

2.3 Simulating a Channel

Generating the Rayleigh envelope described in the previous sections is only a single step toward simulating a channel. The question still remains of how to use the envelope. For the purposes of the network model the requirement is to be able to tell relatively quickly how many errors occur, if any, when a packet is transmitted over a given channel. Hence the envelope needs to be translated into channel conditions which are controllable and which can be used to determine if and where errors would occur. A brief outline of how this is achieved is that the fade envelope is assumed to be a received signal level then noise is added to produce a signal to noise ratio. This can be converted to a probability of bit error and the location of any errors detected.

2.3.1 Conversion of Fade Envelope to Signal to Noise Ratio

To save memory all faded points were saved as two byte integers instead of four byte real numbers. This involves scaling the points up to the maximum integer and rounding them. It does introduce a small amount of quantisation error but this is minimal and can be neglected. As points are read in they are scaled to give the required root mean square signal level before noise is added.

Addition of noise to the signal level is fairly flexible. A noise signal can either be calculated as the simulation progresses or can be precalculated and stored on disk in the same manner as the signal level. However when information on the amplitude variations of noise was sought very little accurate data on the ranges of interest could be found. Since equipment was available to make tests some brief trials were conducted. In urban areas interference from man-made sources is a severe problem. However this is extremely variable, time dependent and difficult to characterise in short term trials Samples of background noise were recorded at different frequencies though and the amplitude distributions of these were calculated. The results show that the noise fits a Gaussian distribution surprisingly well, a typical result is given in figure 2.3. Hence a Gaussian variable generator, similar to those used to produce values for the Gaussian modulators when generating Rayleigh fading, was implemented as a noise source.

When noise of a suitable frequency, variance and mean level has been generated it is divided into the signal level to produce an instantaneous signal to noise ratio.

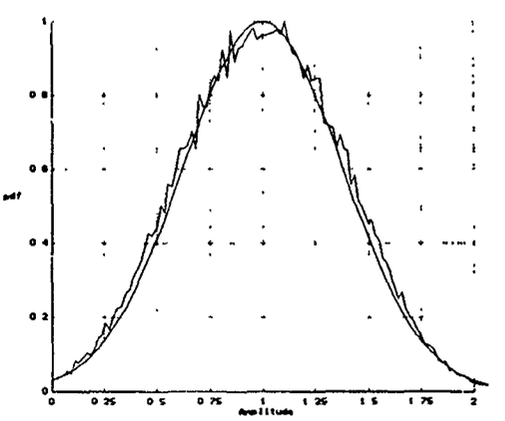


Figure 2.3 - Normalised distribution of noise with perfect Gaussian distribution superimposed.. Sampled at 22kHz at frequency of 201.2MHz.

2.3.2 Probability of Error

Conversion of the signal to noise ratio to a probability of bit error is performed using standard expressions for each type of modulation [13]. The equations relate steady signal conditions to a received probability of error under Gaussian white noise conditions. At present low level techniques such as FSK, ASK, PSK and DPSK have been implemented. The application of such expressions to each individual signal to noise ratio point gives an instantaneous probability of error. The suitability of the equations used for this purpose is not certain. If the signal to noise ratio of each point is assumed to be constant for a short length of time then it would seem reasonable to apply such steady state expressions to give a fixed probability of error even though it is for a very brief interval. The assumption of Gaussian white noise that these expressions include means that the addition of Gaussian noise to the signal level is not necessarily required. A constant noise level can be assumed which improves the speed of the simulation and reduces its complexity.

The probability of error is used to predict where bit errors will occur using Monte Carlo simulation, ie. generating random numbers and checking them against an error threshold. A minor complication arises when fade rate or the data transmission rate change. Except in exceptional circumstances there are always several packet bits to one fade point. Thus the number of bits per point needs to be calculated and the probability of error slightly altered. For example a typical transmission rate is 16kbit/s and at a mobile speed of 30mph there are 4000 fade points per second. Hence there are 4 transmitted bits per fade point. It is assumed that these bits will all suffer the same probability of error and it is required to find the probability that none are in error. This is done using the expression :

$$p(\text{no errors in B bits}) = (1 - p_e)^B \qquad (5)$$

where p_e is the probability of error, constant for all B bits. Thus errors can be detected as both the fade rate and transmission rate change. Figure 2.4 demonstrates the type of output produced by the channel simulation.

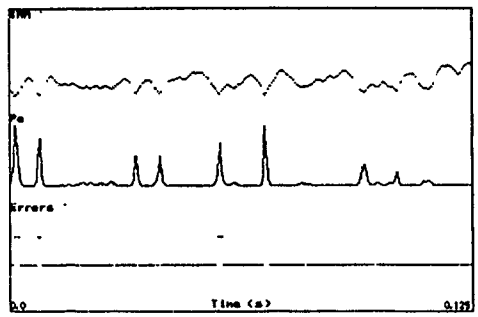


Figure 3.1 - Mapping of Signal to noise ratio to probability of error. RMS SNR is 10dB for a packet of length 2000bits at 16kbit/s and mobile velocity of 30mph. Modulation type assumed is FSK.

2.4 Summary of Channel Simulation

The main aim of the channel simulation was that it should be flexible and that all the parameters should be variable in a controllable manner. This has been achieved in a simulation that is relatively efficient and fast. It should also be noted that, although the channel is based on fading with a Rayleigh distribution and the generation of this has been described, it is possible to change the amplitude distribution of the received signal without altering the simulation. If a data file is generated containing data representing fading with Ricean, Gaussian, log normal plus Rayleigh or any other distribution it could be substituted in directly to replace the Rayleigh fade data. In time a library of channel conditions could be developed.

Despite the apparent flexibility of the system there are a number of factors which are not implemented. These include frequency selective fading, Doppler shift, propagation delay and the FM noise caused by effectively random phase changes at the receiver. However this was considered to be acceptable as long as it is borne in mind when assessing results and if techniques which are unduly affected by any of the above are avoided.

An area on which no information could be found was the effect of two ground level terminals communicating in built up areas. In a distributed network there are no high up base stations to relay packets, however the signal level between two mobile terminals has not been extensively analysed. Since a Rayleigh distribution is the product of randomly scattered waves then the fading should still exhibit a Rayleigh pattern. The mean received signal levels will be far lower though and the signal should conform to a Rayleigh distribution more closely especially in less urban areas.

The channel simulation includes the inherent assumption that there is a perfect error detection scheme in use in the transmitted packet. This can not only detect all errors but can accurately locate the position of every error without fail. In a practical system this is an impossibility. If a non-perfect code is required a probability of detection can be introduced. In all tests conducted so far the only fact used is whether there are any errors present or not. If coding schemes are to be assessed the channel simulation will be useful in producing realistic error patterns for the type of channel desired.

3. The Network Model

The model for the network is a terminal based simulation. The software simulates the functions and actions of each terminal as opposed to modelling the network traffic. This makes the model ideal for the simulation of a distributed network where all terminals are identical. A control protocol for a mobile radio network should, if efficiently planned, be different from any other network protocol only in the lower levels. In relation to the ISO Open Systems Interconnection reference model [14,15,16] only the network and data link layers would "know" that the channel was not fixed and need to be altered accordingly. Hence the functions performed by these layers are all that it is required to implement for each terminal in the network.

The channel simulation is obviously an implementation of the physical layer. However, since it also detects errors and is responsible for the selection and control of quality of service parameters it does in addition perform many of the functions of the data link layer This base is then used to support the rest of the network, which is basically an implementation of the network layer in that it controls routing, delivery confirmation, congestion control, error notification and decisions on quality of service parameters.

The ISO reference model is not implemented directly since it was considered that this would cause a large increase in the simulation complexity for little or no performance gain. In addition large parts of the model would be redundant when only the functions of the network layer are required. However ideas of the ISO model are used, such as the modular structure gained by using "service primitives" to perform each operation. This makes the network more flexible since it is simple to change and does not require large amounts of programming to alter small protocol details. When firm protocol recommendations are available such a structure will make it easy to translate them from the network simulation into a real mobile network which is based on a standard reference model.

At present the channel and network simulation is capable of modelling access schemes in which the channel is single user only. Channel access schemes such as code division multiple access (CDMA) using spread spectrum techniques and frequency division multiple access (FDMA), in which the channel is accessed simultaneously by two or more users, cannot be implemented. However time division multiple access (TDMA), carrier sense schemes, random access methods such as Aloha or any hybrid mix of these are all possible. This means that there is a great deal of scope for testing existing and experimental control protocols.

The network model has been designed to operate with a relatively small number of terminals. Any number up to about ten could be included, but beyond this the model would become extremely complex and simulation speeds would be severely impaired. Five terminals is about the optimum number but in

specific tests only two or three terminals are required to be implemented. For most purposes this is sufficient. Even when testing protocol details which require a complex network configuration, such as routing algorithms, ten terminals would be enough.

Network configuration has been assumed to relate to the signal to noise ratios and overall channel conditions between individual terminals. These can be calculated by giving each terminal a location, either fixed or mobile, defined by coordinates and using them to measure distances between them and hence signal to noise ratios. However these signal to noise ratios are arbitrary and the distances are used only to get the channel conditions to vary relative to one another in the correct manner between required limits. Equations are available which relate distance to signal to noise ratio in urban areas, but these require detailed information about building statistics in the vicinity. During most trials it is required that only the network channel conditions vary in a prescribed manner not necessarily the distances between terminals. Hence for individual trials almost any network configuration can be implemented in which the channel conditions vary in the required manner.

4. Testing the Network - Optimum Packet Length

The network model could be used to examine many techniques. However there are two main areas in which the model may be used to obtain results. When conducting tests on low level techniques such as synchronisation, coding or modulation the results should include data on link performance for individual channel conditions as well as total network throughput. Higher level tests on network control strategies will require detailed information on packet loss rates, individual terminal states and again on total network throughput. The model should be structured in such a way that all the above data is easily available. The best way to determine whether this is the case, and to assess the effectiveness of the model, is to perform some trials.

Therefore a study of optimum, variable length packets was conducted. This is a simple way of maintaining network throughput as channel conditions degrade. Apart from being useful in its own right the work should highlight any problems with the network model. In addition to requiring low level information on link performance a specific protocol design is necessary so that the optimum lengths can be tested under realistic network conditions and their performance compared to a range of fixed length packets.

4.1 Test Network Models

Before any work can be started a definition of 'optimum packet length' is required. For the purposes of this investigation it will mean the packet length which maximises individual channel throughput. Therefore the restrictions on throughput in terms of packet length need to be clearly identified. The first and most obvious factor is that packet loss rate will increase as packet length is increased due to an higher number of errors within packets. To offset this there is the fact that as packets get shorter more time is wasted in the gaps between packets, ie the overheads increase. Thus the optimum length is a balance between long packets failing more often and the increased number of inter-packet gaps as lengths decline.

Therefore the factors which affect throughput in addition to packet length are the channel conditions - signal to noise ratio and fade rate, and the inter-packet gap. The inter-packet gap length will be assumed to be constant for all packet lengths, which is more or less true in most systems. A definition of the gap is also required. For the purposes of this work it will be assumed to include the actual period of time between transmissions plus the preamble, in fact all times when data is not being transmitted. Hence simulations are required in which the signal to noise ratio and fade rate vary in a controllable manner. This needs to be repeatable so that both the optimum and various fixed length packets can be tested at different inter-packet gap lengths.

