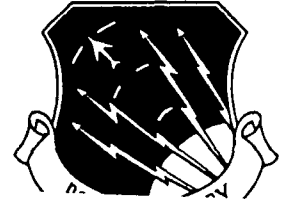AD-A238 279

ACOUSTIC BACKGROUND NOISE VARIATION IN AIR FORCE PLATFORMS AND ITS EFFECT ON NOISE REMOVAL ALGORITHMS

ARCON Corporation

Philip A. LaFollette
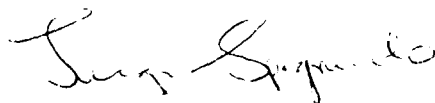
91-04959

91   1   058

This report has been reviewed by the Rome Laboratory Public Affairs Office (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

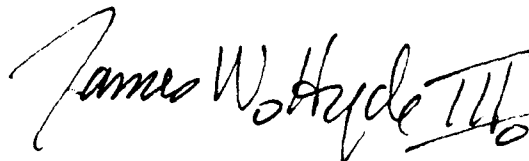RL-TR-91-97 has been reviewed and is approved for publication.

APPROVED:

LUIGI SPAGNUOLO
Project Engineer

APPROVED:

JOHN K. SCHINDLER
Director of Electromagnetics

FOR THE COMMANDER:

JAMES W. HYDE III
Directorate of Plans & Programs

REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	June 1991	Interim Aug 88 – Feb 90

4. TITLE AND SUBTITLE
ACOUSTIC BACKGROUND NOISE VARIATION IN AIR FORCE PLATFORMS AND ITS EFFECT ON NOISE REMOVAL ALGORITHMS

5. FUNDING NUMBERS
C - F19628-88-C-0190
PE - 33401F
PR - 7820
TA - 03
WU - 59

6. AUTHOR(S)
Philip A. LaFollette

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
ARCON Corporation
260 Bear Hill Road
Waltham MA 02155

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Rome Laboratory (EEV)
Hanscom AFB MA 01731-5000

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
RL-TR-91-97

11. SUPPLEMENTARY NOTES
Rome Laboratory Project Engineer: Luigi Spagnuolo/EEV/(617) 478-4249
NOTE: Rome Laboratory (formerly Rome Air Development Center/RADC)

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

In this study of short-term noise variation in Air Force platforms, we followed two avenues of investigation. First, we applied quantitative measures of variation to individual noise recordings, and compared the results across various aircraft. Some measures used were simple descriptive statistics, but we also measured attenuation obtained by spectral restoration (spectral subtraction), applied to the noise signal alone. The noise attenuation obtained for real aircraft environments was in most cases about the same as predicted theoretically for white Gaussian noise, but in some instances was considerably higher, especially in the presence of propeller noise. Second, we applied the nonparametric Mann-Whitney statistic to test the stationarity of power spectrum estimates on time scales of 200 to 800 ms. There was little or no evidence of nonstationarity in large jet or turboprop aircraft. In fighter aircraft and helicopters, there was some evidence of nonstationarity confined to more or less narrow frequency ranges. The nonstationarity found did not appear to limit the performance of spectral restoration algorithms. The noise recordings used were taken from the RADC/EEV database of field recordings made in the E-3A, E-4B, EC-135, E-130, P-3C, F-15, F-16, F-4, A-10, HH-53 and Tornado aircraft.

14. SUBJECT TERMS Aircraft interior noise, Acoustic noise, Nonstationary acoustic noise, Digital voice processing, Speech compression, Noise suppression, Speech enhancement, Spectral subtraction

15. NUMBER OF PAGES
114

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	U/L

Contents

List of Tables

List of Figures

v

Acknowledgments

Chapter 1

Introduction

1.1   The Acoustic Noise Problem in Vocoders

A major problem with narrowband digital voice processors is the degradation of
their performance by background acoustic noise. Digital voice communication
systems utilized by the Air Force are required to operate on a large variety of
military platforms. The acoustic noise environment is a function of the specific
platform, the operational mode of the platform, the location of speaker and mi-
crophone, and the noise-shielding (or noise-introducing) characteristics of equip-
ment such as oxygen masks. Various noise reduction and noise removal methods
have been tried to solve this problem for specific platforms and processors, but
none have been successful at improving the intelligibility of narrowband voice
communication in a noisy environment, as measured by the accepted Diagnostic
Rhyme Test.

It would be futile to attempt to characterize acoustic noise completely as
a deterministic function of all the operational variables, but it can be help-
ful to examine the *range* of acoustic noise phenomena facing Air Force speech
communication systems. Until recent years, no research had been directed at
characterizing and categorizing the broad range of noise environments of interest
to the Air Force, as they affect narrowband voice processors and noise reduction
techniques. Research conducted by ARCON [35] surveyed the long-term char-
acteristics of those noise environments, using the acoustic noise data library of
the RADC/EEV Speech Processing Facility. That study obtained representa-
tive noise power spectrum estimates for recordings made aboard various classes
of aircraft.

This report contains the results of a further study directed at the short-term
variation of noise in the same environments. The time intervals over which
we look for variation may be as short as 20 ms. For comparison, we note
that vocoders analyze speech in 20-40 ms time increments, and conversational

speech contains about 10 phonemes per second [13]. For reasons related to noise-removal methods, we are interested in variation over periods on the order of 1-2 s as well.

In Appendix B, we include some further information obtained about long-term noise characteristics aboard additional aircraft not covered in the earlier report [35].

In discussing the acoustic noise problem for airborne speech communication systems, we have to distinguish between two types of acoustic noise in operational aircraft. On one hand there is what we will call "inherent" noise of the aircraft, arising from

1. Turbulent airflow and mechanical vibration associated with the engines, turbines, exhausts, and propellers;

2. Turbulent airflow around the rest of the aircraft;

3. Vibration of the aircraft's structure excited ultimately by the above two sources.

This "inherent" noise is the noise arising because the aircraft is flying in a certain control configuration through a certain external aerodynamic environment. Contrasted with "inherent" noise is noise arising from operations within the aircraft, such as the acoustic noise caused by weapons, communications equipment, or other speakers. It is difficult to predict or classify the effects of such "operational" noise sources, and they will not be discussed in this report.

Before the acoustic noise and the acoustic speech signal are processed by narrowband speech systems, they are influenced by other factors. The absolute level of the speech itself is controlled, within a certain range, by the speaker. Noise-cancelling microphones may reduce noise levels. Time-domain noise cancellation or frequency-domain spectral restoration may be employed as a signal processing step before the vocoder input. A further factor is the noise-suppression effect of oxygen masks. In aircraft such as the F-15, the microphone is inside the oxygen mask, and the mask itself provides much-needed attenuation of the aircraft's acoustic noise [11]. On the other hand, this attenuation is variable due to the opening and closing of the valves, and the mask introduces some distortion of the pilot's speech.

The overall acoustic noise power is probably the most often quoted attribute of an aircraft's acoustic noise environment. However, for the processing of speech against this noise background, other attributes of the noise may be more significant. Speech processing is not affected by noise outside the passbands of the analog anti-alias and high-pass filters commonly employed in analyzers. Thus, for narrowband systems, the frequency range of prime importance extends from

100 Hz to about 4000 Hz. Noise outside this range could affect a listener located in the aircraft, but would not directly corrupt an analyzer's voice processing.

Furthermore, to assess the impact of acoustic noise on speech processing, frequency-domain properties of acoustic noise should be compared with the frequency-domain properties of speech. Although speech is inherently highly variable in its spectral shape, the physical shape of the human vocal tract produces a long term average spectral distribution that is not flat. In fact, one motivation for the preemphasis typically applied as the first stage of digital speech processing is to compensate for the decline in typical speech energy above about 500 Hz. Although this "average shape" is to some extent dependent on the individual speaker, the long-term spectral distribution of speech is roughly flat from 100 Hz to 500 Hz and then declines about 6 dB per octave above 500 Hz. In discussing acoustic noise spectra, we should compare the spectral shape of the noise with this long-term "average" distribution of speech energy.

In evaluating noise environments from the point of view of narrowband voice communication, there are three possible classes of noise measures. The first class includes measures of the noise itself: later in this report we discuss some of these, such as short-time spectral estimation, the Mann-Whitney statistic, the standard-deviation-to-mean ratio as a function of frequency, the residual RMS deviation of noise estimates as a function of averaging time. The second class includes measures of the effect of noise variation on the performance of noise stripping techniques, notably the techniques linked to spectral restoration.

The third class, not discussed in this study, would include measures of the effect of noise variation on speech coding algorithms themselves. Such measures depend on a choice of coding algorithm; for instance, the effects of additive noise on the standard LPC-10E may be separated into:

1. Distortion of the reflection coefficients

2. Degradation of the voiced/unvoiced decision

3. Distortion of the pitch estimate

4. Distortion of the energy estimate

For characterizing the *stationary* aspects of the acoustic noise background, power spectrum estimation is the most appropriate tool, since the second-order statistics of a zero-mean stationary random process are completely characterized by its power spectrum.

Some noise backgrounds contain relatively well-behaved periodic components. If these components are well separated, they may be susceptible to noise-reduction processing techniques based on tone removal, but they might pose a special problem to strategies that assume background noise has a continuous spectral distribution.

1.2   Previous ARCON Study

In the long-term noise characterization study [35], we found that the noise environments were grouped in four classes.

In the first group are large aircraft with wing-mounted jet engines. Aboard these aircraft, the bulk of the acoustic noise power is concentrated at frequencies less than 1000 Hz. Above 1000 Hz, the noise power drops off at 6-12 dB per octave, compared to the decline of 6 dB per octave in typical speech signals. From the narrowband communication point of view, such a shape is desirable both in terms of reduced competition with higher-formant information in speech and in terms of susceptibility to noise-cancelling microphones, which perform better at low frequencies.

The second group consists of large aircraft with wing-mounted turboprop engines. These turboprop engines and their propellers produce more low-frequency noise than jet engines do, and the difference is apparent in long-term noise power spectra.

The third group consists of smaller fighter aircraft with jet engines. In these aircraft there is substantial noise power distributed all across the frequency range studied, and even higher. Noise-cancelling microphones are of little help with this high-frequency noise. Moreover, in our previous study we found strong line components varying in frequency, which would be expected to cause severe problems for spectral restoration processing techniques.

The fourth group consists of helicopters, in which we have also found substantial noise power distributed all across the frequency range, apparently in harmonically related narrow bands.

Our past research has documented differences in acoustic noise from one compartment to another in the same aircraft, and also substantial and repeatable differences between noise measured by two microphones as little as 30 inches apart, as shown in Figure 1.1. These differences show that we should not expect high correlation between acoustic noise at two locations near one another, and further imply that we should be careful not to over-interpret details of particular acoustic noise spectra.

1.3   Other Research

In the past five years, several researchers have studied the effectiveness of two-microphone noise cancellation methods in simulated cockpit environments. Harrison *et al.* [17] used actual aircraft noise recordings played through a single loudspeaker. Darlington *et al.* [10, 28] simulated the cockpit more thoroughly, with multiple loudspeakers, and Darlington [11] measured the effect of an oxygen mask and helmet. Rodriguez [30] and Rodriguez and Lim [31] used multiple loudspeakers and gradient microphones. These studies did not address temporal variation in the noise, because two-microphone noise cancellation is not ham-
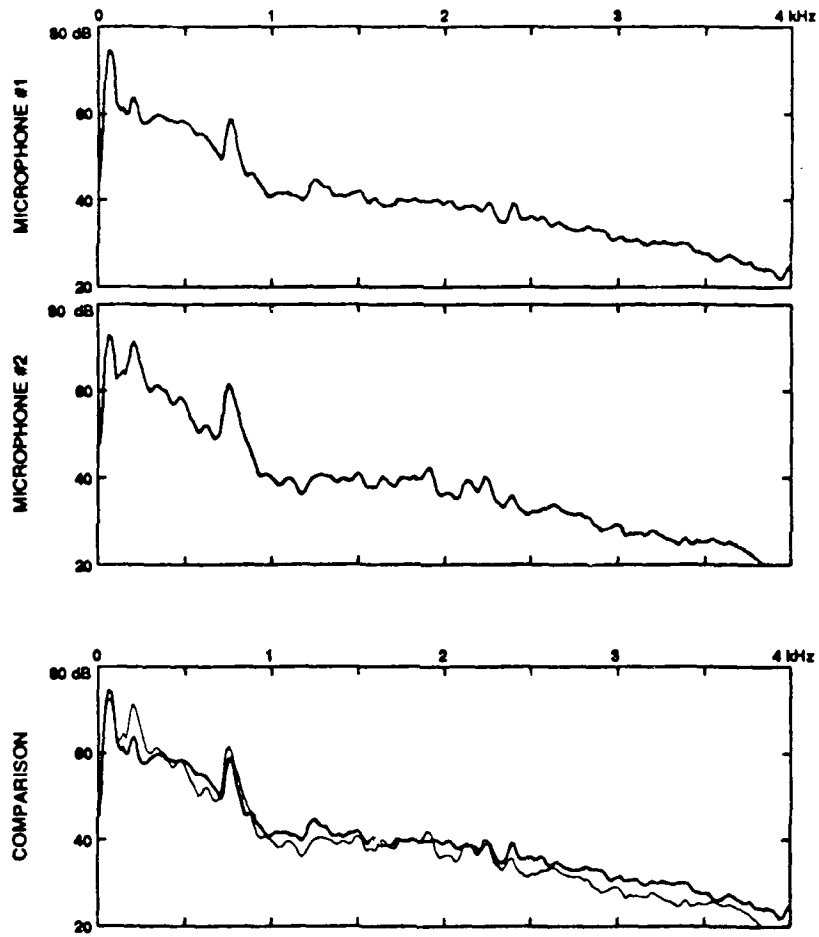
Figure 1.1: Noise spectra, EC-130, microphones 30 inches apart

pered by such variation. (*Spatial* variation, on the other hand, is important for such cancellation methods.)

Aschkenazy and Weiss [38, 4, 5, 6] have applied separate methods of tone removal, impulse removal, and spectral restoration to cockpit noise, with an emphasis on enhancement for speech recognition applications.

Since 1975, the US Air Force Armstrong Aerospace Medical Laboratory (AAMRL) has been gathering, analyzing, and reporting on the acoustic noise environment of USAF systems, including aircraft, for the purpose of assisting in environmental assessments of the exposure of flight crews, aircraft passengers, ground crews, other flight personnel, and airbase communities to acoustic noise [9]. In the series of which [9] is the first volume, they have published a large amount of data on third-octave band and octave-band analyses of aircraft acoustic noise, including interior noise. Although these third-octave analyses give a suggestion of the broad spectral shape of the environments studied, they do not have a fine enough frequency resolution to be used to judge the effects of noise environments on narrowband speech processing. (For example, a vocoder with a frame size of 25 ms has an effective analysis bandwidth of about 40 Hz.) The third-octave resolution was chosen by AAMRL largely because their emphasis was on bioenvironmental noise studies, not on the effects of acoustic noise on communications equipment.

Some of the AAMRL recordings were used in another study [25] which found an undocumented low-frequency rolloff in the recording system used to make the F-15 recordings. In addition, some recordings have been adjusted for system response utilizing a 1/3 octave filter set.

1.4   Spectral Restoration

Spectral restoration methods, as described in Section 4.3.1, are the only widely-applied general approach to the removal of noise from single-microphone speech in communication systems. Two-microphone methods, such as adaptive noise cancellation, have limited applicability in aircraft, both because of the need for a second microphone and because of incoherent noise fields.

Spectral restoration has been used in several forms [7, 8, 12, 24, 27]. The idea is to transform a noisy signal into the frequency domain and then, using an estimate of the noise power spectrum, to correct the frequency-domain representation of the noisy signal. When spectral restoration methods are applied as a preprocessor to LPC-10 speech coding, intelligibility is generally not improved, but perceived speech quality has been shown to be enhanced in aircraft noise environments [20]. Some researchers have suggested that some of the observed limitations on the performance of spectral restoration could arise from nonstationarity in the noise. Because of the importance of spectral restoration, we have taken care in this study to consider the effect of short-term noise variation on spectral restoration methods.

1.5 Time Variation: This Study

Our previous study of long-term acoustic noise characteristics [35] relied on modeling the noise as a stationary stochastic process. Likewise, spectral restoration methods rely on estimates of noise spectra based on previous measurements during silent (or unvoiced) intervals. In some cases, there is an obvious mismatch between the assumption of stationarity and the reality of a noise background. But if there is no obvious reason to suspect nonstationarity, might the noise source still be nonstationary in a more subtle way? In Chapters 5 and 6 we present the results of statistical tests of stationarity, applied to acoustic noise recorded in aircraft.

Even if the noise *is* well-modelled as the output of a station ..y stochastic process, the noise will not be the same from frame to frame nor will the discrete Fourier transform of the noise be any more predictable (from one frame to the next) than the noise itself. In Chapter 4 we also discuss measurements of time variation in terms of the consistency of discrete Fourier transforms, and in terms of the performance of spectral restoration.

Chapter 2

Sources of Acoustic Noise Data

2.1 RADC/EEV Acoustic Noise Database

The acoustic noise samples used for this study come from the Acoustic Noise Database of RADC/EEV. This database consists of noise recordings made aboard aircraft in flight. The recordings were made on various occasions from 1976 onward.

All these recordings were made with high-quality microphones with an essentially flat frequency response across the audio range. These microphones were used in preference to the resident communications microphones. One reason for this approach was the desire to separate the effects of microphone characteristics from the noise field itself. A second reason was that tapping into the resident audio system can present problems with flight qualification. A third reason was the desire to use the noise recordings as source material for sound-chamber simulations of acoustic noise fields. These simulations were used to generate intelligibility and quality test tapes using various microphones [32]. The resulting tapes provide for the evaluation of voice communication systems. A disadvantage of the use of instrumentation microphones is that the recordings do not show reductions in the effective noise field due to noise cancelling or frequency selectivity in the communications microphone.

Until recently, all the original recordings in the Acoustic Noise Database were analog tapes. In 1989, noise recordings began to be made directly in PCM formats. The database also contains PCM copies of selected analog originals [39].

2.1.1   Conditions of Recording

With a few exceptions, the recordings in the database come from efforts in which noise-only recordings were made with a view to preparing source material for speech intelligibility and quality testing. There are three sources of such data:

1. Field recordings made by RADC/EEV and contract personnel between 1979 and 1984 [34, 37, 35];

2. Field recordings made by BBN personnel in 1981 [25];

3. Field recordings made by Ketron personnel in 1978 [36].

RADC/EEV Tapes

Noise recordings were made at a number of locations in an E-3A (AWACS) by C. P. Smith in August 1979. These recordings include electrical calibrations and have sound-level documentation. Although all the recordings have audible talkers in the background, only one has a talker near to the recording microphone. Two-microphone recordings were made, the microphones placed at the left and right side of each operator location.

In June–July 1982, single-microphone recordings were made by C. P. Smith with C. Teacher of KETRON Corp., aboard an E-4B Airborne Command Post and an EC-135 command and control aircraft. These are single-microphone recordings with electrical calibration, supported with noise level documentation.

In June–August 1984, further noise recordings were made by D. Robitaille and L. Spagnuolo of RADC/EEV. The aircraft covered by these recordings were an HH-53 search and rescue helicopter, an EC-130 turboprop with an ABCCC module, and an HC-130 turboprop. These recordings were made with a two-microphone configuration, and supported with noise level documentation.

BBN Tapes

The database includes originals and PCM copies of noise recordings made in January 1981 under the supervision of Miller *et al.* of Bolt Beranek and Newman Inc., who were working under contract to M. I. T. Lincoln Laboratory. The purpose of this effort was to obtain noise recordings that could be used as source material for sound-chamber simulations of field environments relevant to the JTIDS program. Aircraft covered are the F-15A, F-15B, F-16A, A-10, and the F-4E. These are two-microphone recordings with full acoustic calibration. One microphone was located on a pilot's helmet, and the other was placed inside his oxygen mask near the communications microphone.

KETRON Tapes

The database also includes copies of 1978 noise recordings made by C. Teacher and H. Watkins of KETRON, Inc. [36] Their effort obtained extensive recordings

of noise and wordlists in environments ranging over aircraft, ships, and ground vehicles. The aircraft covered were the RH-53 helicopter, the EC-135 command and control aircraft, the KC-135 tanker, the C-130 and P-3C turboprop, and the C-141. These recordings were supported with noise level documentation.

2.1.2 Recordings Selected

From the RADC/EEV acoustic noise database, Table 2.1 lists the source recordings that were judged appropriate for the purposes of this study.

2.2 Calibration Methods

If absolute noise levels are needed, it is always necessary to compensate for the various gains and sensitivities in the recording and playback systems. When digital recordings are made for the Noise Database, these systems extend all the way from the instrumentation microphone to the input to an A/D converter. The calibration process measures the power gain, from the original acoustic field to the input of an A/D converter, stated as a ratio of the mean square of the output (expressed in V^2) to the mean square of the input overpressure (in units of the square of the SPL reference, an overpressure of 20 μPa). Then a system gain of 0 dB means that an acoustic field with an RMS overpressure of 20 μPa produces a signal with an RMS of 1 V at the input to the A/D converter. For a specific A/D conversion, the system gain all the way to the digitized file can be specified, replacing V^2 with squared A/D converter counts, and a system gain of 0 dB means that the same acoustic field produces a digitized file of integers with an RMS of 1.

Depending on the recording, several different calibration methods are used. An acoustic calibrator is a device that fits over the element of an instrumentation microphone and creates a known sound field at the microphone, typically a sine wave of 1000 Hz at a sound pressure level (SPL) of about 95 dB. If the microphone's output in this condition is then recorded on the same system used to record data, then the overall gain of the recording and playback systems can be deduced and compensated for.

In the absence of a full acoustic calibration, it is still possible to measure the gain of the recording/playback system using the acoustic noise itself, provided that the noise power remains essentially unchanged throughout the recording. In this case, the original acoustic noise level is measured at the microphone position, either as SPL or as a weighted sound level, while the recording is being made. On playback, the mean square of the playback signal is measured (in V^2), using the identical weighting if any. Comparing the two measurements gives us an overall system gain, as a ratio of the mean square of the output (in V^2) to the mean square of the input overpressure (in units of the square of the SPL reference, an overpressure of 20 μPa).

Table 2.1: Noise recordings available for this study

Aircraft	Location	Tape #	Duration	Comments	Ref.
EC-135	Radio compartment	28N	12 min		[37]
	Battle Staff area	21N	13 min		[37]
E-3A	Senior Dir. (#4)	206	16 min		[34]
	Air Surv. Tech. (#10)	203	15 min	nearby talker	[34]
	Weapons Dir. (#13)	201	13 min		[34]
	Console #25	205	14 min	distant talkers	[34]
	Console #30	204	14 min		[34]
E-4B	Battle Staff area	12N	5 min		[37]
	Briefing room	12N	5 min		[37]
	NCA compartment	15N	5 min		[37]
P-3C	NAVSEA pos.	ND	40 s		[36]
EC-130	ABCCC module	C	5 min		[35]
	Seat #1	213-1A	13 min	nearby talker	[35]
HC-130	Radio operator	1A	5 min		[35]
HH-53	Cockpit rear bulkhead	1AA	14 min		[35]
F-15A	Near pilot	HHNT	8-12 s	many segments, outside mask	
F-15A	Pilot (helmet)	5-2	32 min		[25]
	Pilot (helmet)	10-2	26 min		[25]
	Pilot (helmet)	9-2	39 min		[25]
	Pilot (O_2 mask)	5-1	32 min	held breath 30 s	[25]
	Pilot (O_2 mask)	10-1	26 min	held breath 16 s	[25]
F-16A	Pilot (helmet)	6-2	39 min		[25]
	Pilot (O_2 mask)	6-1	39 min	held breath 10 s	[25]
F-4E	Pilot (helmet)	11-2	39 min		[25]
	Pilot (O_2 mask)	11-1	39 min	held breath 14 s	[25]
A-10	Pilot (helmet)	8-2	39 min		[25]
	Pilot (O_2 mask)	8-1	39 min	held breath twice	[25]
Tornado	Pilot	TOR	8 min	outside mask	
Tornado	Navigator	TOR	8 min	outside mask	
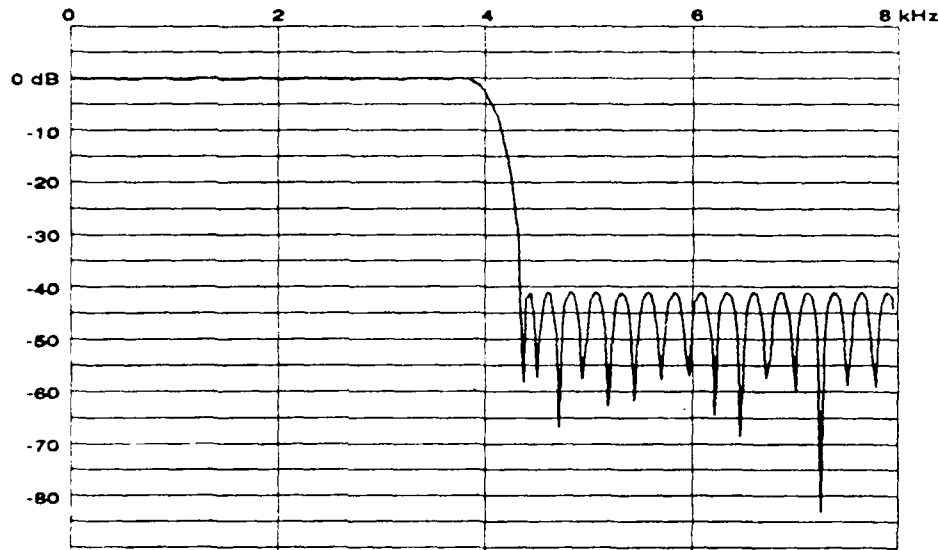
Figure 2.1: Low-pass filter used for downsampling

Finally, some recordings have neither a full acoustic calibration nor a noise-level calibration. These may have an electrical calibration consisting of a tone recorded with a known electrical signal at the input to the record amplifier, so that acoustic levels can be inferred only if the sensitivity of the microphone is known.

2.3   Selection of Representative Segments

Appendix A contains detailed information about the digitized noise segments used. In this section, we discuss the general issues addressed in the selection and preparation of the noise segments.

Table 2.2 lists the digital acoustic noise files used by this study from the pre-existing RADC/EEV Acoustic Noise Database. For this study, digitized noise records for several aircraft were added to the existing database, as shown in Table 2.3. Most of these new records are taken from BBN tapes [25]. Each of the new records is a sample of noise about 5 seconds long, recorded at a time when the pilot was holding his breath in accordance with experimenters' instructions. In each case, the in-mask and on-helmet recordings were synchronized approximately by use of the recorded synchronization signal at the beginning of each tape. All files are supplied with standard data base headers.

In the case of both old and new files, the 8-kHz files were downsampled from the original 16-kHz files using a 64th-order linear-phase FIR dealiasing filter.

File name	Tape	Sampling
EC135B.NOI	21N	16 kHz
EC135B.FLT	21N	8 kHz
E3AC13.NOI	201	16 kHz
E3AC13.FLT	201	8 kHz
E3AC04.FLT	206	8 kHz
E4BBS.NOI	12N	16 kHz
E4BBS.FLT	12N	8 kHz
EC130A.NOI	C	16 kHz
EC130A.FLT	C	8 kHz
HC130A.NOI	1A	16 kHz
HC130A.FLT	1A	8 kHz
HC130B.NOI	1A	16 kHz
HC130B.FLT	1A	8 kHz
P3C.NOI	ND	16 kHz
P3C.FLT	ND	8 kHz
F15C33.NOI	HHNT	16 kHz
F15C33.FLT	HHNT	8 kHz
F15C418.NOI	HHNT	16 kHz
F15C418.FLT	HHNT	8 kHz
F15C53.NOI	HHNT	16 kHz
F15C53.FLT	HHNT	8 kHz
F15C59.NOI	HHNT	16 kHz
F15C59.FLT	HHNT	8 kHz
F15C68.NOI	HHNT	16 kHz
F15C68.FLT	HHNT	8 kHz
HH53.NOI	1AA	16 kHz
HH53.FLT	1AA	8 kHz

Table 2.2: Pre-existing digitized noise files used

The magnitude transfer function of this filter is shown in Figure 2.1. The filter's 3dB cutoff is 4kHz, resulting in essentially flat response up to 3.8 kHz at the cost of a small but tolerable amount of "folding" of energy in the 4.0 to 4.2 kHz band.