Two different test simulations were developed. The first and simplest was to use just the basic channel simulation with variable mean signal to noise ratio and fade rate. The fade rate was varied by changing the mobile terminal speed. In order to assess packet performance for each individual channel condition both signal to noise ratio and fade rate were varied separately in discrete steps. Each channel state was held for a length of simulated time and separate results compiled for the performance of the current packet length and inter-packet gap length in each of the different channel conditions. Totals were calculated for each packet and inter-packet gap length to show the overall performance as the channel varies. This configuration is analogous to a simplex link between two mobile terminals using no handshaking or control protocols. The simulation was used for two purposes. Initially optimum lengths had to be found hence different packet lengths were tried with different inter-packet gap lengths for each channel condition. When the optimum lengths were being tested it was used to analyse the performance in each channel condition.

A more complex network configuration was created. This model simulated a more realistic protocol. To include the effect of packets being repeated by other terminals the network consisted of two fully functioning mobile terminals with a dumb repeater terminal somewhere between them. The repeater terminal simply rebroadcasts all packets which it receives without errors. It cannot change the packet length to keep it optimum. This configuration represents a small part of a network in which two users have a session connection and the broadcast packets are routed by an intermediate terminal. A positive acknowledgement system was used to control errors. The two terminals transfer data between each other and when a packet is successful, ie. it has no errors, an acknowledgement is sent which can be 'piggybacked' on the next data packet. If an acknowledgement is not received for a given packet the source terminal will retransmit that packet.

The channel access method used is a persistent carrier sense (CSMA) scheme. As soon as the channel is quiet a terminal will transmit. This allows variable length packets to be used. To measure the maximum throughput both terminals always have packets to go. Since for a random CSMA scheme such heavy loading would cause the network to collapse each terminal transmits in turn in a token passing manner. To ensure that the repeater does not consistently favour one terminal it is given the opportunity to transmit after both terminals, if it has a packet to repeat.

For a more realistic simulation the channel conditions were varied continuously. It is required to vary the fade rate and signal to noise ratio between maximum and minimum limits which cover the range over which the throughput degrades. The fade rate received at each of the terminals was varied by changing the velocity of the mobile terminals independently in a sinusoidal manner. It is important that the repeater terminal be between the other two. Therefore the positions of the terminals were varied, making sure that the repeater was between the others. Then the distances between the three points were calculated and the signal to noise ratios varied in proportion to these distances. The mobiles were moved sinusoidally at different frequencies between limits which gave suitable maximum and minimum signal to noise ratios. Due to the wide standard deviation of Rayleigh fading the minimum limit is about 10dB at which very little can get through the channel. The maximum limit is about 30dB where the errors have ceased to occur in bursts and approach a random Gaussian distribution.

4.2 Algorithms for Optimum Packet Length

Before an optimum algorithm was developed for variable signal to noise ratio and fade rate the two parameters were examined independently. Two channel simulations were started. One held the fade rate constant at a value corresponding to a mobile speed of 30mph and the signal to noise ratio was varied in steps from 9dB to 30dB. The other held signal to noise ratio constant at 20dB and varied the mobile speed between 10mph and 80mph. These produced figures giving the probability that a packet of a given length will succeed, ie. have no errors, under each particular channel condition.

The probabilities of packet success can be converted to a theoretical maximum channel throughput for different inter-packet gap lengths using equation 6.

$$\text{Throughput} = \frac{D.L.t.p}{L + IPG} \qquad (6)$$

where D = data rate (bit/s)
L = packet length (bits)
t = time (seconds)
p = probability that packet length L will be successful
IPG = inter-packet gap length (measured in bits)

Since the data rate and time are constant and arbitrary for all packet

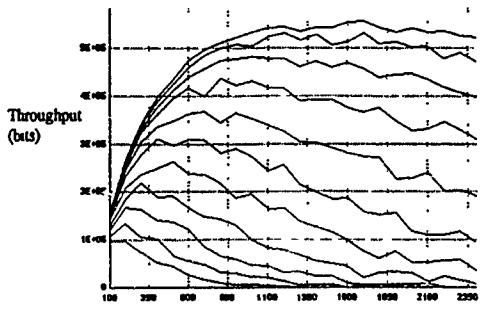


Figure 4.1 - Channel throughput in 100 seconds against packet length for SNRs 12dB to 22dB at an IPG length of 500 bits

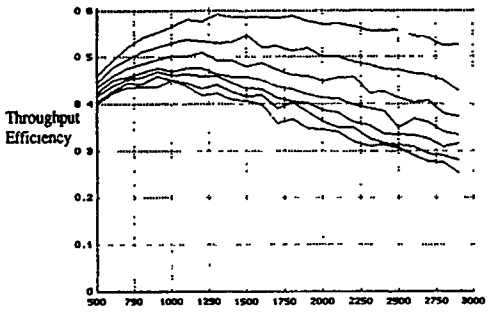


Figure 4.2 - Throughput efficiency of various packet lengths at speeds from 10mph to 70mph and at 20dB SNR at an IPG length of 500bits

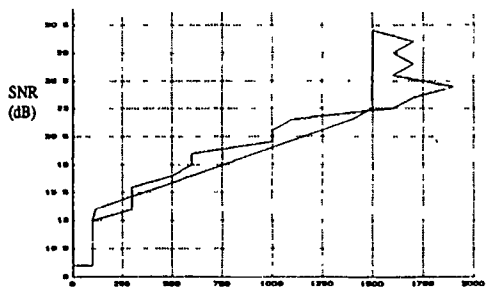


Figure 4.3 - Optimum packet length at an IPG length of 500 bits, algorithm prediction superimposed

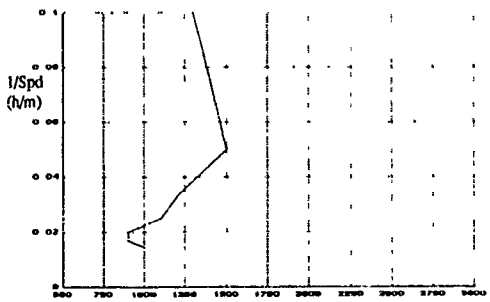


Figure 4.4 - Optimum packet lengths plotted against reciprocal of speed at IPG length of 500 bits

lengths they can be dropped to give a throughput efficiency, equation 7.

$$\text{Throughput Efficiency} = \frac{L_p}{L + IPG} \qquad (7)$$

Figures 4.1 and 4.2 show throughput at an inter-packet gap length of 500 bits. The optimum lengths are the maximum points of the curves in figures 4.1 and 4.2. These were therefore identified and plotted against their respective channel conditions, as in figures 4.3 and 4.4.

An optimum length algorithm should approximate the curves in figures 4.3 and 4.4. A minimum packet length of 100 bits has been applied and it can be seen that the curves for signal to noise ratio approach a maximum asymptote. Hence suitable maximum and minimum limits can be identified. Between these the curves are almost straight hence a linear approximation can be used. This was done for each inter-packet gap length and then the parameters of these lines were themselves approximated linearly. The resulting algorithms are given in equations 8 to 15.

$$\text{Fade Rate Opt Length} = \frac{A}{\text{Speed}} + B \text{ bits} \qquad (8)$$

$$\text{SNR Opt Length} = C.SNR - D \text{ bits} \qquad (9)$$

$$\text{where } A = 5.64*IPG + 2864 \qquad (10)$$

$$B = 0.47*IPG + 558 \qquad (11)$$

$$C = 0.205*IPG + 58.5 \qquad (12)$$

$$D = 2.71*IPG + 783 \qquad (13)$$

$$\text{with Minimum Length} = 100 \text{ bits} \qquad (14)$$

$$\text{Max for SNR opt length} = 1.5*IPG + 750 \text{ bits} \qquad (15)$$

These may appear to be very rough approximations however they

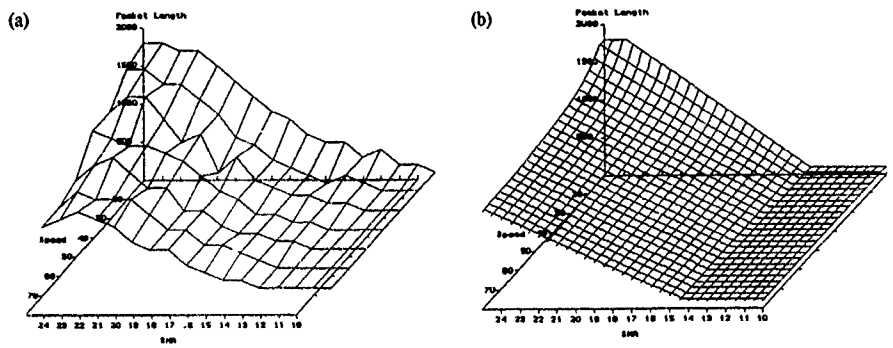


Figure 4.5 - Comparison of simulated and algorithm optimum packet lengths for varying SNR and fade rate at IPG of 400 bits.(a) Simulated optimum lengths and (b) ratio algorithm of equation 17

are surprisingly close to the simulated curves at all inter-packet gap lengths. If they perform well when tested it will prove that optimum length packets are beneficial even if they are rough approximations to the ideal optimum.

To derive an optimum packet length for use when both signal to noise ratio and fade rate change the two algorithms found were combined. They intercept at 20dB and 30mph. The optimum packet length at this point is coincidentaly 1000 bits. Therefore there are two options available. The algorithms can be added and 1000 subtracted.

$$Sum\_Opt = SNR\_Opt + Fade\_Rate\_Opt - 1000 \text{ bits} \quad (16)$$

Alternatively the two can be multiplied and then divided by 1000.

$$Ratio\_Opt = \frac{SNR\ Opt \cdot Fade\ Rate\ Opt}{1000} \text{ bits} \quad (17)$$

A channel simulation was started which varied both signal to noise ratio and fade rate in discrete steps between 9dB and 30dB and 10mph and 80mph respectively. The resulting probabilities were converted to throughput and the optimum packet lengths found. Figure 4 5 gives a comparison of these results and equation 17. It would be possible to use the simulation results to derive an optimum length algorithm which was a better fit. However this would be a long and tedious process so both equations 16 and 17 will be tested.

4.3 Testing the Optimum Length Algorithms

The four algorithms found were assessed and their performance compared to a range of fixed length packets at different inter-packet gap lengths. The channel simulation allowed comparisons at particular channel conditions. When testing the optimum algorithm for varying signal to noise ratio only the signal to noise ratio was changed and the fade rate held constant. Similarly for the fade rate algorithm only the fade rate was varied and the signal to noise ratio held constant. Typical results are shown in figure 4.6. The algorithms for both fade rate and signal to noise ratio were tested varying both channel parameters. Figure 4.7 gives some sample results.

All four algorithms were tested on the network model with a repeater terminal and the results compared to various fixed length packets. Again only the relevant parameters were varied, but for this simulation they changed continuously as described in section 4.1. This means that no results are available for each individual channel state, only over the whole path. Typical results are shown in figure 4.8. It should be noted that the results are specific to the particular path simulated and the way that the channel parameters were varied on the path.