For each file listed in Table 2.3, the "ENERGY/POWER" field in the database header gives the calibrated system gain all the way from the microphone input to the sampled data file itself, expressed as dB re one squared A/D count per sample per 400 squared μPa. (400 = 20^2.) This number was obtained for each tape by digitizing the tape's 95.5dB calibration signal and observing the level of the sampled data. To compute the true noise power of a segment of sampled data, one can use the ILS CST command to compute the level of sampled data in dB re one squared A/D count (displayed by CST as "DB LEVEL") and then subtract the system gain given in the header, yielding the (band-limited) noise power at the microphone in standard SPL units, dB re 20 μPa. For example, for the file F15HT05.FLT, the CST command shows that the entire file has a power of 43.8 dB (actually displayed as 43.849). The header gives the system gain as -62.8 dB. Therefore the noise power in the 0-4-kHz band is 43.8 - (-62.8) = 106.6 dB in SPL units.

Both the in-mask microphone and the helmet microphone were used, as shown in Table 2.3. Separate calibrations allow a direct comparison between recordings made with the two microphones.

Digitized files were prepared from both tapes 10-1 and 10-2, recorded in an F-15A. However, in the PCM tape made from tape 10-2, the waveform shows symptoms of analog tape saturation. Since the PCM tape was made by from the original analog tape recorded in the field, we conclude that the original tape 10-2 was recorded at too high a level and saturation of that tape occurred. Therefore the F-15A files from tape 10-2 were not used in this study.

2.4   Oxygen Mask Effects

When a communicator wears a helmet/oxygen mask combination, the mask has a strong effect on the acoustic noise appearing at a vocoder's input. There are four effects:

1. Some additional noise is introduced by the rush of air through the valves of the respiration system.

2. The speech itself is distorted by the mask as a resonant chamber, compared to what the speech would be if the talker's vocal tract were terminated in a larger enclosure.

3. The communications microphone normally used in an oxygen mask [19] is of a noise-cancelling type, attenuating noise at low frequencies.

4. The mask attenuates noise originating outside it.

File name	Tape	Sampling	In-Band SPL (dB)	Comments
A10H09.16K	8-2	16 kHz		On helmet
A10H09.FLT	8-2	8 kHz	104	On helmet
A10M08.16K	8-1	16 kHz		In mask
A10M08.FLT	8-1	8 kHz	104	In mask
A10H25.16K	8-2	16 kHz		On helmet
A10H25.FLT	8-2	8 kHz	107	On helmet
A10M24.16K	8-1	16 kHz		In mask
A10M24.FLT	8-1	8 kHz	103	In mask
F15HT05.16K	5-2	16 kHz		On helmet
F15HT05.FLT	5-2	8 kHz	107	On helmet
F15MT05.16K	5-1	16 kHz		In mask
F15MT05.FLT	5-1	8 kHz	99	In mask
F15MT10.16K	10-1	16 kHz		In mask
F15MT10.FLT	10-1	8 kHz	108	In mask
F16H06.16K	6-2	16 kHz		On helmet
F16H06.FLT	6-2	8 kHz	110	On helmet
F16M06.16K	6-1	16 kHz		In mask
F16M06.FLT	6-1	8 kHz	101	In mask
F4EH11.16K	11-2	16 kHz		On helmet
F4EH11.FLT	11-2	8 kHz	104	On helmet
F4EM11.16K	11-1	16 kHz		In mask
F4EM11.FLT	11-1	8 kHz	98	In mask
F4EHL.16K	11-2	16 kHz		On helmet
F4EHL.FLT	11-2	8 kHz	108	On helmet
Tornado Recordings, All Outside Masks				
TOR13.FLT	TOR:13	8 kHz	106	Pilot
TOR21.FLT	TOR:21	8 kHz	102	Pilot
TOR22.FLT	TOR:22	8 kHz	99	Navigator
TOR34.FLT	TOR:34	8 kHz	115	Pilot

Table 2.3: Acoustic noise files digitized 1989-90

(In [33], Singer studied the speech distortion effect and concluded that the effect did not interfere seriously with the use of narrowband (LPC-10) vocoders.)

In figures 2.2 through 2.5 we present comparisons quantifying the mask attenuation effect. These figures show the difference between noise power spectra estimated from recordings made inside and outside the oxygen masks in four aircraft. In all four cases, the pilot was holding his breath. These are not truly measurements of the mask's attenuation of outside noise, because some noise originates inside the mask/respirator system. However, the comparisons are suggestive. The mask does not provide much attenuation at low frequencies, but seems to provide 15-30 dB of attenuation at most frequencies above 800 Hz. Dips in the apparent attenuation at some higher frequencies may be due to noise originating inside the mask, or to resonances in the cavity inside the mask, or to both.

It should be noted that the mask's exhaust valve, which would be expected to be closed while the pilot holds his breath, opens when the pilot speaks. Thus we would expect the mask to attenuate outside noise less efficiently during actual speech.

The difference between broadband noise levels inside and outside the mask is only about 8 dB, because much noise power is concentrated in lower frequencies where the mask's attenuation is less effective. But it should be noted that the measurement microphones used here were instrumentation microphones, whereas these masks are equipped with noise-cancelling communications microphones which attenuate far-field noise below about 1000 Hz.
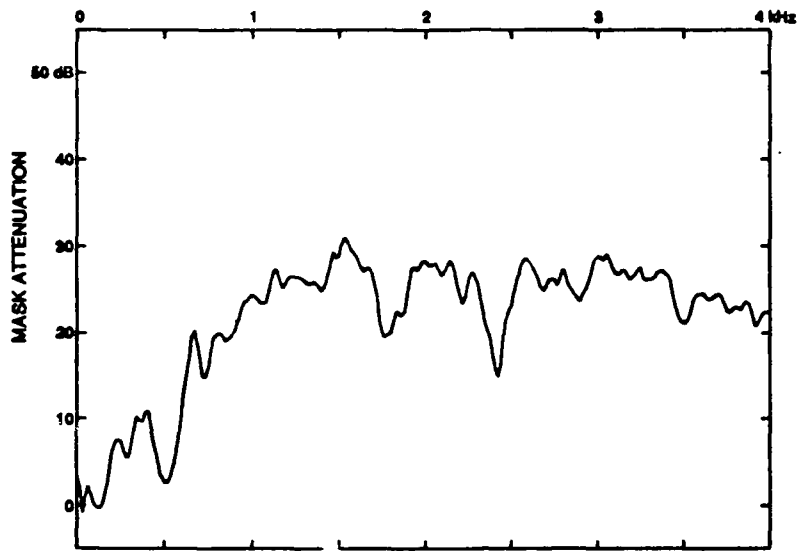
Figure 2.2: Calculated "mask attenuation," F-15A (F15HT5,F15MT5)
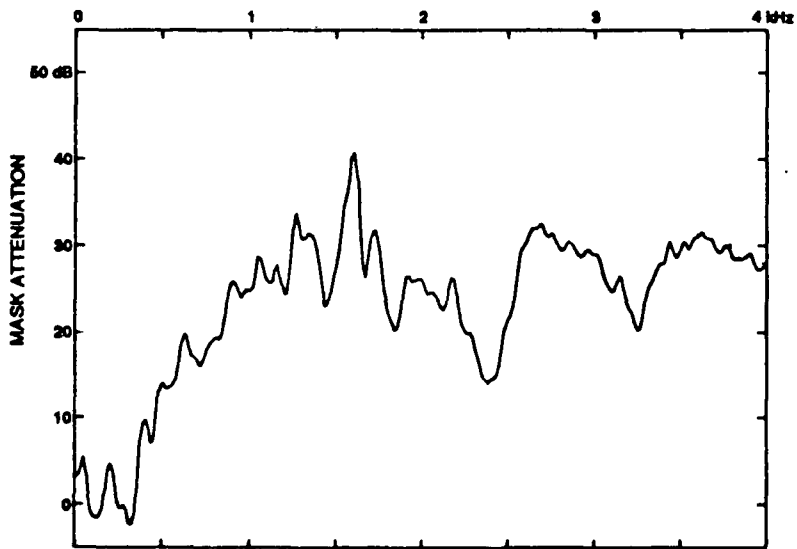
Figure 2.3: Calculated "mask attenuation," F-16A (F16H06,F16M06)
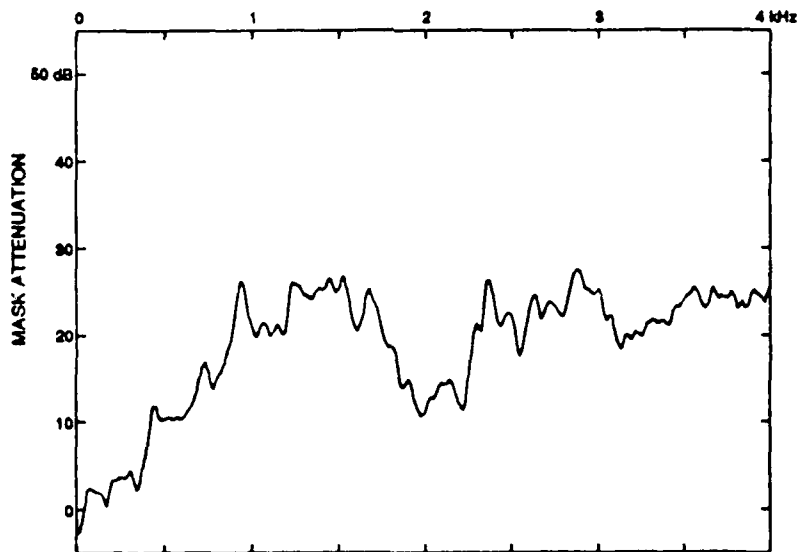
Figure 2.4: Calculated "mask attenuation," F-4E (F4EH11,F4EM11)
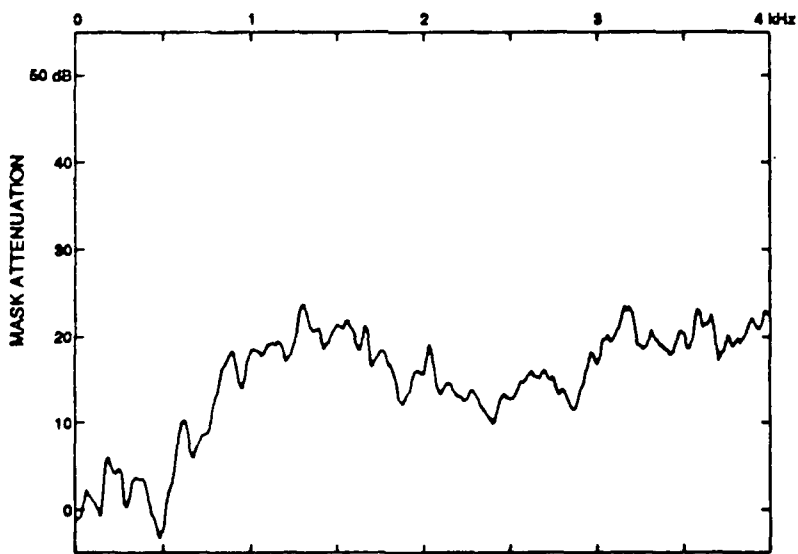
Figure 2.5: Calculated "mask attenuation," A-10 (A10H24,A10M25)

Chapter 3

Spectrum Estimation Methods

In order to detect and characterize rapid changes in noise statistics, analysis based on relatively short time segments is required. A typical segment of data might consist of 20-50 ms of recorded noise sampled at 8 kHz, providing 160-400 samples for digital processing. A spectral analysis of such segments provides *the basic information with which we will evaluate time varying noise properties.* This chapter discusses issues related to the choice of power spectrum estimators (PSE's).

3.1   Random Process Notation and Definitions

For completeness, we define in this section some basic discrete-time random process terminology that will be used in this report.

For our purposes, a *discrete-time random process* is an infinite sequence of real or complex random variables $\{x_n\}$ where n ranges over all integers.

1. The process $\{x_n\}$ is *stationary* if, for every finite set of integers $n_1, \cdots, n_k$ and every integer l, the joint distribution of the random variables $x_{n_1}, \cdots, x_{n_k}$ is identical to the joint distribution of $x_{n_1+l}, \cdots, x_{n_k+l}$.

2. The process $\{x_n\}$ is *Gaussian* if, for every finite set of integers $n_1, \cdots, n_k$, the random variables $x_{n_1}, \cdots, x_{n_k}$ have a multivariate Gaussian joint distribution.

3. The process $\{x_n\}$ is *white* if the random variables x_n are all independent.

4. *Frequency:* Throughout this report, the random processes will represent data sampled at a rate of F_s samples per second, and we can express angular frequency ω (in radians per second) in terms of frequency F (in Hz) or normalized

dimensionless frequency $f = F/F_s$,

$$\omega = 2\pi f = 2\pi F/F_s.$$

In this report, F_s will always be either 8000 Hz or 16000 Hz.

5. *The power spectral density (PSD):* For a stationary random process, the expectation $\mathcal{E}(\bar{x}_{n-k} x_n)$ (if it exists) is independent of n, and its value is denoted r_k. The complex sequence $\{r_k\}$ is called the autocorrelation function of $\{x_n\}$, and if $\{r_k\}$ has a Fourier transform

$$P_f = \sum_{k=-\infty}^{\infty} r_k \exp(-2\pi jkf), \quad -1/2 \leq f \leq 1/2, \tag{3.1}$$

then P_f is the *power spectral density* of $\{x_n\}$. (Note that P_f is only defined when $\{x_n\}$ is stationary.)

6. *The short-time Fourier transform (STFT):* Given a window function of length N, $\{w_k, k = 0, \cdots, N - 1\}$, the short-time Fourier transform $X_f(t)$ is defined as

$$X_f(t) = \sum_{k=0}^{N-1} w_k x_{t+k} \exp(-2\pi jkf), \quad -1/2 \leq f \leq 1/2. \tag{3.2}$$

3.2   Power Spectrum Estimators

In choosing a PSE method, the principal concerns are:

1. Resolution—the ability of the method to show distinct features of the PSE at neighboring frequencies, e.g. separating two peaks.

2. Accuracy—the ability of the method to produce estimates close to the theoretical power spectrum although based on a limited sample of data.

3. Sensitivity—the vulnerability of the estimate to a misjudgment of the character of the underlying random process.

In addition to these main characteristics, individual PSE techniques may suffer from such quirks as a tendency to produce split peaks where only one exists.

The available PSE algorithms always show a tradeoff among the three basic factors. For example, within the classical technique of smoothed periodogram analysis one may continuously trade off between resolution and accuracy by adjusting the smoothing window. Some techniques such as "maximum entropy" spectral estimation, which assumes an underlying autoregressive (AR) process model, achieve improved resolution and accuracy at the expense of sensitivity to the correctness of the model assumptions. Essentially then, a PSE method selection and adjustment can only achieve an appropriate balance of quality factors for the intended application.

The various PSE techniques can be organized in three broad categories: the "classical" methods based on Fourier analysis, those based on determining a filter model which might have been used to generate the data, and finally some special techniques which will not be discussed here, such as Capon's "maximum likelihood estimator." (The latter is based on adapting a data filter at each frequency so as to pass the frequency in question, but minimize the response to all other frequencies. This method, like most of the others, relies on an evaluation of the sample autocorrelation, but uses it differently.)

3.2.1 Model-Based Methods

The autoregressive (AR), moving average (MA) and combined autoregressive, moving average (ARMA) methods of PSE analysis are based upon the notion that the noise record x_n has been generated by passing a white noise sequence u_n through a constant coefficient linear filter. The most general ARMA generator is

$$x_n = \sum_{k=0}^{q} b_k u_{n-k} - \sum_{i=k}^{p} a_k x_{n-k}, \tag{3.3}$$

where $b_0 = 1$. For an AR process the b_k coefficients are assumed zero for $i > 0$, and for an MA process the a_k coefficients are zero. Now if given the x_n one can estimate the a_k, the b_k, and $\text{Var}(u)$ for this model, a PSE can be written down immediately as

$$\hat{P}(\omega) = \text{Var}(u) \left| \frac{\sum_{k=0}^{q} b_k \exp(-j\omega k)}{1 + \sum_{k=1}^{p} a_k \exp(-j\omega k)} \right|^2. \tag{3.4}$$

In terms of normalized frequency, then,

$$\hat{P}_f = \text{Var}(u) \left| \frac{\sum_{k=0}^{q} b_k \exp(-2\pi j k f)}{1 + \sum_{k=1}^{p} a_k \exp(-2\pi j k f)} \right|^2. \tag{3.5}$$

The AR case is by far the easiest to handle and a considerable number of methods have been devised to handle it. One satisfactory method is the modified covariance method. It is based on a particular estimate of the sample autocorrelation function of the observed x_n followed by the solution of a set of linear equations to obtain the a_k coefficients of the model, as in Linear Predictive Coding.

The MA model is more difficult to work with since an analogous development leads in this case to a set of nonlinear equations. One satisfactory approach is provided by Durbin's method, which first fits the process to a high order (large p) AR model. Then as a second step, a set of b_k for an MA model (small q) is determined which approximates the AR model. This second step also leads to linear equations so that the computational problem becomes tractable.

Treating the general ARMA model is even more complex. The full ARMA model methods provide the best possible PSE's when the model truly fits the data, but are sensitive to this assumption and are more computationally demanding and cranky than other procedures.

For AR, MA, and ARMA models, the orders p and q must be either chosen by the experimenter (on the basis of hypotheses about the process being estimated) or else computed from the data. Automatic selection of p and q is especially complex in the case of ARMA models, which require estimates of both p and q. Algorithms for estimating these parameters are the subject of continuing research.

3.2.2   Classical Methods

Classical Fourier methods make minimal assumptions about the underlying noise process and consequently produce only modest resolution and accuracy for short data records. They are particularly ᵤₐₛy to implement using FFT algorithms. *If the noise source is only locally stationary, the accuracy and/or resolution cannot be improved by utilizing an average of the PSE's over many data records (long term averaging).*

The Fourier methods are based either on the Blackman-Tukey method, which proceeds from autocorrelation estimates, or on the periodogram estimator, which proceeds from the discrete Fourier transform of the noise signal itself. Periodogram-based methods are especially significant because they are directly related to spectral restoration methods for noise removal. For both the Blackman-Tukey estimator and the periodogram estimator, the principal drawback is that we must choose either modest frequency resolution or large estimate variance for short records. In addition, there may be a problem with bias (frequency sidelobes) when a large amount of power is concentrated in narrow bands.

When a discrete random process $\{x_n\}$ is regarded as sampled data with a sampling rate of F_s, we will use a normalized frequency variable $f = F/F_s$, where F is frequency in Hz. Then the basic periodogram power spectrum estimate for $\{x_n\}$ is defined by

$$\hat{P}_f(t) = \frac{1}{N} \left| \sum_{k=0}^{N-1} w_k x_{t+k} \exp(-2\pi j f k) \right|^2, \tag{3.6}$$

where t is the starting time of the sample "frame" being used, and w_k ($k = 0, \cdots, N-1$) is a window function of length N. In terms of Section 3.1, $\hat{P}_f(t)$ is $1/N$ times the squared magnitude of the short-time Fourier transform with the same window function:

$$\hat{P}_f(t) = \frac{1}{N} |X_f(t)|^2. \tag{3.7}$$

In Chapter 4 we will make use of the fact that the statistical distribution of $|X_f(t)|^2$ is known, at least approximately, for certain choices of frequency spacing and certain classes of random processes.

3.2.3 Windows

The window function w_k $(k = 0, \cdots, N-1)$, referred to in the previous section, affects the bias and the apparent resolution of the periodogram PSE. Although many different data windows are in common use [16], we will restrict ourselves to discussing three types of windows:

1. Rectangular window, $w_k \equiv 1$.

2. Hamming window, $w_k = 0.54 - 0.46\cos(\frac{2\pi k}{N-1})$.

3. Trapezoidal window

$$w_k = \begin{cases} (k+1)/(R+1) & 0 \le k < R \\ 1 & R \le k < N - R \\ (N-k)/(R+1) & N - R \le k < N \end{cases}$$

Windows such as the Hamming window are commonly applied when estimating spectra that are suspected of having moderately spaced sinusoidal features. Compared with the rectangular window, the Hamming window offers lower "sidelobe" levels. This means that, with a Hamming window, a sinusoidal component of the noise will cause less bias at frequencies far from its own frequency. Figures 3.1 and 3.2 show the frequency sidelobes of a pure sinusoid with rectangular and Hamming windows, respectively.

In practice, a trapezoidal window is often used for spectral restoration as discussed in Section 4.3.1. The frequency characteristics of a trapezoidal window depend on R, the length of its "ramp". For a ramp size of 76 and a window length of 256 (as used by Kang and Fransen in [20]), the frequency sidelobes of a pure sinusoid are as plotted in Figure 3.3.

3.2.4 Prewhitening

One widely recommended PSE method employs a prewhitening filter to the noise before conducting the periodogram PSE analysis. This prewhitening filter may be determined as in the autoregressive (AR) model procedures for PSE. Then the final PSE becomes the product of the PSE's found from the AR model and from periodogram analysis of the residual whitened noise. This estimate should have improved accuracy because the periodogram is unbiased for white noise. The prewhitening technique is intended to mitigate the biases often encountered when a spectral estimator is applied to noise whose power spectrum has pronounced peaks or valleys.
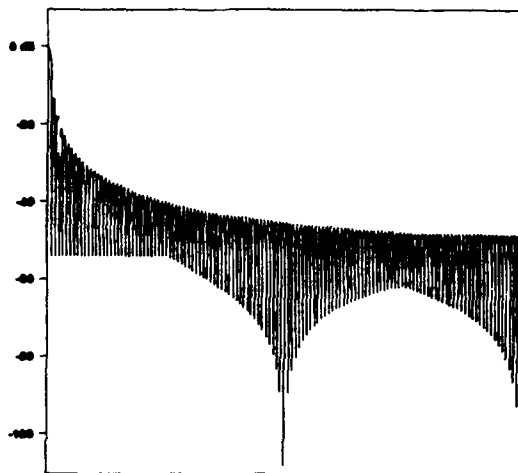
Figure 3.1: Frequency sidelobes of rectangular window
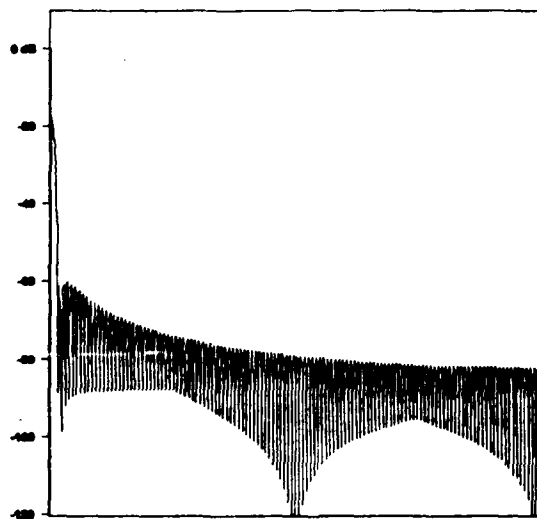
Figure 3.2: Frequency sidelobes of Hamming window
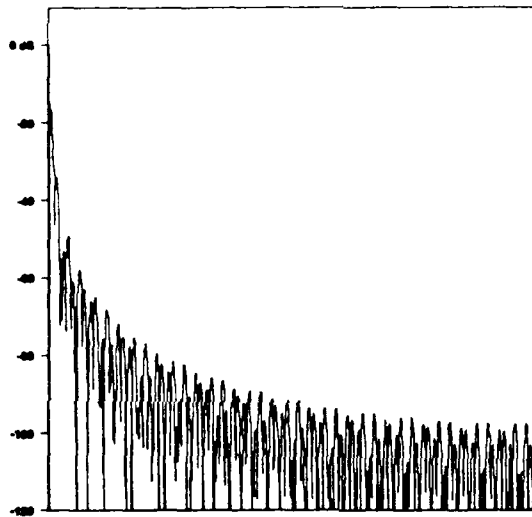
Figure 3.3: Frequency sidelobes of trapezoidal window (N=256, R=76)

The procedure is

1. apply a filter to make the original signal more spectrally flat;

2. estimate the power spectrum of the whitened signal;

3. compensate for the prewhitening filter by dividing by its power transfer function.

The resulting spectrum estimate is not critically dependent on the exact parameters of the prewhitening filter, since the filter is compensated for in the third step.

In the second step, any spectrum estimation method could be used. If the periodogram method is chosen, then the data window will have an effect on the estimates obtained.

Intuitively speaking, the prewhitening method should be most helpful in situations where a large amount of energy at certain frequencies causes biases (sidelobes) at other frequencies. In order to evaluate the effectiveness of prewhitening, we have analyzed noise samples from several aircraft by six methods:

1. Non-prewhitened averaged periodogram method, rectangular window;

2. Prewhitened averaged periodogram method, rectangular window;

3. Non-prewhitened averaged periodogram method, Hamming window;

4. Prewhitened averaged periodogram method, Hamming window.

5. Non-prewhitened averaged periodogram method, trapezoidal window;

6. Prewhitened averaged periodogram method, trapezoidal window.

Figures 3.4, 3.5 and 3.6 compare prewhitened and non-prewhitened estimates for noise recorded inside an oxygen mask in an F-15A aircraft. The prewhitened estimate was obtained using an 8-pole prewhitening filter. The heavier curve is the prewhitened estimate, and the lighter one is the estimate obtained without prewhitening. Figure 3.4 compares the two methods using a rectangular window ($N = 256$). Prewhitening has a strong effect on the spectrum estimate in this case, because of the sidelobes of the low-frequency peaks.

On the other hand, Figure 3.5 compares the two methods using a Hamming window ($N = 256$). The effect of prewhitening is much less pronounced in Hamming-windowed estimates, because the window lowers those sidelobes. Figure 3.6 makes the same comparison for a trapezoidal window ($N = 256, R = 76$), with similar results: estimates obtained without prewhitening are within 1 dB of the estimates obtained with prewhitening.

The preceding comparisons for actual aircraft noise show little advantage for prewhitening, provided that a Hamming or trapezoidal window is used. On the other hand, Figures 3.7, 3.8, and 3.9 present similar comparisons for *synthetic* noise that is the sum of a 70.5 Hz sinusoid sampled at 4 kHz and white gaussian noise. Because of the large discontinuity in the spectrum at 70.5 Hz we would expect a bias problem in the periodogram estimates. The difference is again more pronounced in the case of the rectangular window, but *for this synthetic noise* the difference is still substantial at some frequencies, even for the other windows.

Because prewhitening has little effect on estimates obtained for real aircraft noise with Hamming or trapezoidal windows, our conclusion is that prewhitening is not advantageous for our purposes, except in a situation where a rectangular window is being used.