4.4 Analysis of Results

In terms of the optimum packet length investigation the main result, though predictable, is that optimum packet lengths are always as good as or better than any single fixed length at maximising throughput. However, more importantly, the results show that the channel simulation and network model are effective. The particular protocol implemented worked well and all the information required to analyse performance was available. There were several specific points to arise during the investigation though, and these will be discussed here.

The performance of the optimum packet length in the network simulation was not as good as it was in the simple channel simulation relative to the fixed length packets. This is because the repeater terminal changes the effective probability that a packet will succeed so that an optimum length based on channel conditions is no longer optimum. However since all the links in a network will have different routes and local configurations this cannot be accounted for in the optimum length algorithm. In addition sending acknowledgements increases the redundancy of the system. If a packet is successfully received but the acknowledgement is not then the packet is needlessly repeated.

The fade rate algorithm is not as effective as the signal to noise ratio algorithm in increasing the throughput. The explanation for this is simply that a varying fade rate does not have such a severe effect on the channel as a varying signal to noise ratio. Hence it is more important to keep the packet length optimum

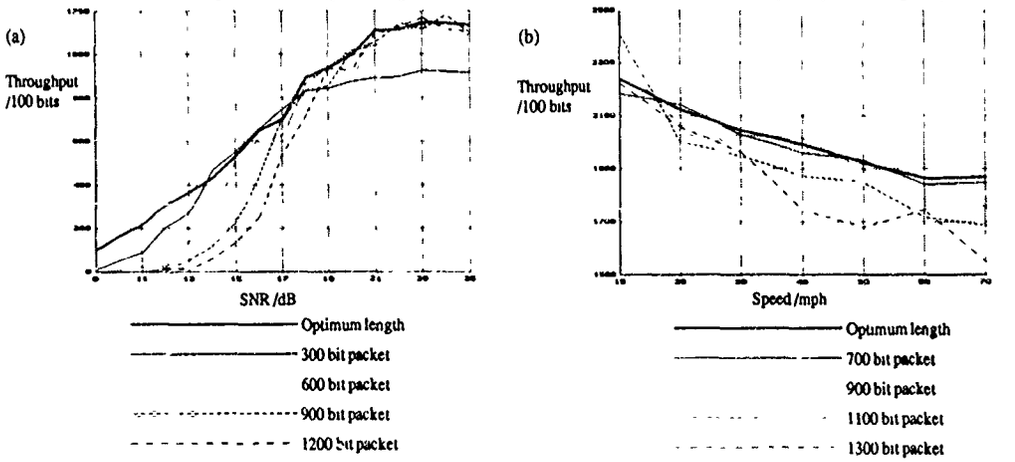


Figure 4.6 - Sample throughputs of optimum and fixed length packets at an IPG length of 200 bits (a) for signal to noise ratio algorithm and (b) for fade rate algorithm

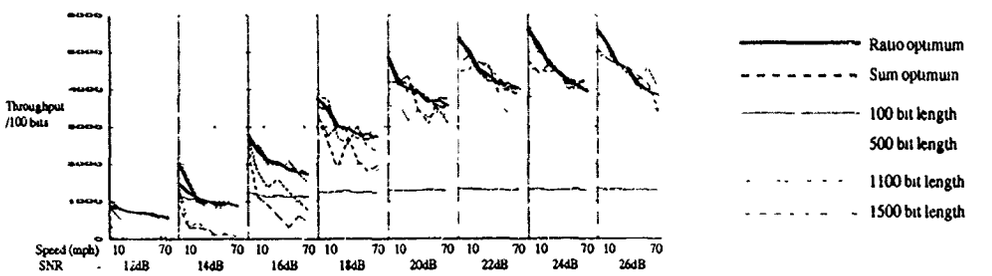


Figure 4.7 - Throughput of optimum and fixed length packets as both SNR and fade rate vary

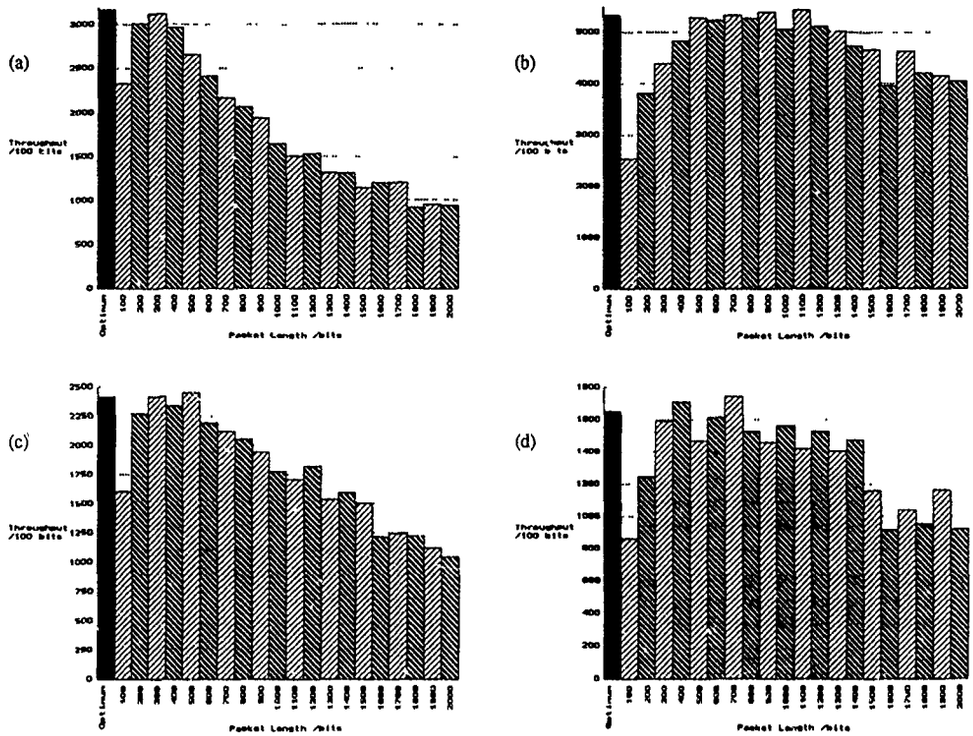


Figure 4.8 - Total throughput of optimum and fixed length packets over repeater network configuration with an IPG length of 200 bits : (a) Signal to noise ratio algorithm, (b) fade rate algorithm, (c) ratio algorithm and (d) ratio algorithm with an IPG length of 500 bits

relative to the signal to noise ratio than the fade rate.

Both combined algorithms for varying fade rate and signal to noise ratio had, except is specific cases, very similar performances. However a close examination of the figures reveals that equation 17, the ratio optimum algorithm, is marginally more effective When it is considered that these expressions are much worse fits to the simulated optimum lengths than either the individual signal to noise ratio or fade rate algorithm then they perform surprisingly well.

A major point which had to be considered was upon which individual link in a network route should the optimum length be based? If the destination is 'aimed' for directly then the advantages of having terminals repeat or route a packet are minimal. Since repeating terminals will not re-optimise a packet to a different length then using the channel conditions of the first link en route is likely to lower throughput. The method used here was to base the optimum length on the channel with the lowest signal to noise ratio likely to be encountered on the packet route. However when the route is longer and not certain this method may not be as effective.

Variable length packets could not be used effectively for all channel access methods. In particular a strict time division multiple access system would find no advantage in optimising packet lengths. To maintain a minimum information throughput lengths could be lowered within each slot but this would increase the redundancy of the channel enormously and could only be an emergency measure. However optimum length packets are ideally suited to all forms of unslotted carrier sense access schemes and it is here that the largest benefits could be found. The use in frequency or code division multiple access schemes may also be feasible

The major problem to be identified in the investigation was the amount of channel information which terminals would require. Each terminal in a distributed network would need to have accurate and up to date information on every usable link in the network. This would result in a large increase in control information flowing round the network. In a distributed network, with no stations to gather and assess such information, this should be avoided.

5. Conclusion

The design of a flexible radio network simulation and some initial trials to assess its effectiveness have been described The simulator and the methods used to implement it have been shown to be successful. A small set of computer simulations now exist which are capable of performing all the functions of a distributed radio network. However the model is not yet finished, as work progresses and trials are conducted the simulation will be updated and improved continuously.

The examination of optimum lengths proved to be useful in itself. The results show that the implementation of such a system is possible and that it can have a beneficial effect on throughput. The simulation also allowed the problems and pitfalls which may be encountered in designing a system to be identified In addition the investigation showed that the methods used for the channel simulation do work and gave the opportunity to iron out all the small bugs in the channel simulation and network model

References

[1] Kahn, R E., Gronemeyer, S.A., Burchfiel, J. and Kunzelman, R.C.: "Advances in Packet Radio Technology", Proc IEEE, Vol 66, 1978, pp. 1468-1497

[2] Reudink, D.O.: "Properties of Mobile Radio Propagation above 400MHz", IEEE Trans. Vehicular Technology, VT-23, 1974, pp 143-159

[3] Davis, B.R and Bogner, R.E.: "Propagtion at 500MHz for mobile radio", IEE Proc F, Vol 132, 1985, pp 307-320

[4] Parsons, J.D. and Ibrahim, M.F.: "Signal Strength Prediction in Built up Areas Part 2 : Signal variability", IEE Proc. F, Vol 130, 1983, pp.385-391

[5] Olivier, P. and Tiffon, J.: "Fast Fading Characterisation in Urban Mobile Propagation at 855MHz in Paris", Electronics Letters, Vol 21, 1985, pp.830-832

[6] Clarke, R.H.: "A Statistical Theory of Mobile Radio Reception", Bell System Technical Journal, Vol 47, 1968, pp.957-1000

[7] Ralphs, J.D. and Sladen, F.M.E.: "An H.F. Channel Simulator Using a New Rayleigh Fading Method", Radio & Electronic Engineer, Vol 46, 1976, pp.579-587

[8] Arredondo, G.A., Chriss, W.H. and Walker, E.H.: "A Multipath Fading Simulator for Mobile Radio", IEEE Trans Communications, COM-21, 1973, pp.1325-1328

[9] Ball, J.R : "A Real-Time Fading Simulator for Mobile Radio", Radio & Electronic Engineer, Vol 52, 1982, pp.475-478

[10] Arnold, H.W. and Bodtmann, W.F.: "A Hybrid Multichannel Hardware Simulator for Frequency Selective Mobile Radio Paths", IEEE Trans Communications, COM-31, 1983, pp.370-377

[11] Bodtmann, W.F. and Arnold, H.W.. "Fade Duration Statistics of a Rayleigh Distributed Wave", IEEE Trans Communications, COM-30, 1982, pp.549-553

[12] Jakes, W.C.: "Microwave Mobile Communications", Wiley-Interscience, 1974

[13] Darnell, M.. "Review of Modulation, Coding & Speech Digitisation Techniques", Agard Report No.744, 1986, pp.2-1 to 2-17

[14] Marsden, B.W.: "Communication Network Protocols", 2nd Edition, Chartwell-Bratt, 1986, Chapters 14 & 15

[15] ISO 7498, "Reference Model for Open Systems Interconnection", 1984

[16] ISO 8648, "Open Systems Interconnection : Internal Organisation of the Network Layer", 1985

DOPPLER-MULTIPATH TOLERANT VOICE COMMUNICATION

by
R. M. Harris
Royal Aerospace Establishment
Farnborough, Hampshire
GU14 6TD, UK

SUMMARY

Line-of-sight communication between high-performance aircraft has been found to be subject to a peculiar form of multipath radio-wave propagation - Doppler multipath. It degrades analogue voice reception on the standard-fit UHF radio, producing low-frequency random noise and warbling.