3.3   Choice of Method

The periodogram is based on the short-time Fourier transform, which is also central to spectral restoration, the only widely used general method of single-microphone noise removal. Therefore, any information we can obtain about the time variation of periodogram estimates has direct implications for the

performance of spectral restoration. For this reason we have chosen to use a periodogram-based PSE as our short-term spectral estimator. For the time scales we are interested in, the periodogram provides adequate frequency resolution. We have chosen to use Hamming and trapezoidal windows for our estimates, again partly because of their use in spectral restoration. Finally, we have chosen not to use prewhitening because its effect is marginal for real aircraft noise with these windows.
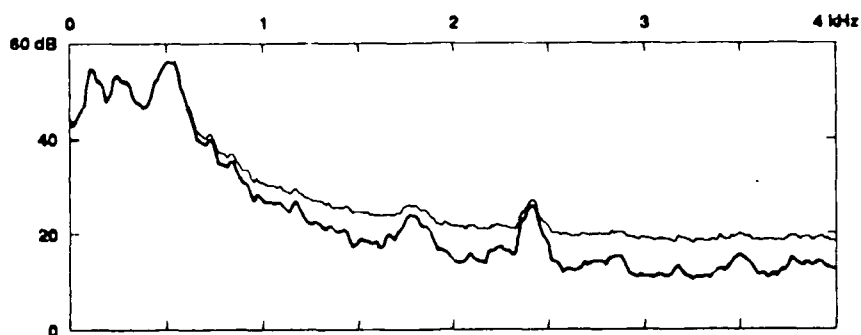
Figure 3.4: Prewhitening effect: Aircraft noise, rectangular window, PSE with and without prewhitening (F15C59)
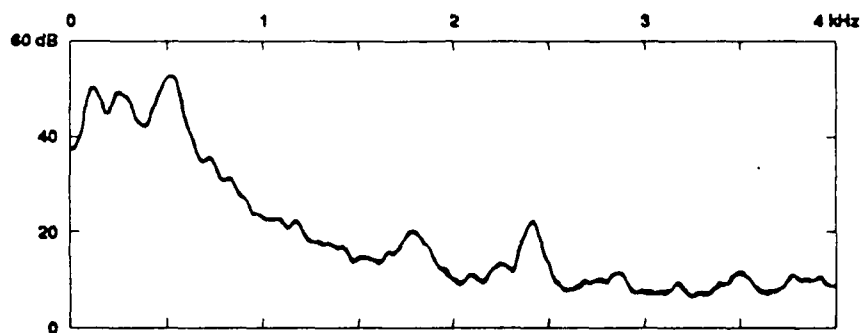
Figure 3.5: Prewhitening effect: Aircraft noise, Hamming window, PSE with and without prewhitening (F15C59)

Figure 3.6: Prewhitening effect: Aircraft noise, trapezoidal window, PSE with and without prewhitening (F15C59)
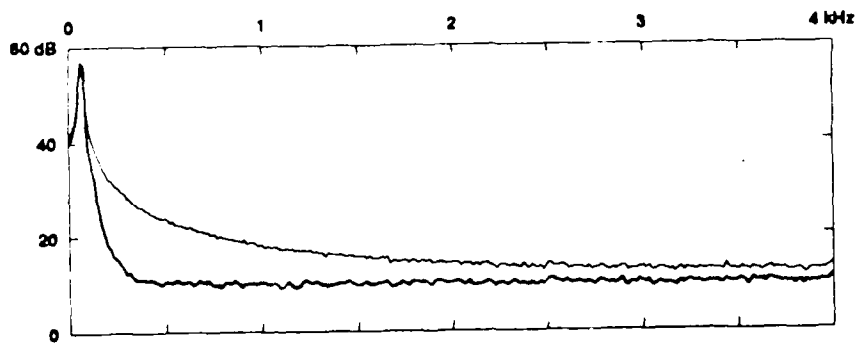
Figure 3.7: Prewhitening effect: Synthetic signal, rectangular window, PSE with and without prewhitening (SINGAU)
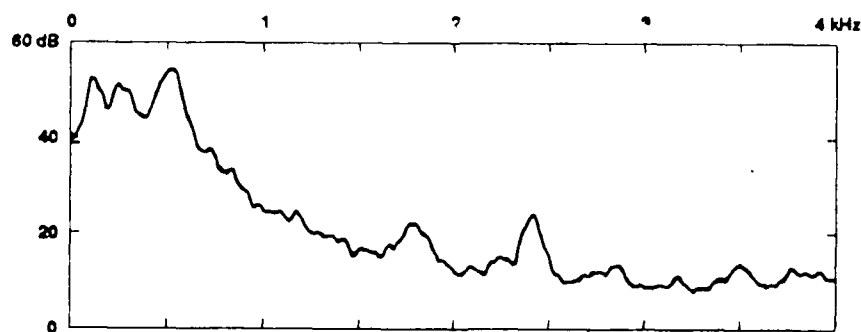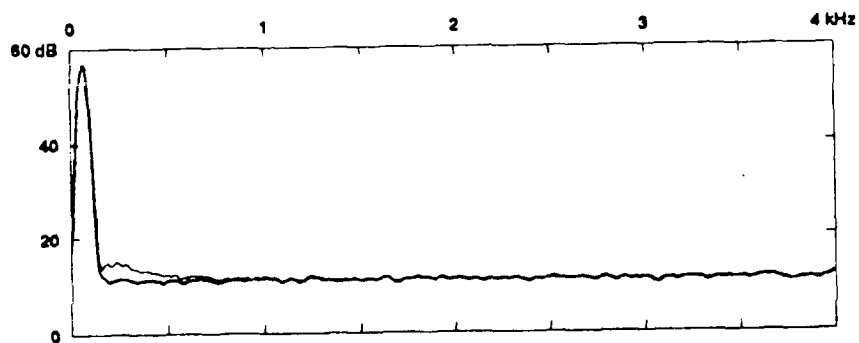

Figure 3.8: Prewhitening effect: Synthetic signal, Hamming window, PSE with and without prewhitening (SINGAU)
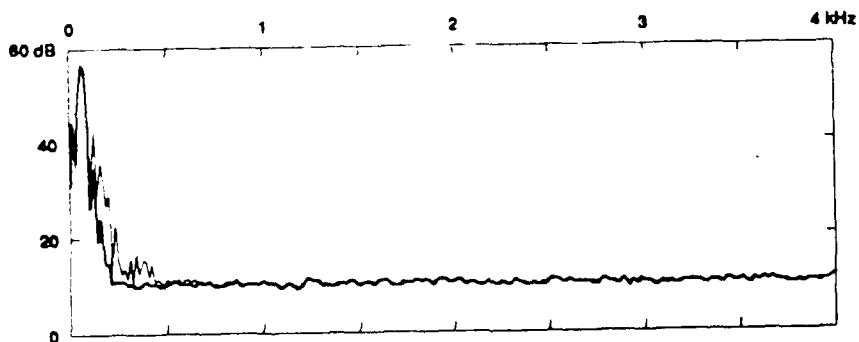
Figure 3.9: Prewhitening effect: Synthetic signal, trapezoidal window, PSE with and without prewhitening (SINGAU)

Chapter 4

Variation of Noise

The output of a random noise source, whether it is stationary or nonstationary, will exhibit variation from one sample (or one analysis frame) to another. In this chapter we compare known time-varying properties of stationary random processes with characteristics of aircraft noise environments. We devote special attention to the relationship between noise variation and spectral restoration.

4.1 Stationary Noise and the Short-Time Fourier Transform

If $\{x_n\}$ is a stationary random process, then for given values of f and t, $X_f(t)$ is a random variable. When $\{x_n\}$ is a white Gaussian process, we can characterize the distribution of $X_f(t)$, at least at some frequencies f [3, 18, 22]. If a *rectangular* window is used for the STFT and if f is any multiple of $1/(2N)$ except for 0 or $\pm 1/2$, then $|X_f(t)|^2/(P_f/2)$ has the χ^2 distribution with two degrees of freedom.[1] (Equivalently, $|X_f(t)|/\sqrt{P_f/2}$ has the Rayleigh distribution.) From the known probability density function of χ_2^2, it follows that in these cases the density function of $v = |X_f(t)|^2$ is $(1/P_f)\exp(-v/P_f)$ for $v \geq 0$:

$$\Pr(|X_f(t)|^2 < u) = \int_0^u (1/P_f)\exp(-v/P_f)dv. \qquad (4.1)$$

In other words, the measured spectral magnitudes squared

$$|X_f(0)|^2, \quad |X_f(M)|^2, \quad |X_f(2M)|^2, \quad \cdots$$

for frames spaced M samples apart will be a stochastic process whose individual variables will have the χ_2^2 distribution, up to the scale factor $P_f/2$. Moreover,

---

[1] Or one degree of freedom, if f is 0 or $\pm 1/2$.

because the noise is assumed white, the successive magnitudes are independent if frames do not overlap.

To a certain extent, these properties of $|X_f(t)|^2$ can be extended to non-white or non-Gaussian stationary processes, to intermediate frequencies, and to other windows. If a rectangular window is used, then $|X_f(t)|^2/(P_f/2)$ will still [3, 18] be approximately distributed as χ_2^2, provided that the frame length N is large enough, that the spectral density P_f is a smooth enough function of f, and that f is not close to 0 or $\pm 1/2$. (It should be noted that this condition specifically excludes noise sources with a strong sinusoidal component.) For other commonly used windows, the distribution of $|X_f(t)|^2/(P_f/2)$ can be approximated [18] by χ_ν^2 for some larger number of degrees of freedom ν; in the case of a Hamming window, $\nu = 4$.

4.2   Quantitative Measures of Variation

4.2.1   Standard-deviation-to-mean ratio

One measure of the variation of $|X_f(tM)|$ is the RMS deviation of $|X_f(tM)|$ from its long-term time average. Taking a five-second record of aircraft noise, at each frequency f we treat the successive magnitude estimates $|X_f(tM)|$, $t = 0, 1, \cdots$ as samples from a common distribution and compute the sample standard deviation and mean. Since we expect the sample standard deviation to be directly proportional to the mean, we normalize the quantity obtained by dividing the sample standard deviation by the sample mean.

In the case where $|X_f(tM)|^2/(P_f/2)$ has the χ_2^2 distribution (as for Gaussian white noise), it can be calculated from Equation (4.1) that the ratio of sample standard deviation to sample mean has an expected value of $\sqrt{4/\pi - 1}$, or approximately 0.52. In the case of a strong sinusoid, the ratio is close to 0 because the succesive magnitudes $|X_f(tM)|$ are nearly the same.

Figures 4.1, 4.2, and 4.3 show the standard-deviation-to-mean ratio of spectral magnitude estimates, as a function of frequency, for three aircraft noise records in the RADC/EEV acoustic noise data base.

4.2.2   Residual RMS Error

Another simple measure of variation is motivated specifically by the spectral-restoration application. Instead of the RMS deviation of $|X_f|$ from its long-term mean, we can examine the RMS deviation of $|X_f|$ from short term estimators of $|X_f|$. The specific short-term estimator we chose is simply a moving average over L frames, where L is variable. Thus we are concerned with the RMS
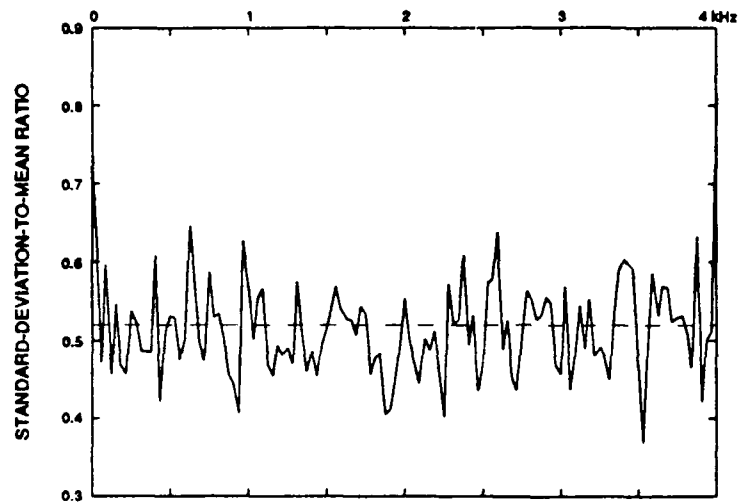
Figure 4.1: Ratio of standard deviation to mean, E-4B Battle Staff area (**E4BBS**)
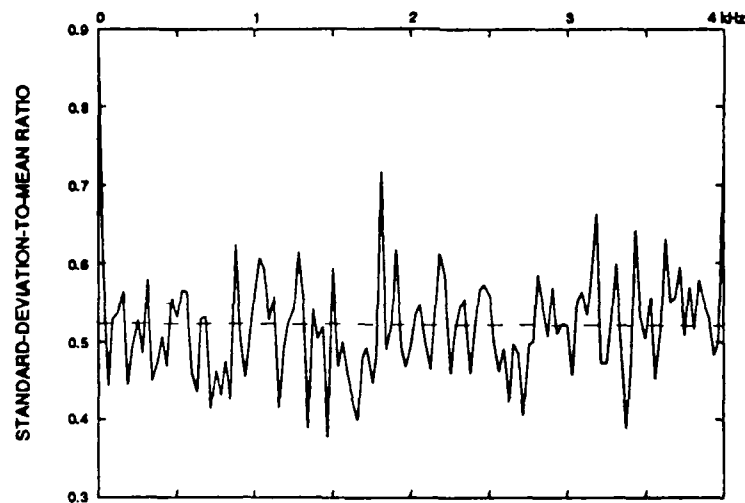
Figure 4.2: Ratio of standard deviation to mean, F-15A outside mask, 1.2 Mach (**F15C33**)
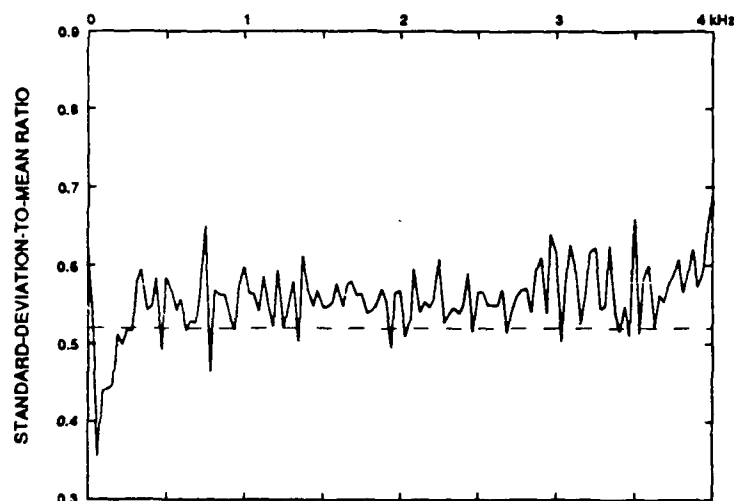
Figure 4.3: Ratio of standard deviation to mean, EC-130 ABCCC radio operator (**EC130A**)

deviation over a T-frame interval

$$E_f^L = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(|X_f(tM)| - Q_f^L(tM))^2} \qquad (4.2)$$

where $Q_f^L(t)$ is the moving average of the L frames preceding frame t,

$$Q_f^L(t) = \frac{1}{L}\sum_{u=1}^{L}|X_f((t-u)M)|. \qquad (4.3)$$

The quantity E_f^L can be described as the residual RMS error of Q_f^L considered as an estimator of $|X_f|$.

Taking the same five-second records of noise used in Figures 4.1, 4.2, and 4.3, we have computed DFT magnitudes of successive frames ($M = 256$). At each frequency, we have normalized the residual RMS error by dividing by the long-term mean magnitude of $|X_f|$. Figures 4.4, 4.5, and 4.6 show the normalized residual RMS error of spectral magnitude estimates, as a function of frequency. Each plot shows curves for a number of different averaging periods (2, 3, $\cdots$, 20, and 40 frames).

4.2.3 Conclusion

Both standard-deviation-to-mean ratio and normalized residual RMS error are simple measures of variation of $X_f(tM)$ over time. Measures like these could
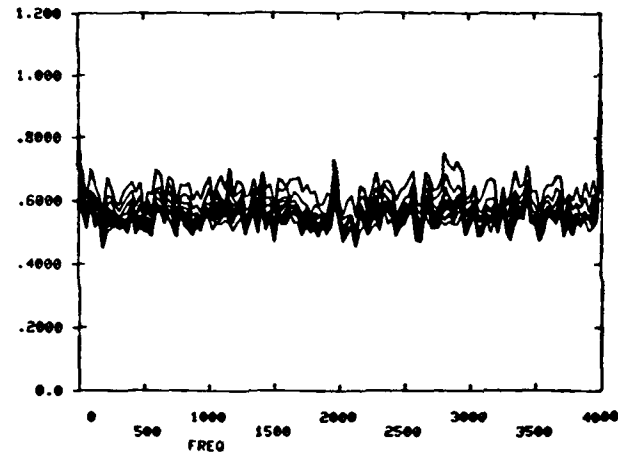
Figure 4.4: Residual prediction error, E-4B Battle Staff area (E4BBS)
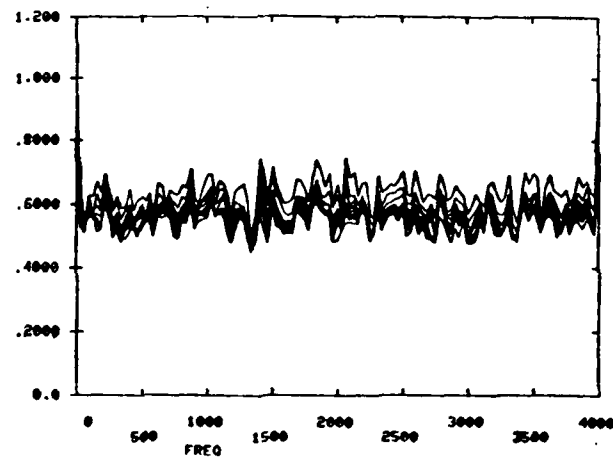
Figure 4.5: Residual prediction error, F-15A outside mask, 1.2 Mach (F15C33)
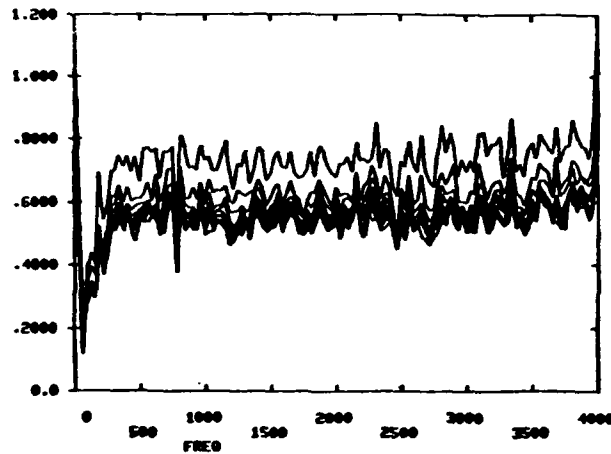
Figure 4.6: Residual prediction error, EC-130 ABCCC radio operator (**EC130A**)

be used in a practical algorithm that adapts noise-removal strategies to noise variation, frequency by frequency. However, as we have seen from the examples, both measures seem to have a large amount of sample variation that might makes them difficult to apply. For this reason, we go on in the next section to a more complicated measure of variation, based on the performance of spectral restoration algorithms.

4.3 Noise Variation and Spectral Restoration

4.3.1 Spectral Restoration Methods

Spectral restoration methods, often grouped under the name "spectral subtraction," attempt to recover a signal process $\{s_n\}$ corrupted by an additive noise process $\{d_n\}$ from the measured sum $\{x_n\} = \{s_n + d_n\}$, using information about the power spectrum of the noise process $\{d_n\}$. The input signal $\{x_n\}$ is processed in overlapped segments $\{x_0 \ldots x_{N-1}\}$ of length N.

To each segment, a fixed time-domain window $w_0 \ldots w_{N-1}$ is applied and a zero-filled FFT is used to evaluate the short-time Fourier transform of the windowed segment at a spacing[2] of $1/(2N)$, yielding a frequency-domain representation $X = \{X_{f_k}\} = \{X_{k/N}\}$, where k ranges over the $2N$ discrete values $-N, \cdots, N - 1$. Because the windowing and Fourier transform operations are

---

[2] This spacing is most common, the rationale being to accommodate multiplicative suppression rules; the issue of spacing is discussed in [29, 2, 15].

linear, $X = S + D$ where S and D are the corresponding (unknown) transforms of the signal and noise processes. For each f, an estimate $\hat{S}_f$ of S_f is formed, using the known X_f and an estimate $\hat{D}_f$ of the magnitude $|D_f|$. The rule for estimating S_f is called the *suppression rule*. Many popular suppression rules take the form

$$\hat{S}_f = \arg(X_f)(|X_f|^\mu - \beta \hat{D}_f^\mu)^{1/\mu} \tag{4.4}$$

or

$$\hat{S}_f = \arg(X_f)\max(c\hat{D}_f, (|X_f|^\mu - \beta \hat{D}_f^\mu)^{1/\mu}) \tag{4.5}$$

for appropriate values of β, μ, and c. The rule expressed in Equation (4.5) applies a "spectral floor" to prevent total suppression of the input signal, and has been found to reduce the "musical noise" effect otherwise found in the enhanced signal.

Once the estimates $\hat{S}_f$ are formed, an inverse FFT is used to form a signal estimate $\hat{s}_0 \hat{s}_1 \dots \hat{s}_{N-1}$. Signal estimates from the overlapping noisy segments are then added together [29, 2, 15] to produce the final enhanced signal estimate.

The designer of such a spectral restoration system has three choices to make:

1. The window: the data window is usually chosen so that, when overlapped copies of it are added together, the sum is a constant.

2. The noise estimator: how $|D_f|$ is estimated.

3. The suppression rule, including the noise floor c.

4.3.2   Limitations of Spectral Restoration Methods

Before we discuss the consequences of variation in $|D_f|$, let us examine the problem of estimating S_f from X_f when the magnitude of D_f is *known in advance*. The complex number X_f is the sum of S_f and D_f as in Figure 4.7. If we knew the magnitude of D_f and not its phase, then we would know only that S_f lay somewhere in the complex plane on the circle centered at X_f with radius $|D_f|$. As the figure shows, the magnitude of S_f could be anything from $|X_f| - |D_f|$ to $|X_f| + |D_f|$. There is a large uncertainty in estimating either the phase or the magnitude of S_f here, even in the "best case" situation where the magnitude of D_f is known.

4.3.3   Predictions, Simulations, and Actual Results

Noise variation of any kind will affect the performance of spectral restoration methods. Generally, the estimate of $|D_f|$ is updated during intervals judged to have little or no speech signal present, as determined by a speech-detection algorithm. Therefore the noise magnitude estimate applied to any single frame of noisy speech will have been estimated at some earlier time. With a nonstationary noise source, this magnitude estimate will be out of date. Variation in
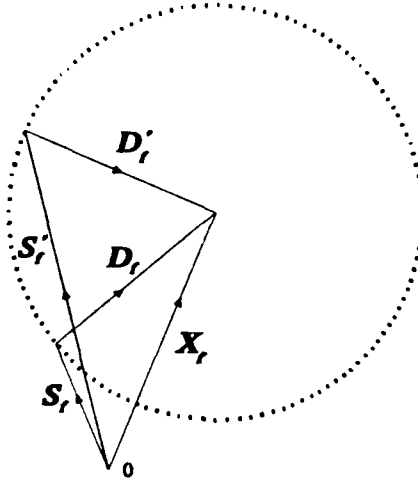
Figure 4.7: Effect of phase uncertainty in spectral restoration

the noise statistics should therefore be reflected in poorer performance of the spectral restoration procedure.

Gauging the performance of a spectral restoration algorithm for speech is not a straightforward matter. One could synthetically add a known reference signal to noise, apply the noise removal algorithm, and apply some distortion measure to compare the enhanced signal to the known reference. This approach is dependent on the specific reference signal. We have chosen to consider the effect of applying spectral restoration techniques to a noise-only signal, and measuring the ratio of the enhanced signal power spectrum to the input (noise-only) signal power spectrum. This ratio can be approximated by the ratio $|\hat{S}_f|^2/|X_f|^2$, with slight error due to the overlapping of frames; measurements with real signals have shown this error to be negligible.

This attenuation ratio should not be interpreted as a figure of merit for the suppression rule, since it does not take into account the distortion of any original signal. (For example, a large value of β with a small value of c leads to total suppression of any input, signal and noise alike.) Instead, we use the attenuation ratio to suggest the relative performance of a fixed suppression rule at different frequencies, and for different noise records.

For a white Gaussian noise source, we can apply the known distribution of Equation (4.1) to predict the ratio of output power $|\hat{S}_f|^2$ to input noise power $|X_f|^2$. Assuming the suppression rule of Equation (4.5),

$$|\hat{S}_f| = \max(c\hat{D}_f, \left(|X_f|^\mu - \beta\hat{D}_f^\mu\right)^{1/\mu}),$$

Rule	Predicted	Actual
$\mu = 2, \beta = 1$, c=1/12	4.3 dB	4.3 dB
$\mu = 1, \beta = 1$, c=1/12	10.3 dB	10.2 dB

Table 4.1: Predicted and actual noise attenuation (white Gaussian noise)

or in terms of squared magnitudes,

$$|\hat{S}_f|^2 = \max(c^2 \hat{D}_f^2, \left((|X_f|^2)^{\mu/2} - \beta \hat{D}_f^\mu \right)^{2/\mu}). \tag{4.6}$$

Then from Equations 4.1 and 4.6, the expected output power is

$$\mathcal{E}(|\hat{S}_f|^2) = \int_0^\tau c^2 \hat{D}_f^2 \exp(-v/P_f)/P_f \, dv$$

$$+ \int_\tau^\infty \left(v^{\mu/2} - \beta \hat{D}_f^\mu \right)^{2/\mu} \exp(-v/P_f)/P_f \, dv, \tag{4.7}$$

where τ is the threshold value of $|X_f|^2$ below which the spectral floor $c^2 \hat{D}_f^2$ is applied in Equation 4.6,

$$\tau = (\beta + c^\mu)^{2/\mu} \hat{D}_f^2.$$

Equation (4.7) can be used to calculate the attenuation to be expected when spectral restoration is applied to a noise-only input consisting of white Gaussian noise. Table 4.1 shows the results of evaluating (4.7) for representative values of μ, β, and c, with a rectangular window. For comparison and verification, this table also shows the results measured when a spectral restoration algorithm with the same parameters was actually applied to a 5-second sample of simulated white Gaussian noise. The same results would be expected for non-white Gaussian noise with any smooth enough spectral density, because the approximation (4.1) is still valid for such noise. If a non-rectangular window were used, the density (4.1) would have to be replaced by another density, such as the χ_4^2 density for a Hamming window.

Figures 4.8 through 4.20 show the noise attenuation obtained with a variety of noise-only inputs, using spectral restoration with the parameters

$$\mu = 2, \beta = 1, c = 1/12.$$

In other terminology, we have used the energy subtraction rule, with no over-subtraction, and a spectral floor 21 dB below the expected noise. These parameters are representative of those used in other studies [7, 20].