Various modifications were carried out on the aircraft's communications system but the problem remained; all the evidence points to a natural phenomenon.

The reported observations are corroborated by theoretical studies and laboratory simulations of multipath radio-wave propagation between 2 points moving relative to a diffusely scattering reflector. Theoretical predictions of Rician fading have explained the disruption of speech transmitted using conventional dsb(am) modulation. This also indicated suppressing the carrier as a radical cure.

Double sideband suppressed-carrier radios have been developed for airborne evaluation in comparison with standard dsb(am). The air-to-air flying trials proved the superior performance of the suppressed-carrier system under conditions of Doppler multipath.

1 INTRODUCTION

Tactics in air warfare, as in any other area of military activity, are under continual review, and changes take place over a period of time. Recent times have seen the need for changes in tactics for strike aircraft in response to vast improvements achieved in radar performance. Strike aircraft are now required to approach enemy territory at very low altitude, effectively flying underneath searching radar beams, in order to escape detection for as long as possible.

Low-level formation flying entails a high workload for the pilot, and the last thing that must take his attention is the reliability of the tactical UHF radio communications. But, under certain flight conditions, excessive interference effects began to be encountered on the air-to-air link.

The interference observed was characterised by a sporadic rumbling noise, sometimes accompanied with howls and warbling of the speech; occasionally it was intense enough to obliterate the message. Its causes have remained elusive for some considerable time.

As will be explained, the scientific investigation of the problem was circuitous and spawned a number of interesting developments. The initial thesis that a specific form of multipath radio-wave propagation was responsible led to adoption of the term "multipathing" for the phenomenon. Development of the theory of multipath wave propagation with Doppler shift led to various attempts to model the natural mechanism in the laboratory by subjective comparison of the audible effects of "multipathing". Direct investigation of the radio-frequency (r.f.) phenomenon was impractical.

The search for a more direct means of verifying the so-called Doppler-multipath hypothesis led to a form of carrierless analogue voice modulation whose performance would depend heavily on the nature of the perturbations in the r.f. channel.

2 CONFIRMATION OF "MULTIPATHING"

Suggested candidates were cockpit acoustic noise, aircraft vibration, random motion in turbulent air, and malfunctioning of the communications electronics. Minor modifications were made to the Communications Control System (CCS), the Voice Operated Gain Adjusting Device (VOGAD), and the automatic gain control (agc) systems were investigated [1]. Whilst some improvements resulted, the essential interference persisted. Mechanisms involving the airframe itself were discounted since "multipathing" was reported on 4 or 5 substantially different types of military aircraft.

3 REPRODUCTION OF "MULTIPATHING"

The interference had been monitored in Service aircraft by equipping aircrew with personal miniature tape recorders. Subjectively similar recordings were obtained from special trials conducted at RAE [2] in transport aircraft. Three quite disparate radio receivers (a valve and a transistorised military transceiver, and a measurement receiver) were connected in parallel, and the audio outputs were recorded during air-to-air transmissions of conventional dsb(am) speech over a uhf radio link. When "multipathing" occurred the audio outputs were indistinguishable to the human ear. Third-octave spectra also confirmed that the respective "multipathing" signals were identical. Attempts to separate the speech from the noise by high-pass filtering failed to improve the quality or intelligibility; there appeared to be too much spectral overlap [3].

4 MULTIPATH PROPAGATION

Evidently it was unnecessary to invoke any sort of receiver malfunction to account for the observed effects. The cause seemed to be a multipath electromagnetic wave-propagation mechanism not previously understood. Indeed, the average severity of the "multipathing" was found to depend on the spatial configuration of the 2 aircraft and was inversely proportional to their

height above average rolling countryside (see Table 1).

Line-of-sight radio communications between aircraft are normally subject to the effects of multipath propagation. For many flying configurations the main effect is a relatively slow rate of signal fluctuation, which is usually disguised by the compensating action of the receiver's automatic gain control (agc).

This "multipathing" was different; it was virtually independent of agc action and it increased in proportion to both aircraft speed and carrier frequency.

5 THEORETICAL TREATMENT OF "MULTIPATHING"

The suggested link with multipath radio-wave propagation was studied and a theoretical basis for the observed phenomenon developed, namely, 2 aircraft moving together over an uneven reflecting surface (Fig 1a). The problem may be analysed either kinetically (by invoking the concept of multiple Doppler shifts arising from oblique scattering from the ground) or quasi-statically in terms of multipath interference [4].

A time-dependent spatial electromagnetic fringe pattern is generated by interference between the direct and ground-reflected waves from the source. The receiving aircraft then moves partly with and partly through the fringe pattern, experiencing relatively slow fluctuations in the strength of the coherent wave. Super-imposed is a structureless component resulting from diffuse scattering from the uneven terrain.

This Doppler-multipath model predicts an effect which approximates a Rician fading channel. Two parameters need to be specified:

i the ratio of the totally random (Rayleigh) field component to the coherent wave, and

ii the bandwidth of the Rayleigh component.

The first parameter is a measure at any given moment of the strength of terrain-scattered energy, and also the degree of destructive interference between the direct and single (or multiple) waves specularly reflected from the terrain. This parameter fluctuates as the receiving antenna moves in and out of interference minima - hence the sporadic nature of the effect.

The second parameter is related to a terrain roughness index (see Appendix A) and is enhanced by the grazing angle of the reflected waves, the aircraft speed, and the carrier frequency.

6 MICROWAVE MODELLING OF DOPPLER MULTIPATH

The multipath simulator took the form of a 1/30th scaled model of aircraft flying over rough ground, but the aircraft remained stationary and the landscape moved beneath them. Rough ground was represented by crinkled aluminium foil attached to the periphery of a 6 ft diameter wooden drum, and this was rotated at speeds equivalent to ground speeds of between 400 and 800 knots. The military UHF Band transmission frequencies were represented by frequencies 30 times higher in the I/J band. Aircraft antennas were represented by 2 horn antennas mounted above the rotating aluminium surface; one was used as the emitter and the other was used to collect the scattered radiation. The direct wave between emitter and receiver was achieved by inserting a waveguide coupler and attenuator between the horn waveguides. The complete system is depicted by Figure 2.

A high-quality recording of speech was used as source material to modulate the transmitter. The output from the receiver passed through an audio amplifier to the loudspeaker. Optimal (amplitude) modulation level and amplifier gain were established with the drum stationary. The drum's rotation rate was adjusted for an equivalent ground speed of 500 knots, the horn antennas were positioned, and the strength of the r.f. signal through the waveguide coupler (representing the direct wave) was adjusted for maximum level of interference on speech from the loudspeaker. Low-frequency audio rumbling predominated, but the high-frequency components encroached well into the speech band. As the drum's rotation rate (ground speed) increased so did this encroachment.

The resulting audio output consisted of the original voice signal plus a continuous background noise occupying sub-audio frequencies and the low audio band. As the direct wave was adjusted, so as to add in antiphase to the reflected wave, the subjective equivalent of "multipathing" was heard. Spectral measurements (see Fig 3) of the low-frequency noise also corroborated the Doppler-multipath hypothesis.

7 REVIEW OF STRATEGY

There is good correspondence between the observations of "multipathing", the microwave modelling, and the rudimentary theory of Doppler multipath. The observed effects were qualitatively and quantitatively within the limits set by the theory. Nevertheless, a more direct investigation of the r.f. phenomenon was felt necessary before a proper solution could be determined. Unfortunately the required flying laboratory was impractical.

Two candidate remedies to the "multipathing" problem had been proposed. The first simply accepted the interference and shifted the voice spectrum away from the band of noise, restoring it after high-pass filtering at the receiver. That method did not depend on an understanding of the interference mechanism. The second involved a change of r.f. modulation and depended for success on the correct understanding of "multipathing".

8 INTERACTION OF DOPPLER MULTIPATH WITH AMPLITUDE MODULATION (AM)

The Rayleigh component of the Rician fading predicted by the Doppler-multipath hypothesis introduces random phase and envelope modulations on top of the wanted signal modulation. They have the effect of

multiplying the wanted double sideband (dsb) spectrum with a noise-like spectrum (see Appendix B) whose bandwidth is determined by the model's input data (speed, carrier frequency, terrain-roughness index, etc).

The r.f. spectrum of dsb(am) is illustrated in Figure 4 (inset). It consists of a fixed carrier and 2 symmetrical sidebands containing the wanted modulation frequencies. The effect of multiplying the sidebands and carrier with a noise-like signal is now considered.

The frequency-domain structure of speech largely comprises 4 "formants", characteristics of the vocal tract [5]. They manifest themselves as broad peaks in the harmonic spectrum of spoken phonemes. With a natural bandwidth of 200-300 Hz they can easily survive the little extra broadening described by the Doppler-multipath process. The subjective effect of so broadening the modulation sidebands is not deleterious, adding a tolerable hoarse or gravelly quality to the speech.

The carrier, on the other hand, is endued with new noise-like sidebands corresponding to maximum (100%) modulation index. Figure 4 illustrates the phenomenon. The noise-like sidebands are superimposed on the (speech) modulation sidebands with the result that noise is added to the audio output of the receiver.

Whilst not an essential element of the dsb(am) modulation standard, in practice am receivers use nothing more sophisticated than non-coherent envelope detectors. It can be shown that such a non-linear detector produces severe harmonic distortion of unconstrained signals such as carriers that have undergone multipath or Rayleigh fading. Such frequency multiplication extends the noise bandwidth after demodulation well beyond the bandwidth of the perturbations of the radio channel, and accounts for further encroachment into the audio pass band.

In conventional "amplitude modulation" - dsb(am) - neither sideband may be more than 6 dB below the carrier at 100% modulation index. Because of the danger of over-modulating speech signals the average modulation index is set much lower, so further disadvantaging the wanted signal relative to this carrier-borne interference.

The corollary is that simply suppressing the carrier prior to transmission will eliminate the additive noise, which has the most harmful effect. Two forms of carrierless modulation are available: single sideband (ssb) and double sideband suppressed carrier (dsb.sc). The effect of channel fading on dsb.sc is illustrated in Figure 5.

9 PILOT-SCALE DSB.SC COMMUNICATION SYSTEM

Implementing a carrierless transmission system is dependent on reconstituting the carrier, in phase, at the receiver. For arbitrary modulating signal the only key to locating the carrier wave is the symmetry of the r.f. spectrum; clearly ssb per se is ruled out. The implicit symmetry of dsb.sc is preserved in transmission because the coherence bandwidth (qv) of the multipath is of the order of 10 times the occupied bandwidth.