Each attenuation plot is accompanied by a plot of the estimated spectral density of the noise record used as input. These plots show behavior very
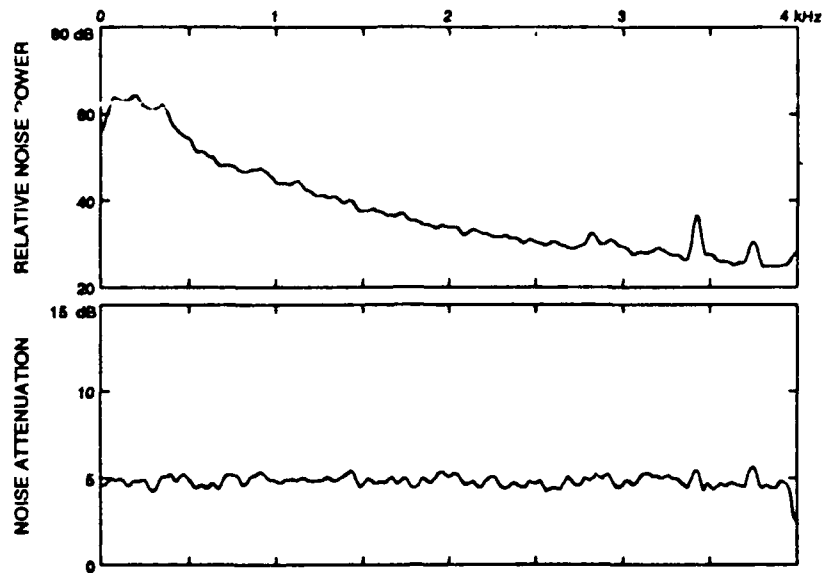
Figure 4.8: Noise attenuation, EC-135 Battle Staff (EC135B)

similar to the Gaussian case over most of the noise sources, but with much better performance in the vicinity of a strong and reliable sinusoid, as in Figures 4.12 and 4.10. For certain sinusoidal noise, as in Figure 4.20 near 1400 Hz and in Figure 4.19 near 3000 Hz, performance is near the Gaussian noise level; we will see in Chapter 6 that these noise sources show nonstationarity at these frequencies. Overall, the noise attenuation obtained with spectral restoration (as measured against noise-only inputs) is generally no worse for any of these noise records than for white Gaussian noise.
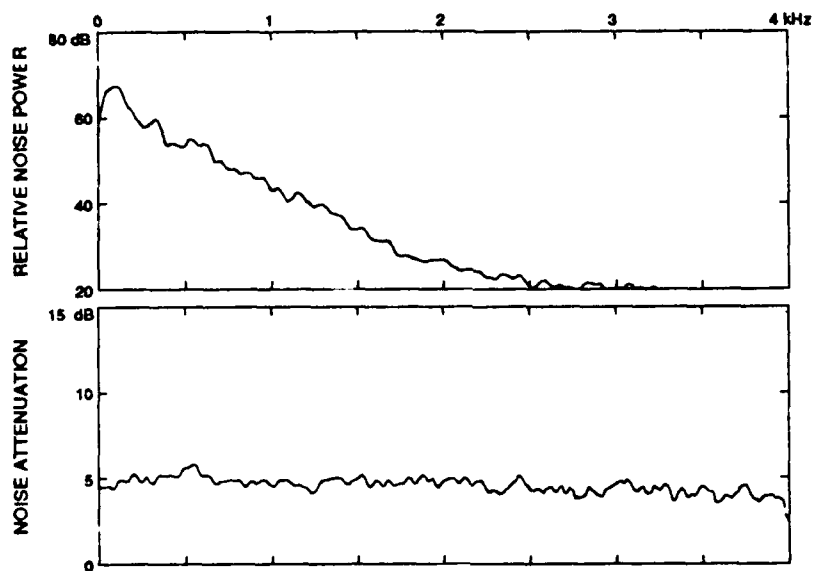
Figure 4.9: Noise attenuation, E-4B Battle Staff area **(E4BBS)**
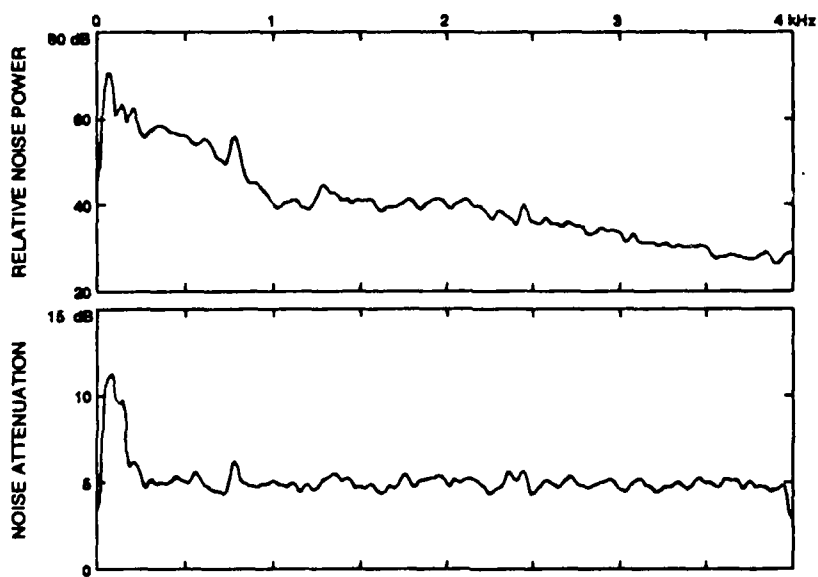
Figure 4.10: Noise attenuation, EC-130 ABCCC radio operator **(EC130A)**
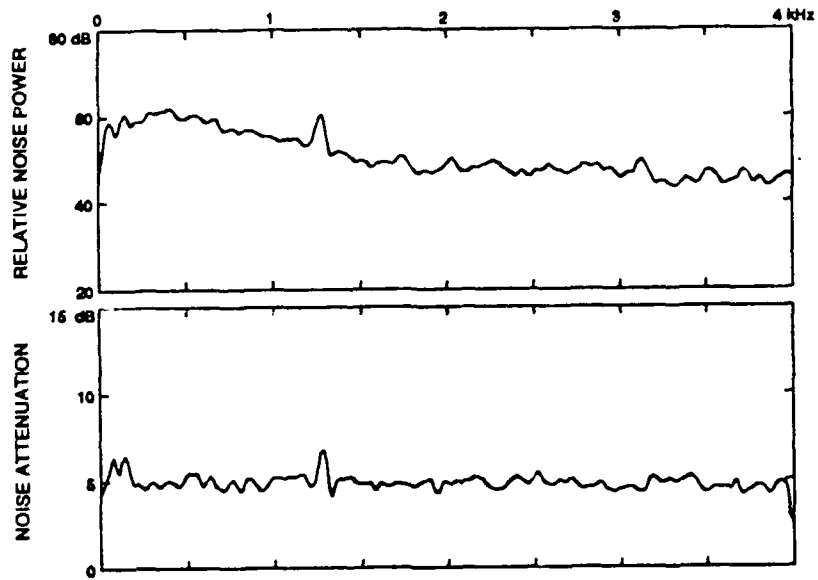
Figure 4.11: Noise attenuation, HC-130 (**HC130A**)
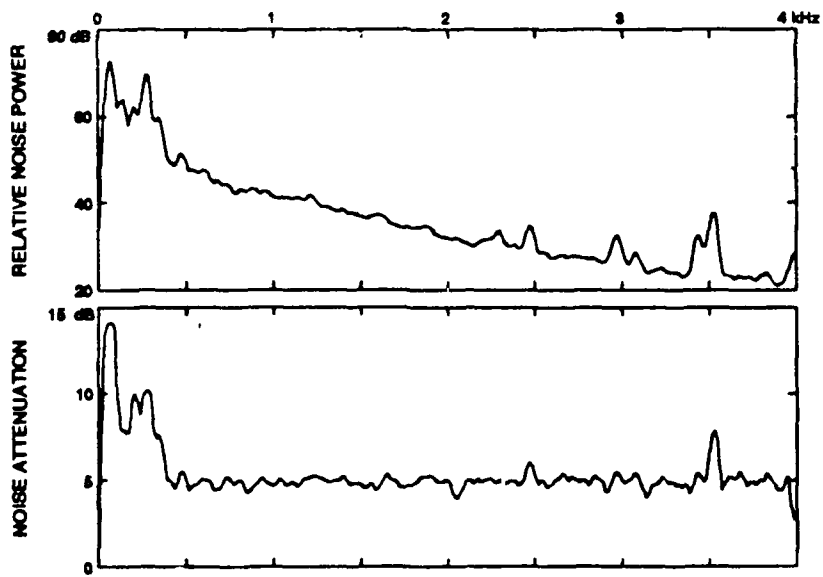
Figure 4.12: Noise attenuation, P-3C (**P3C**)
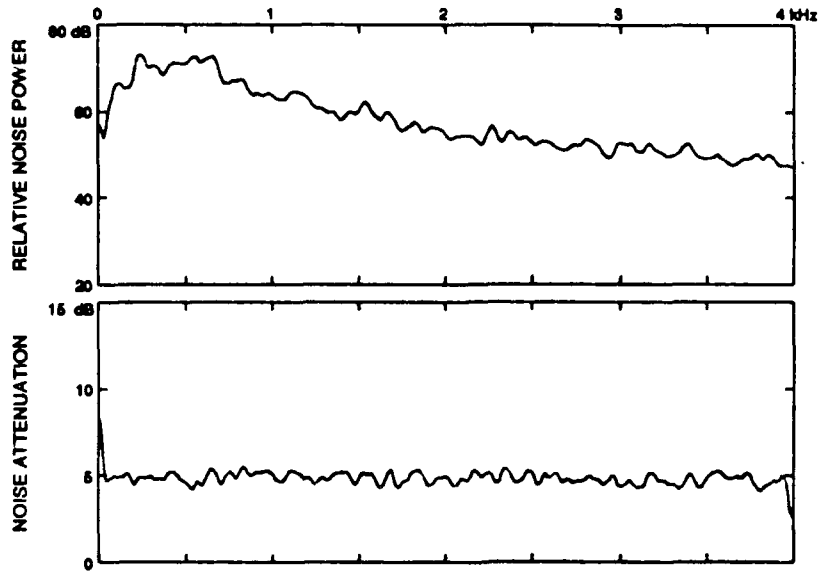
Figure 4.13: Noise attenuation, F-15A outside mask (**F15HT5**)
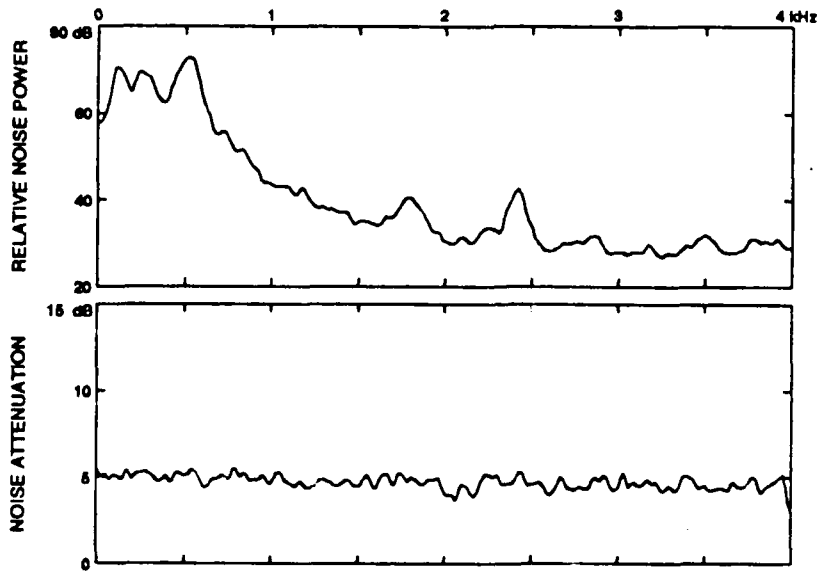
Figure 4.14: Noise attenuation, F-15A inside mask (**F15HT5**)
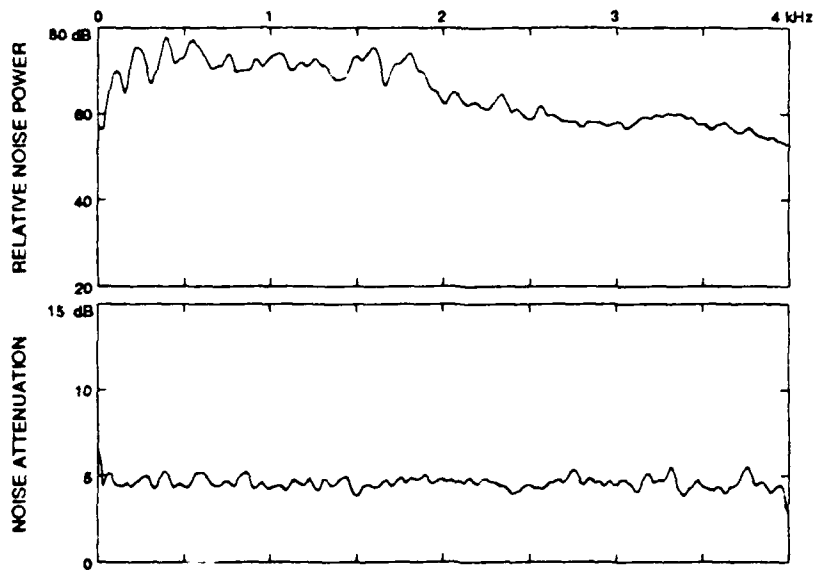
Figure 4.15: Noise attenuation, F-16A outside mask (**F16H06**)
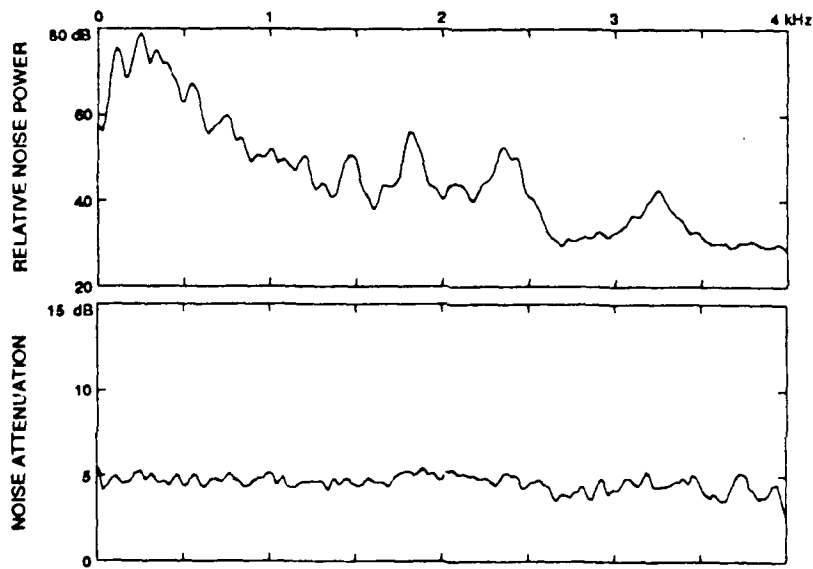
Figure 4.16: Noise attenuation, F-16A inside mask (**F16M06**)
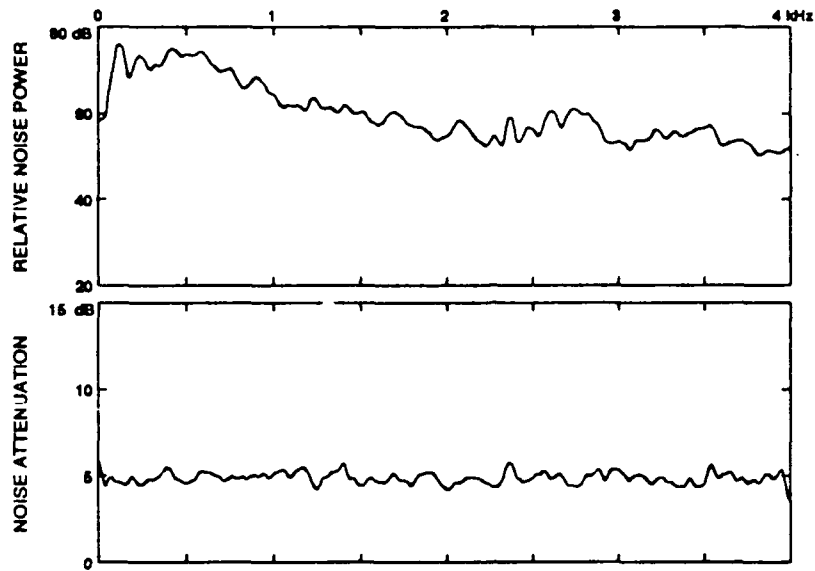
Figure 4.17: Noise attenuation, F-4E outside mask (**F4EH11**)
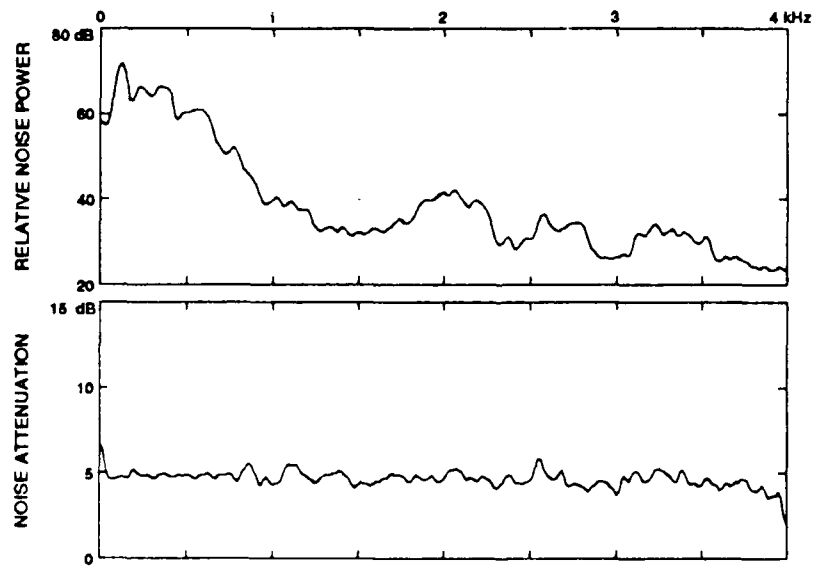
Figure 4.18: Noise attenuation, F-4E inside mask (**F4EM11**)
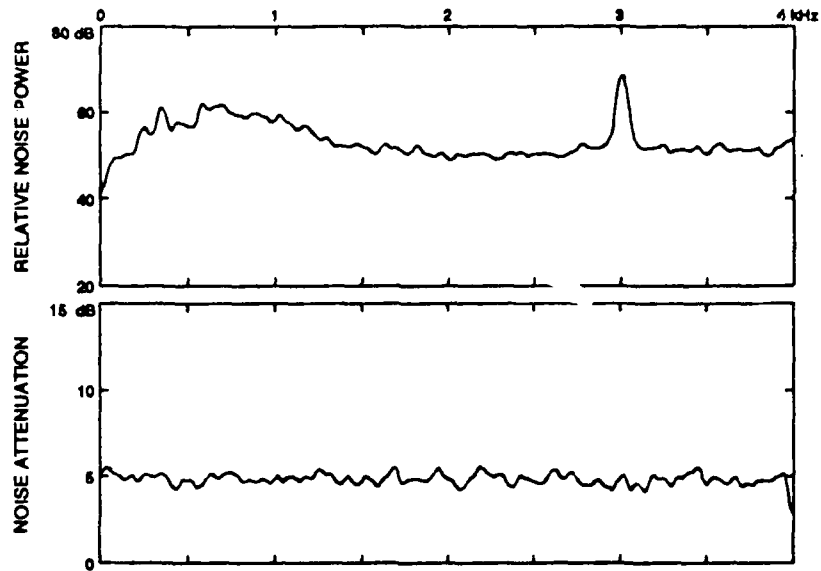
Figure 4.19: Noise attenuation, F-15A outside mask, 1.3 Mach (**F15C59**)
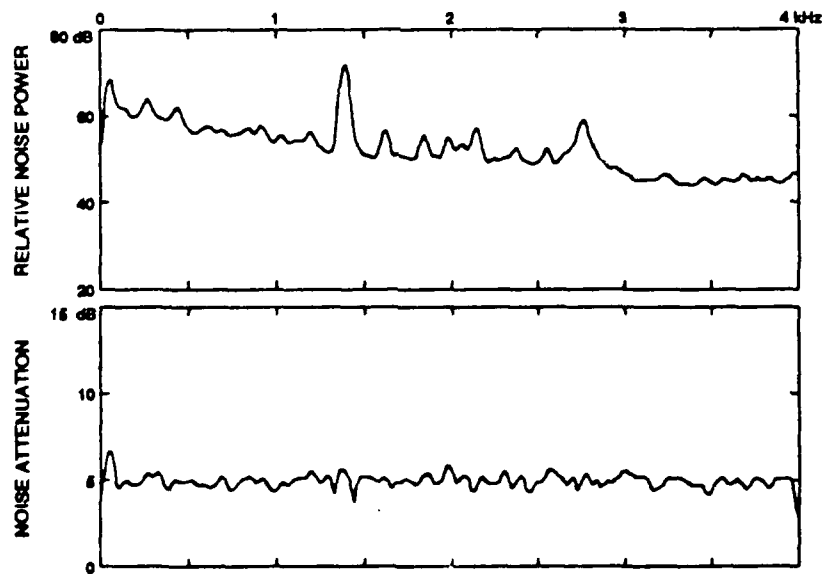
Figure 4.20: Noise attenuation, HH-53 cockpit (**HH53**)