A pilot-scale double-sideband suppressed-carrier communication system was developed [6] using a low carrier frequency of 80 kHz to ease electronic circuit design. Most of the functions of a full-scale uhf radio communication system were represented including: linear modulation, automatic gain control (agc), carrier reconstitution, and coherent demodulation. Out of several different carrier-recovery methods [7,8,9, 10] the one selected at this stage was the so-called 2f method. The whole r.f. signal is first squared to generate a phase-coherent carrier at twice the frequency f of the original. A fast-acting phase-locked loop then cleans up the 2f carrier before it is divided by 2 and used to reference the coherent demodulator. An auxiliary audio tone was added to the speech signal to keep the receiver on track during pauses in speech. Placing this tone above the conventional audio pass band at 5 kHz made it easy to filter it out at the receiver.

10 RICIAN FADING TRIALS OF DSB SUPPRESSED CARRIER

A Rician fading simulator was interposed between the sender and receiver of the pilot-scale system (see Figure 6). The bandwidth and relative intensity of the Rayleigh component could be varied at will. Thus the fading channel could be varied continuously from nearly pure Rayleigh, through all intermediate combinations, to steady coherent transmission.

The dsb.sc system allowed the carrier to be diminished progressively from 0 dB to -50 dB relative to the audio sideband. The rms and peak noise produced at the audio output was measured using a Bruel & Kjaer Type 2425 meter and was found to vary in proportion to the residual carrier (see Figure 7). The resulting signal-to-noise ratio steadily increased as the carrier was progressively diminished down to 25 dB below the audio sideband; full suppression was unnecessary. Even with full Rayleigh fading the spoken sounds, whilst sounding hoarse or gravelly, were yet intelligible and the interstices free of noise. The acquisition time of the carrier reconstitutor was practically instantaneous.

For comparison a conventional dsb(am) system, using an envelope detector, was substituted for the suppressed-carrier system. The additive noise reappeared and its intensity was proportional to the strength of the Rayleigh component of the fading. The noise filled in the gaps between phonemes and, at its most intense, obliterated the speech. High-pass filtering of the audio signal failed to restore intelligibility because the noise spectrum seemed to match that of the speech. That was significant since the bandwidth of the noise-like excitation of the Rician fading generator had been set at 170 Hz, the value for a typical low-level flying sortie.

11 REQUIREMENTS FOR AIRBORNE TRIALS OF DSB.SC

In order to compare different communications systems the exposure to "multipathing" must be consistent. The sporadic nature of the observed effect was entirely consistent with the Doppler-multipath hypothesis. Magnification of the Rayleigh component is contingent upon the receiving aircraft placing its aerial in a minimum in the spatial interference pattern of the direct and indirect waves. For exact repeatability the 2 aircraft would have to navigate to an accuracy of 0.1 m, each holding precisely the same heading and bank angle; in practical terms the interference is unrepeatable. That consideration ruled out system comparison on a flight-by-flight basis.

Parallel testing would entail 2 separate UHF channels (for the 2 modes are different from sender to receiver). They would have to be closer in frequency than the coherence bandwidth of the interference pattern, about 500 kHz, and such allocations were not available. Rapid serial comparison of dsb(am) and dsb.sc modes seemed to be the only recourse. Fortunately under experimental conditions the "multipathing" tended to remain consistent over periods lasting for several seconds. That would permit fair comparison of short transmissions in alternate modes provided the switch over could be done quickly.

12 DEVELOPMENT OF UHF DSB.SC AIRBORNE RADIOS

For economy, a one-way radio communication system was specified comprising a sender and receiver - each of which could be made to select suppressed-carrier mode or conventional dsb(am) virtually instantaneously. In view of the proximity of the aircraft, a low output power of 1 watt was considered adequate.

The suppressed-carrier facility was implemented by modifying a pair of ex-Service UHF transceivers. Whilst, at this stage, a Costas Loop might have given superior performance it would have entailed a re-design of the existing i.f. strip, and so an elaboration of the "2f" method [11] was employed for carrier reconstitution. As in the pilot-scale version, the acquisition of the dsb.sc signal was practically instantaneous. In all, 3 printed circuit boards and an FET ring modulator were added to the basic radios.

Prior to the flight trials the sender and receiver were fully characterised, paying particular attention to the carrier-tracking range as a function of r.f. input power. The complete link was then tested on the ground in both modes with steady and rapidly fading r.f. channels.

13 AIRBORNE MULTIPATHING TRIALS OF SCB.SC

Survivability in the air, and general EMC, was tested by flying air-to-ground and then ground-to-air radio links. The experimental radio link allowed communication in only one direction. Two-way communications were furnished by a separate radio link.

The first air-to-air flight trials established where and under what flying configurations reliable "multipathing" could be produced. The 2 aircraft began by flying in line astern formation at 250 knots. The favoured configuration for inducing severe "multipathing" was to have the second aircraft pull forward from trailing position at an altitude of 500 to 1000 ft below the lead aircraft until it was beneath it. Very low flying was ruled out on grounds of safety, but some sorties were flown below 2500 ft.

The final proof trials took place over southern England and the Midlands in early 1990, using 2 BAC 1-11 aircraft fitted out as flying laboratories. The main piece of instrumentation was a TEAC professional 7-track tape recorder to record the received speech, and monitor the agc line and the transmitted envelope. Standard test sentences were read and transmitted in alternately conventional dsb(am) and suppressed-carrier modes. Several hours of recordings were analysed subjectively, resulting in the production of exemplary episodes of severe "multipathing" in which the presence of the interference came and went as the transmission mode was switched between dsb(am) and suppressed carrier (see Figure 8). Fast Fourier Transform (FFT) spectra were derived from portions of "multipathing" that occurred when there was no voice modulation (see Figures 9,10); the noise reduction is virtually total. A statistical analysis of the trials tapes was made in which all episodes of multipath interference were recorded and the quality of the following dsb.sc transmission noted. The results are shown in Figure 12 for 2 levels of interference. Severe multipath interference made normal am unintelligible.

14 CONCLUSION

Frequency modulation has been identified as an alternative to dsb.sc. The Doppler-multipath hypothesis predicts phase fluctuations as well as envelope fluctuations such that, for comparable channel bandwidths, (narrowband) fm will fare no better than dsb(am). The prospect of overcoming the interference by using wideband fm carries the penalty of substantially increasing the occupied bandwidth in an already overcrowded radio band. Dsb.sc, on the other hand, is an inherently narrow band and linear modulation standard which, in addition, is compatible with conventional dsb(am) broadcasts.

The employment of double-sideband suppressed-carrier in the presence of "multipathing" type inteference has successfully demonstrated the application of the Doppler-multipath hypothesis.

A fully engineered dsb.sc system will outperform conventional UHF dsb(am) short-range air-to-air radio communications. A 10-fold reduction in message failure rate due to "multipathing" is predicted for dsb.sc.

It is recommended that a fully engineered 10 W version of the UHF dsb.sc radio be trialled in Service high-performance aircraft.

REFERENCES

1 R. M. Harris; Characterisation of the dynamical response of receivers to fading. Clerk Maxwell Commem. Conf. on Radio Receivers and Associated Systems, Conf. Proc. No. 50, Leeds, 79, p235, 7-9 July 1981.

2 J. S. Webb; Multipath propagation interference experiments and comparative radio performance tests. RAE Tech. Rep. 82079 (1982).

3 J. S. Webb; Multipath propagation interference on uhf communications between aircraft and a method for its suppression. RAE Tech. Rep. 83054 (1983).

4 R. M. Harris, N. A. D. Pavey; Air-to-air UHF radio communications under severe multipath propagation conditions. RAE Tech. Rep. 84078 (1984).

5 G. A. Kopp, H. C. Greene; Basic Phonetic Principles of visible speech. J. of the Acous. Soc. of America, 18 (1946).

6 R. M. Harris; A multipath-tolerant aeronautical communication system, Part II. RAE Tech. Rep. 82061 (1982).

7 J. P. Costas; Synchronous communications. Proc. I.R.E., 44, p1713, Dec 1965.

8 R. S. Badessa; A communication detector with a signal synthesized reference. IEEE Trans. Comm. Tech., Com-19, No. 5, Oct 1971.

9 D. A. Tong; Phase-locked detector for double-sideband diminished-carrier reception. Wireless World, 79, Sept 1981.

10 F. G. Apthorpe; Homodyne reception. Electronic Engineering, 19, p238, July 1947.

11 R. C. V. Macario; A vhf surveillance receiver adapted for the reception of suppressed-carrier double-sideband transmissions. The Radio & Electronic Engineer, 43, No. 7, p407, July 1973.

12 H. R. Reed, C. M. Russell; Ultra-high frequency propagation. Chap. 5, Chapman and Hall (1966).

13 P. Beckmann, A. Spizzichino; The scattering of electromagnetic waves from rough surfaces. Pergamon Press (1963).

14 P. E. Redmill; The rf spectrum of a carrier phase-modulated with noise. RAE Tech. Rep. 71239 (1971).

15 P. A. Bello; Aeronautical channel characterization. IEEE Trans. Com-21, No. 5, May 1973.

GLOSSARY

agc	automatic gain control
am	amplitude modulat(ion/ed)
CCS	Communications Control System
dsb	double-sideband
dsb.sc	double-sideband suppressed-carrier
dB	decibel
EMC	electromagnetic compatibility
FET	field-effect transistor
FFT	fast Fourier transform
f.m.	frequency modulat(ion/ed)
i.f.	intermediate frequency
I band	= 8.0 - 10.0 GHz
J band	= 10.0 - 20.0 GHz
r.f.	radio frequency
rms	root mean square
ssb	single sideband
uhf	ultra high frequency (300-3000 MHz)
UHF	military Ultra High Frequency band (225-400 MHz)
VOGAD	Voice Operated Gain Adjusting Device

APPENDIX A

The subject of multipath due to specular reflections of aircraft-borne radio waves from the ground is well established [12]. In contrast, the diffusely scattered signal is less well understood and is too variable to give rise to any structured interference fringes. It is most usefully treated [13] as a random noise-like variable, and is characterised by its rms power ρ and bandwidth $\pm\Delta f$. The composite reflected wave may be described by the addition of a time-varying vector in phase space to the discrete specular components. If the aircraft-ground-aircraft configuration is such that the principle indirect wave propagation path makes an angle of Ψ to the horizontal, then a criterion for "rough terrain" is that $h/\lambda \sin \Psi >> 1$ where λ is the wavelength and h is the rms surface height above the mean tangent plane (horizontal).

The crucial factor in the analysis is the ratio R of the dimensions of the first Fresnel zone to the auto-correlation distance D of the surface features. If $R>>1$ an effect more like phase modulation of the indirect wave occurs whereas if $R<<1$ a truly diffuse scattering process takes place, corresponding to classical Rayleigh fading. It turns out that in both cases the bandwidth of the perturbations, whether phase or phase and amplitude, is a more or less constant function of the aircraft ground speed, v , the grazing angle, Ψ , and the rms slope, α , of the surface. Terrain topology is described statistically by a Gaussian distribution with standard deviation h of the surface height above its mean, and autocorrelation distance D . Bellow [15] defines the rms slope α thus:

$$\tan \alpha = \sqrt{2}\, h/D \; . \qquad \ldots\ldots (1)$$

$$\Delta f = \beta\sqrt{2}\alpha V \sin \Psi/\lambda \qquad \ldots\ldots (2)$$

where $1<\beta<\sqrt{2}$.

Δf is defined at the $\exp(-\frac{1}{2})$ points of the power spectrum.

Typical data are:

$\alpha = 0.39$ (45 deg)
$V = 310$ m/s (600 knots)
$\Psi = 45$ deg
$\lambda = 1$ m

ergo $121.0 < \Delta f < 171.6$ Hz.

The various results of the interference between the 2 waves may be explained by reference to the phase-space diagrams at Figure 11. When it happens that the receiving aircraft is positioned near a minimum in the spatial interference pattern the vector representation of Figure 11a obtains. The specular component of the reflected (indirect) wave (B) is comparable in magnitude but opposite in polarity to the direct wave (A). The instantaneous reflected wave vector, R(t), is the sum of B and the noise-like scattered component, represented by the time-varying vector D(t). The circle round the tip of B represents the locus of a vector having a random phase, and rms amplitude ρ - the statistical representation of D(t). Two-thirds of the energy in the diffusely scattered wave is contained within the ρ circle.

The electric field experienced by the receiving antenna is the sum of the instantaneous wave vectors - considered as vectors in respect of both the geometrical space, in which the waves propagate, and in phase space. It is assumed that the electric-fields are sufficiently close to the vertical to make the phase-space vectors the determining factor. The resultant is given by the vector sum illustrated in Figure 11b:

$$V(t) = A + B + D(t) \qquad \ldots\ldots (3)$$

V(t) can be described as the sum of a phase-coherent component, C, and the noise-like component, D(t), with rms amplitude ρ

where $C = A + B$. $\qquad \ldots\ldots (4)$

Eq.(4) is an expression of the specular interference pattern. A fuller treatment of the wave polarisation cannot be undertaken here but it is sufficient to state that, in general, the coherent component, C , will be weakened relative to the noise-like component.

The phase modulation which D(t) imparts to both the indirect wave alone ϕ_b and to the resultant V(t) ϕ_V is given by the following approximations:

$$\phi_b = 2/\pi \;\; \rho/B \qquad \ldots\ldots (5)$$

$$\phi_V = 2/\pi \;\; \rho/C \qquad \ldots\ldots (6)$$

where $\rho<<B$ $\rho<<C$.

In the vicinity of minima in the interference pattern, $C<<B$ and therefore the resultant phase modulation ϕ_V may exceed the phase modulation ϕ_b of the indirect wave by the factor B/C. If either ρ increases or C decreases such that $\rho<<C$ no longer holds, then Eq.(6) becomes invalid and the vector representation of Figure 11c becomes appropriate. Here C is swamped by D(t) and the locus of the rms diffuse component encompasses the origin of C as shown by the circle, radius ρ . The phase of the resultant is almost as random as D(t) itself and fills the range 0 to 2π .

The amplitude distribution is also close to Rayleigh.

Figures 11a,b illustrate the way in which random amplitude modulation is produced in addition to phase modulation. For $B>>\rho$ the intrinsic amplitude modulation of the indirect wave is given by

$$m_b = \rho/B \qquad \ldots\ldots (7)$$

and for the resultant:

$$m_b = \rho/C \qquad \ldots\ldots (8)$$

$C>>\rho$

Again, the amplitude-modulation index is increased by the factor B/C.

APPENDIX B

ENVELOPE DEMODULATION OF MULTIPATH-FADING SIGNALS

The important effect of fast fading (multipath or other forms) on full-carrier am reception is the random, noise-like envelope modulation of the transmitted signal, whether already modulated or not. Ignoring the largely irrelevant phase modulation, the equivalent am fading channel essentially multiplies (exactly) the signal envelope by the time function, g(t)

where

$$g(t) = 1 + n(t) \qquad |n(t)| < 1 \quad ...(B\text{-}1)$$

and n(t) is a zero-mean noise-like function of bandwidth δf and rms magnitude related to the factor ρ/C (see Section

Consider a carrier $y(t) = a \cos \omega t$ modu'ɔ'ed to a depth m_S by the wanted signal s(t) . The transmitted signal y'(t) is given by

$$y'(t) = a\left[1 + m_S s(t)\right] \cos \omega t \quad ...(B\text{-}2)$$

where $m_S \leqslant 1$, $|s(t)| \leqslant 1$.

After passing through the equivalent am fading channel, the received signal y"(t) is given by

$$y''(t) = g(t)y'(t) \quad ...(B\text{-}3)$$

$$= a\left(1 + n(t)\right)\left(1 + m_S s(t)\right) \cos \omega t \quad ...(B\text{-}4)$$

$$= a\left[1 + n(t)\right] \cos \omega t + a m_S s(t) \cos \omega t + a m_S n(t) s(t) \cos \omega t \quad ...(B\text{-}5)$$

After perfect envelope detection at the receiver (no harmonic distortion or intermodulation products) the baseband signal z(t) may be written in normalised form as

$$z(t) = 1 + n(t) + m_S s(t) + m_S n(t) s(t) . \quad ...(B\text{-}6)$$

The first 2 terms on the right-hand side represent the envelope modulation of the pure radio carrier. The dc level is discarded but the second term, n(t) , is an alternating-current signal and is passed to the receiver output. The third and fourth terms are relatively smaller as determined by the modulation index m_S . The third term represents the wanted signal and is unchanged under the application of fading. The last term is a second-order noise-like term which rises in proportion to both the wanted modulation and the fading-envelope modulation.

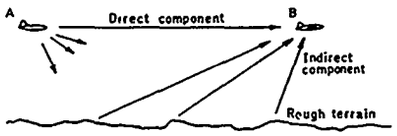
The processes described above are illustrated graphically in Figure 4 for envelope demodulation of a full-carrier amplitude-modulated signal and in Figure 5 for the counterpart dsb.sc signal.

Table 1

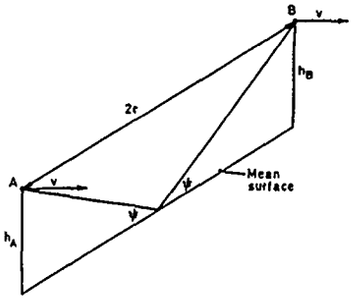
INTERFERENCE INTENSITY AGAINST AIRCRAFT ALTITUDE AND CONFIGURATION

Alt (1000')	Aircraft configuration			
	(a)	(b)	(c)	(d)
10	None	None	None	None
9	None	None	None	None
8	None	None	Slight	Slight
7	None	None	Slight/Med	Slight/Med
6	-	-	Med	Med
5	-	-	Med/Bad	Med/Bad
4	-	-	Bad	Bad
3	-	-	Severe	Severe
2	-	-	V severe	V severe

(a) line astern Canberra leading
(b) line astern BAC 1-11 leading
(c) Canberra 500 ft below BAC 1-11
(d) BAC 1-11 500 ft below Canberra



a Propagation over rough terrain



b Definition of geometry

Fig 1a & b Propagation in ground support flying scenario

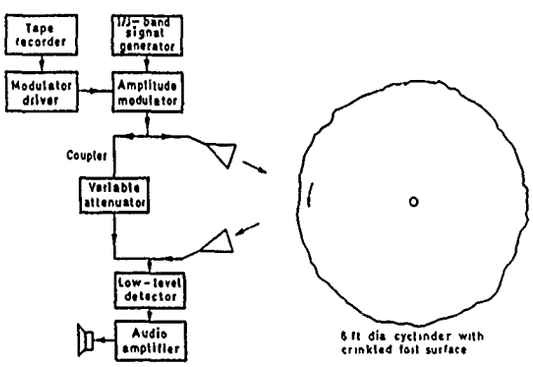


Fig 2 Microwave simulation of multipath propagation

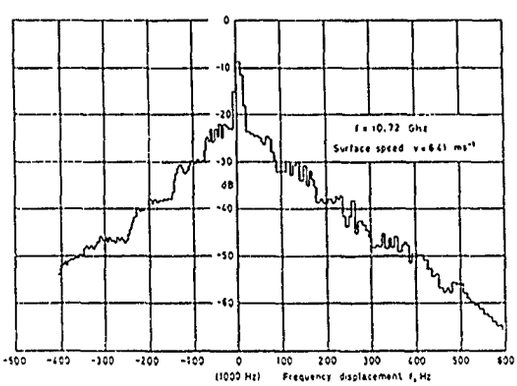


Fig 3 Spectrum from microwave simulation

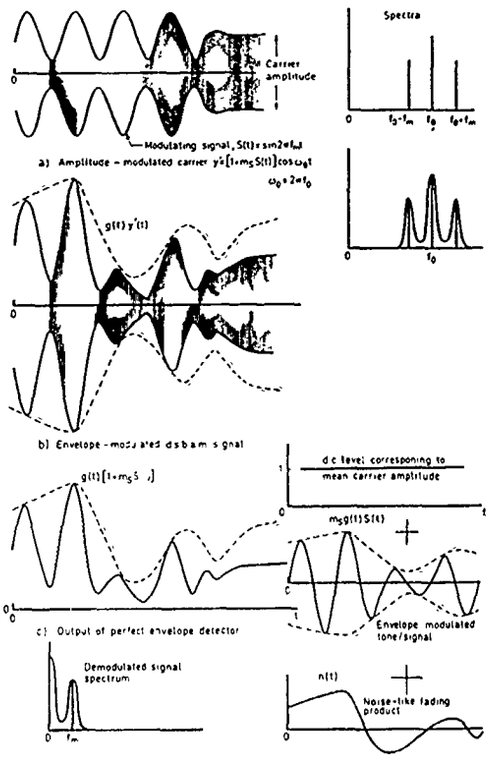


Fig 4a - d Transmission of dsbam in an envelope-fading channel

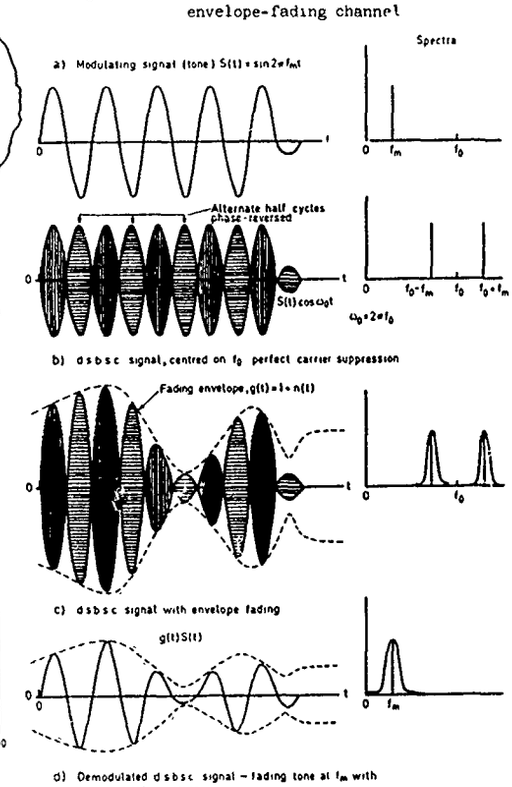


Fig 5a - d Transmission of dsbsc in an envelope-fading channel

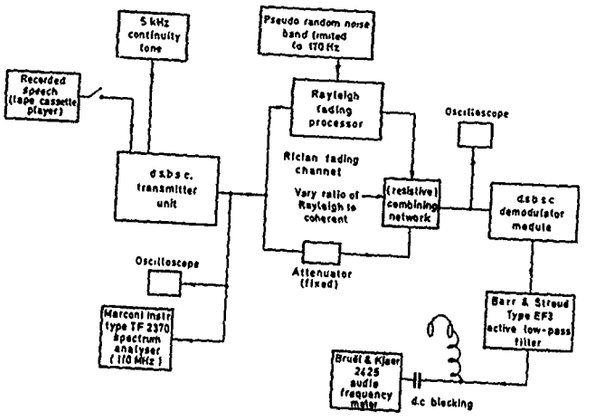
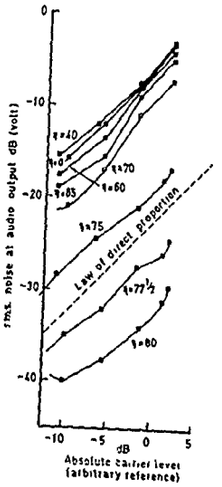


Fig 6 Audio performance test circuit



Parameter 'η'	Absolute level of Rayleigh component dBm
80	-30
77½	-25
75	-21.6
70	18.5
65	16.2
60	-15.5
50	-13.5
40	-12
0	-10

Fig 7 Audio noise versus absolute carrier level

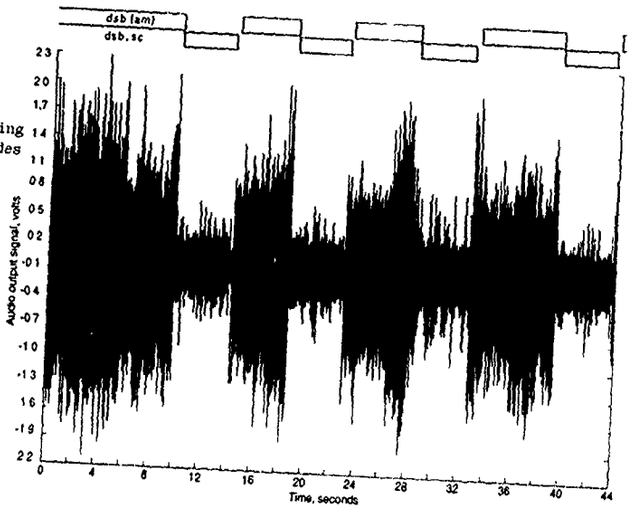


Fig 8 Transcript of tape recording of 'multipathing' in 2 modes

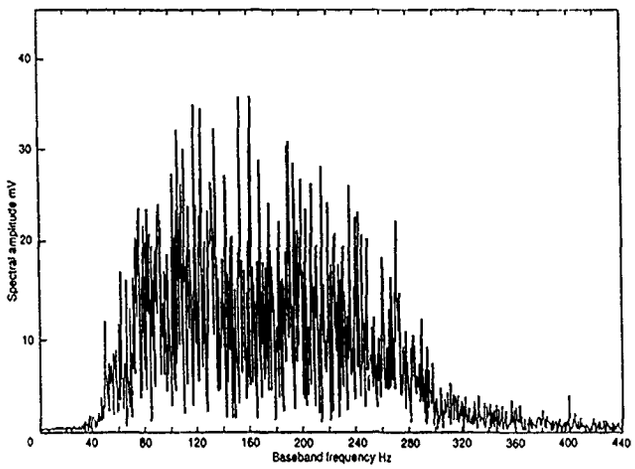


Fig 9 FFT spectrum of multipathing on dsb(am) - no voice modulation

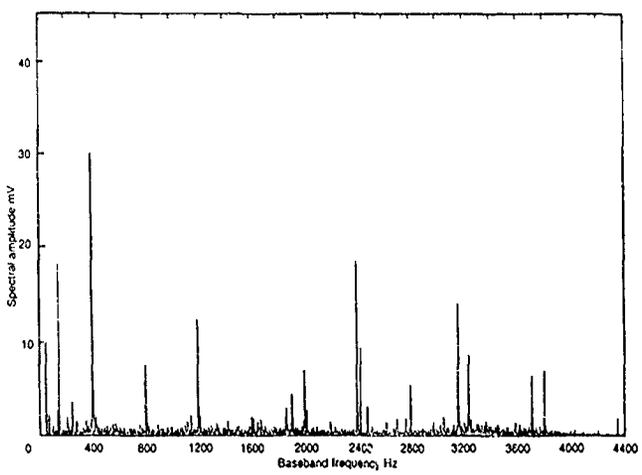
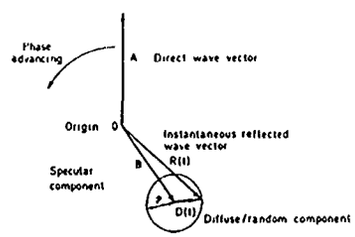
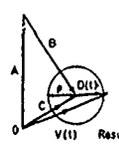


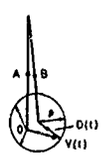
Fig 10 FFT spectrum of multipathing on dsb.sc - no voice modulation



a) Direct and reflected wave vectors



b) Vector addition of direct and reflected waves



c) Predominance of random component

Fig 11a - c Phase-space vector diagrams for multipath interference

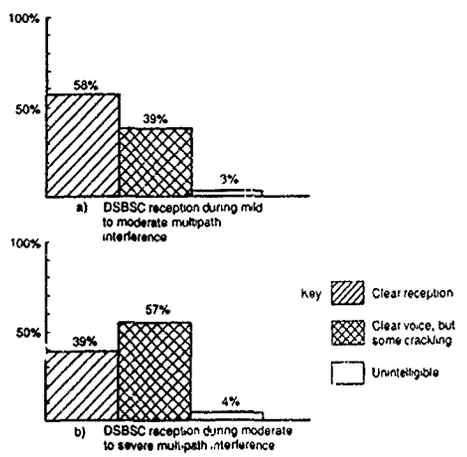


Fig 12 Bar charts showing quality of DSBSC reception during multipath interference

DISCUSSION

C. GOUTELARD

Avez-vous utilisé la même puissance dans les 2 types de modulations ? Si oui, à l'effet que vous signalez concernant les multitrajets s'ajoute le fait que toute l'énergie est concentrée dans les bandes latérales lorsque vous supprimez la porteuse, ce qui améliore la transmission. Avez-vous envisagé d'utiliser la modulation de fréquence ?

AUTHOR'S REPLY

Non, nous n'avons pas utilisé la même puissance totale dans les deux types de modulation. Nous avons mis environ la même puissance dans les bandes latérales pour les deux types de modulation. A l'émetteur nous avons sélectionné un type ou l'autre par l'addition, ou non, de la porteuse.

Oui, nous avons envisagé d'utiliser la modulation de fréquence, mais nous n'avions pas (et maintenant, nous n'avons pas) d'équipement de modulation de fréquence en "UHF". Nous croyons que la modulation de fréquence à bande de fréquence étroite n'améliore pas le bruit de "multipathing".

D. YAVUZ

Can you comment on the existence of this effect on the first generation slow hopping UHF radios (Have Quick) without of course going into the classified details.

AUTHOR'S REPLY

It is expected that "multipathing" interference (g.v.) should affect Have Quick radios in the similar circumstances to those described in the paper, namely air-to-air line-of-sight at relatively low altitudes. Recent (1988) RAE flying trials have demonstrated the effect in relatively slow passenger transport aircraft.

REPORT DOCUMENTATION PAGE

1. Recipient's Reference	2. Originator's Reference	3. Further Reference	4. Security Classification of Document
	AGARD-CP-486	ISBN 92-835-0601-4	UNCLASSIFIED

5. Originator Advisory Group for Aerospace Research and Development
North Atlantic Treaty Organization
7 rue Ancelle, 92200 Neuilly sur Seine, France

6. Title USE OR REDUCTION OF PROPAGATION AND NOISE EFFECTS IN DISTRIBUTED MILITARY SYSTEMS

7. Presented at the Electromagnetic Wave Propagation Panel Symposium, held in Rethymno, Crete, Greece, 15th to 18th October 1990.

8. Author(s)/Editor(s)	9. Date
Various	June 1991
10. Author's/Editor's Address	**11. Pages**
Various	298

12. Distribution Statement This document is distributed in accordance with AGARD policies and regulations, which are outlined on the back covers of all AGARD publications.

13. Keywords/Descriptors

Electromagnetic propagation
Distributed systems
Effects of the medium
ECM and ECCM
Signal processing

14. Abstract

This publication reports the papers presented to a symposium held by the Electromagnetic Wave Propagation Panel at its Fall 1990 meeting.

The topics covered on the occasion of that symposium include:

— Effects on distributed systems of the space, the time and the frequency coherence of waves.

— Jamming and noise reduction effects on distributed systems:
  — coherence jamming and noise versus space, time, frequency;
  — ECM and ECCM.

— Distributed systems:
  — radiocommunications: cellular systems, multi receivers or transmitters;
  — radar systems: synthetic aperture, imagery, holography, multistatic systems, multifrequency systems;
  — satellites: multifrequency systems, range finding;
  — ECM and ECCM in the systems.

AGARD Conference Proceedings No.486
Advisory Group for Aerospace Research and Development, NATO
USE OR REDUCTION OF PROPAGATION AND NOISE EFFECTS IN DISTRIBUTED MILITARY SYSTEMS
Published June 1991
298 pages

This publication reports the papers presented to a symposium held by the Electromagnetic Wave Propagation Panel at its Fall 1990 meeting.

The topics covered on the occasion of that symposium include:

— Effects on distributed systems of the space, the time and the frequency coherence of waves.

P.T.O.

AGARD-CP-486

Electromagnetic propagation
Distributed systems
Effects of the medium
ECM and ECCM
Signal processing

AGARD Conference Proceedings No.486
Advisory Group for Aerospace Research and Development, NATO
USE OR REDUCTION OF PROPAGATION AND NOISE EFFECTS IN DISTRIBUTED MILITARY SYSTEMS
Published June 1991
298 pages

This publication reports the papers presented to a symposium held by the Electromagnetic Wave Propagation Panel at its Fall 1990 meeting.

The topics covered on the occasion of that symposium include:

— Effects on distributed systems of the space, the time and the frequency coherence of waves.

P.T.O.

AGARD-CP-486

Electromagnetic propagation
Distributed systems
Effects of the medium
ECM and ECCM
Signal processing

AGARD Conference Proceedings No.486
Advisory Group for Aerospace Research and Development, NATO
USE OR REDUCTION OF PROPAGATION AND NOISE EFFECTS IN DISTRIBUTED MILITARY SYSTEMS
Published June 1991
298 pages

This publication reports the papers presented to a symposium held by the Electromagnetic Wave Propagation Panel at its Fall 1990 meeting.

The topics covered on the occasion of that symposium include:

— Effects on distributed systems of the space, the time and the frequency coherence of waves.

P.T.O.

AGARD-CP-486

Electromagnetic propagation
Distributed systems
Effects of the medium
ECM and ECCM
Signal processing

AGARD Conference Proceedings No.486
Advisory Group for Aerospace Research and Development, NATO
USE OR REDUCTION OF PROPAGATION AND NOISE EFFECTS IN DISTRIBUTED MILITARY SYSTEMS
Published June 1991
298 pages

This publication reports the papers presented to a symposium held by the Electromagnetic Wave Propagation Panel at its Fall 1990 meeting.

The topics covered on the occasion of that symposium include:

— Effects on distributed systems of the space, the time and the frequency coherence of waves.

P.T.O.

AGARD-CP-486

Electromagnetic propagation
Distributed systems
Effects of the medium
ECM and ECCM
Signal processing

– Jamming and noise reduction effects on distributed systems:
  – coherence jamming and noise versus space, time, frequency;
  – ECM and ECCM.

– Distributed systems:
  – radiocommunications: cellular systems, multi receivers or transmitters;
  – radar systems: synthetic aperture, imagery, holography, multistatic systems, multi-frequency systems;
  – satellites: multifrequency systems, range finding;
  – ECM and ECCM in the systems.

ISBN 92-835-0601-4

– Jamming and noise reduction effects on distributed systems:
  – coherence jamming and noise versus space, time, frequency;
  – ECM and ECCM.

– Distributed systems:
  – radiocommunications: cellular systems, multi receivers or transmitters;
  – radar systems: synthetic aperture, imagery, holography, multistatic systems, multi-frequency systems;
  – satellites: multifrequency systems, range finding;
  – ECM and ECCM in the systems.

ISBN 92-835-0601-4

– Jamming and noise reduction effects on distributed systems:
  – coherence jamming and noise versus space, time, frequency;
  – ECM and ECCM.

– Distributed systems:
  – radiocommunications: cellular systems, multi receivers or transmitters;
  – radar systems: synthetic aperture, imagery, holography, multistatic systems, multi-frequency systems;
  – satellites: multifrequency systems, range finding;
  – ECM and ECCM in the systems.

ISBN 92-835-0601-4

– Jamming and noise reduction effects on distributed systems:
  – coherence jamming and noise versus space, time, frequency;
  – ECM and ECCM.

– Distributed systems:
  – radiocommunications: cellular systems, multi receivers or transmitters;
  – radar systems: synthetic aperture, imagery, holography, multistatic systems, multi-frequency systems;
  – satellites: multifrequency systems, range finding;
  – ECM and ECCM in the systems.

ISBN 92-835-0601-4

NATO OTAN

7 RUE ANCELLE · 92200 NEUILLY-SUR-SEINE

FRANCE

Telephone (1)47.38.57.00 · Télex 610 176
Télécopie (1)47.38.57.99

DIFFUSION DES PUBLICATIONS
AGARD NON CLASSIFIEES

L'AGARD ne détient pas de stocks de ses publications, dans un but de distribution générale à l'adresse ci-dessus. La diffusion initiale des publications de l'AGARD est effectuée auprès des pays membres de cette organisation par l'intermédiaire des Centres Nationaux de Distribution suivants. A l'exception des Etats-Unis, ces centres disposent parfois d'exemplaires additionnels; dans les cas contraire, on peut se procurer ces exemplaires sous forme de microfiches ou de microcopies auprès des Agences de Vente dont la liste suit.

CENTRES DE DIFFUSION NATIONAUX

ALLEMAGNE
Fachinformationszentrum,
Karlsruhe
D-7514 Eggenstein-Leopoldshafen 2

BELGIQUE
Coordonnateur AGARD-VSL
Etat-Major de la Force Aérienne
Quartier Reine Elisabeth
Rue d'Evere, 1140 Bruxelles

CANADA
Directeur du Service des Renseignements Scientifiques
Ministère de la Défense Nationale
Ottawa, Ontario K1A 0K2

DANEMARK
Danish Defence Research Board
Ved Idraetsparken 4
2100 Copenhagen Ø

ESPAGNE
INTA (AGARD Publications)
Pinar 7 Rosales 34
28008 Madrid

ETATS-UNIS
National Aeronautics and Space Administration
Langley Research Center
MS 180
Hampton, Virginia 23665

FRANCE
O.N.E.R.A. (Direction)
29, Avenue de la Division Leclerc
92320, Châtillon sous Bagneux

GRECE
Hellenic Air Force
Air War College
Scientific and Technical Library
Dekelia Air Force Base
Dekelia, Athens TGA 1010

ISLANDE
Director of Aviation
c/o Flugrad
Reykjavik

ITALIE
Aeronautica Militare
Ufficio del Delegato Nazionale all'AGARD
3 Piazzale Adenauer
00144 Roma/EUR

LUXEMBOURG
Voir Belgique

NORVEGE
Norwegian Defence Research Establishment
Attn: Biblioteket
P.O. Box 25
N-2007 Kjeller

PAYS-BAS
Netherlands Delegation to AGARD
National Aerospace Laboratory NLR
Kluyverweg 1
2629 HS Delft

PORTUGAL
Portuguese National Coordinator to AGARD
Gabinete de Estudos e Programas
CLAFA
Base de Alfragide
Alfragide
2700 Amadora

ROYAUME UNI
Defence Research Information Centre
Kentigern House
65 Brown Street
Glasgow G2 8EX

TURQUIE
Milli Savunma Başkanliği (MSB)
ARGE Daire Başkanliği (ARGE)
Ankara

LE CENTRE NATIONAL DE DISTRIBUTION DES ETATS-UNIS (NASA) NE DETIENT PAS DE STOCKS DES PUBLICATIONS AGARD ET LES DEMANDES D'EXEMPLAIRES DOIVENT ETRE ADRESSEES DIRECTEMENT AU SERVICE NATIONAL TECHNIQUE DE L'INFORMATION (NTIS) DONT L'ADRESSE SUIT.

AGENCES DE VENTE

National Technical Information Service
(NTIS)
5285 Port Royal Road
Springfield, Virginia 22161
Etats-Unis

ESA/Information Retrieval Service
European Space Agency
10, rue Mario Nikis
75015 Paris
France

The British Library
Document Supply Division
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
Royaume Uni

Les demandes de microfiches ou de photocopies de documents AGARD (y compris les demandes faites auprès du NTIS) doivent comporter la dénomination AGARD, ainsi que le numéro de série de l'AGARD (par exemple AGARD-AG-315). Des informations analogues, telles que le titre et la date de publication sont souhaitables. Veuiller noter qu'il y a lieu de spécifier AGARD-R-nnn et AGARD-AR-nnn lors de la commande de rapports AGARD et des rapports consultatifs AGARD respectivement. Des références bibliographiques complètes ainsi que des résumés des publications AGARD figurent dans les journaux suivants:

Scientifique and Technical Aerospace Reports (STAR)
publié par la NASA Scientific and Technical
Information Division
NASA Headquarters (NTT)
Washington D.C. 20546
Etats-Unis

Government Reports Announcements and Index (GRA&I)
publié par le National Technical Information Service
Springfield
Virginia 22161
Etats-Unis

(accessible également en mode interactif dans la base de données bibliographiques en ligne du NTIS, et sur CD-ROM)

Imprimé par Specialised Printing Services Limited
40 Chigwell Lane, Loughton, Essex IG10 3TZ

NATO OTAN
7 RUE ANCELLE · 92200 NEUILLY SUR-SEINE
FRANCE
Telephone (1)47.38.57.00 Telex 610 176
Telefax (1)47.38.57.99

DISTRIBUTION OF UNCLASSIFIED AGARD PUBLICATIONS

AGARD does NOT hold stocks of AGARD publications at the above address for general distribution. Initial distribution of AGARD publications is made to AGARD Member Nations through the following National Distribution Centres. Further copies are sometimes available from these Centres (except in the United States), but if not may be purchased in Microfiche or Photocopy form from the Sales Agencies listed below.

NATIONAL DISTRIBUTION CENTRES

BELGIUM
Coordonnateur AGARD – VSL
Etat-Major de la Force Aérienne
Quartier Reine Elisabeth
Rue d'Evere, 1140 Bruxelles

CANADA
Director Scientific Information Services
Dept of National Defence
Ottawa, Ontario K1A 0K2

DENMARK
Danish Defence Research Board
Ved Idraetsparken 4
2100 Copenhagen Ø

FRANCE
O.N.E.R.A. (Direction)
29 Avenue de la Division Leclerc
92320 Chatillon

GERMANY
Fachinformationszentrum
Karlsruhe
D-7514 Eggenstein-Leopoldshafen 2

GREECE
Hellenic Air Force
Air War College
Scientific and Technical Library
Dekelia Air Force Base
Dekelia, Athens TGA 1010

ICELAND
Director of Aviation
c/o Flugstad
Reykjavik

ITALY
Aeronautica Militare
Ufficio del Delegato Nazionale all'AGARD
3 Piazzale Adenauer
00144 Roma/EUR

LUXEMBOURG
See Belgium

NETHERLANDS
Netherlands Delegation to AGARD
National Aerospace Laboratory, NLR
Kluyverweg 1
2629 HS Delft

NORWAY
Norwegian Defence Research Establishment
Attn: Biblioteket
P.O. Box 25
N-2007 Kjeller

PORTUGAL
Portuguese National Coordinator to AGARD
Gabinete de Estudos e Programas
CLAFA
Base de Alfragide
Alfragide
2700 Amadora

SPAIN
INTA (AGARD Publications)
Pinar Rosales 34
28008 Madrid

TURKEY
Milli Savunma Başkanlığı (MSB)
ARGE Daire Başkanlığı (ARGE)
Ankara

UNITED KINGDOM
Defence Research Information Centre
Kentigern House
65 Brown Street
Glasgow G2 8EX

UNITED STATES
National Aeronautics and Space Administration (NASA)
Langley Research Center
M/S 180
Hampton, Virginia 23665

THE UNITED STATES NATIONAL DISTRIBUTION CENTRE (NASA) DOES NOT HOLD STOCKS OF AGARD PUBLICATIONS, AND APPLICATIONS FOR COPIES SHOULD BE MADE DIRECT TO THE NATIONAL TECHNICAL INFORMATION SERVICE (NTIS) AT THE ADDRESS BELOW.

SALES AGENCIES

National Technical
Information Service (NTIS)
5285 Port Royal Road
Springfield, Virginia 22161
United States

ESA/Information Retrieval Service
European Space Agency
10, rue Mario Nikis
75015 Paris
France

The British Library
Document Supply Centre
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
United Kingdom

Requests for microfiches or photocopies of AGARD documents (including requests to NTIS) should include the word 'AGARD' and the AGARD serial number (for example AGARD-AG-315). Collateral information such as title and publication date is desirable. Note that AGARD Reports and Advisory Reports should be specified as AGARD-R-nnn and AGARD-AR-nnn, respectively. Full bibliographical references and abstracts of AGARD publications are given in the following journals:

Scientific and Technical Aerospace Reports (STAR)
published by NASA Scientific and Technical
Information Division
NASA Headquarters (NTT)
Washington D.C. 20546
United States

Government Reports Announcements and Index (GRA&I)
published by the National Technical Information Service
Springfield
Virginia 22161
United States
(also available online in the NTIS Bibliographic Database or on CD-ROM)

Printed by Specialised Printing Services Limited
40 Chigwell Lane, Loughton, Essex IG10 3TZ

ISBN 92-835-0601-4