Chapter 5

Nonstationarity

5.1 Nature of the Nonstationarity Problem

Are aircraft noise sources stationary? When we ask whether a random process is stationary, we are asking whether its joint probability distributions are unchanged by the passage of time. No real-world process can attain this ideal, if only because of the Second Law of Thermodynamics. The question to be asked, then, is whether the statistics of an aircraft noise source change significantly over a certain time scale of importance to us. For the purposes of this study, time scales of interest are in the range from 20 ms to 1 s.

If the form of the random process is known, we can test it for stationarity by estimating parameters of the process at two separated times, and then comparing the parameter estimates. Confidence bounds would be evaluated showing the expected range of variation of the estimatess from frame to frame assuming no change in noise characteristics. Then the observed variations would be compared with these bounds, and if they exceeded the calculated limits the variations would be ascribed to nonstationarity. For example, if the random process being tested were known to be white Gaussian noise, this procedure could be used in conjunction with the properties of the short-time Fourier transform of Gaussian noise processes discussed in Chapter 4.

However, for the acoustic noise environments we are studying, the form of the noise process is not known. Therefore non-parametric statistics are needed to test for variation in the noise.

5.2 Design of the Experiment

In order to isolate noise components that cause variation in narrow frequency bands only, we decided to base our nonstationarity tests on short-term values of power spectrum estimators. Because of the importance of the short-time Fourier

transform in spectral restoration, we chose to concentrate on the periodogram PSE. We therefore apply our non-parametric tests to the squared magnitudes of short-term Fourier transform values, considering these as local estimates of noise power in narrow frequency bands.

Assuming that periodogram power spectrum density estimates, $\hat{P}_f(t) = |X_f(tM)|^2$, have been obtained for successive frames of noise, the problem is then to test the stationarity of the sequence $\{\hat{P}_f(t)\}$ at any particular frequency f.

Our approach is to segment a noise recording into *batches* (typically 0.1 s to 1.5 s in length), and within each batch to test, one frequency at a time, for a difference between the first and second half of the batch. Within each half-batch, we use the short-term estimates $\hat{P}_f(t)$ on individual frames (typically 2 to 8 ms in length), so that for each frequency we have two sets of power estimates, one from each batch. Then we utilize a non-parametric test of distribution difference on the two half-batches of PSE's at each frequency.

For example, suppose that the successive estimates $\hat{P}_f$ are separated into two blocks of consecutive estimates

$$A_1 A_2 \cdots A_k B_1 B_2 \cdots B_k$$

The block size, k, is chosen to be commensurate with the time over which noise characteristics are expected to change. Several useful nonparametric methods are available to test whether the A block represents a significantly different distribution than the B block or whether the difference could occur as a natural fluctuation of a stationary process. Strictly speaking, these tests require the successive $\hat{P}_f$ to be statistically independent, a result that can be assured, for example, by leaving short spaces between the successive data records selected for analysis. Then the $\hat{P}_f$ values are ranked from largest to smallest and tagged with their group: For example, we might obtain:

$$A_{21} A_{11} B_6 A_{15} A_{29} B_{15} \cdots A_2 B_{31} A_8 B_7$$

According to the null hypothesis that no difference exists, this sequence of A's and B's should be a purely random arrangement. If, however, we should observe a sequence

$$A_6 A_{15} A_3 A_7 B_{13} A_2 A_{12} \cdots B_{15} B_{10} A_{11} B_{21} B_{11} B_3 B_{19}$$

containing mostly A's at the beginning and B's at the end we would reject the assumption of no difference.

One statistic for evaluating the randomness of the sequence of A's and B's is the Mann-Whitney U statistic [26], which is reputed to be one of the most powerful of the nonparametric methods for assessing changes in the population. It is sensitive to changes in population shape as well as shifts in mean location. Moreover, for block sizes larger that about 6, the Mann-Whitney statistic is

approximately normally distributed. In order to allow for ties when the A_i and B_j are ranked, we have used the following variant of the Mann-Whitney statistic:

$$U = \sum_{i=1}^{k} \sum_{j=1}^{k} D_{ij} \tag{5.1}$$

where

$$D_{ij} = \begin{cases} 1 & \text{if} \quad A_i > B_j \\ 0 & \text{if} \quad A_i = B_j \\ -1 & \text{if} \quad A_i < B_j \end{cases}$$

Under the null hypothesis that both samples were drawn from the same distribution, U is approximately normally distributed for $k \geq 6$ [14, 26], with mean 0 and variance [26]

$$\text{Var}(U|H_0) = [k^2(2k + 1)/3] \left[1 - \frac{\sum(t^3 - t)}{2k(4k^2 - 1)}\right], \tag{5.2}$$

where the sum is extended over all ties, t being the number of samples tied at a single value. Unless the number of ties is large, we can use the conservative approximation

$$\text{Var}(U|H_0) \approx k^2(2k + 1)/3. \tag{5.3}$$

Using this approximation, it is convenient to work with the normalized statistic Z defined as

$$Z = U/\sqrt{k^2(2k + 1)/3}; \tag{5.4}$$

then (again under the null hypothesis) for $k \geq 6$ the statistic Z is approximately normally distributed with a variance slightly less than 1.

Given a signal to be tested for stationarity, we segment the signal into one or more "batches". Each "batch" consists of $2k$ successive "frames" of the signal, where k is a parameter of the analysis. At each of several frequencies, a power spectrum estimate is obtained for each frame, yielding within each half-batch a sequence of spectrum estimates at each frequency f:

First half-batch:	$\hat{P}_f(1)$	$\hat{P}_f(2)$	$\hat{P}_f(3)$	$\cdots \quad \hat{P}_f(k)$
Second half-batch:	$\hat{P}_f(k+1)$	$\hat{P}_f(k+2)$	$\hat{P}_f(k+3)$	$\cdots \quad \hat{P}_f(2k)$

At each frequency, the Mann-Whitney test is then applied to test the hypothesis that both sequences are drawn from the same distribution. Thus we are testing for variation on a time scale comparable to the size of a half-batch. If the distribution of noise energy at a particular frequency changes between the two half-batches, the change should result in a large value of the normalized Mann-Whitney statistic Z computed between the two half-batches. On the other hand, if both half-batches are drawn from the *same* distribution the statistic Z is approximately normally distributed with a mean of zero and a

variance of unity. If we take stationarity as our null hypothesis, then large values of Z at one frequency would lead to rejection of the null hypothesis that the distribution of spectrum estimates at that frequency is unchanging.

Parameters of the a..a.ysis subject to tradeoffs are:

1. batch size ($2k$),

2. frequency resolution,

3. spectrum estimate bias,

4. time resolution.

If the length (in seconds) of each sampled data record is T, then the length of a half-batch is kT. Since our experimental design is based on the difference, or lack thereof, between two successive half-batches, we must choose the half-batch length kT 'o be no more than the time interval over which we wish to observe changes. On the other hand, as we decrease the record length T we also decrease our frequency resolution or increase the bias of our spectrum estimates. If the batch size k is smaller than about 6, the Mann-Whitney statistic is no longer approximated well by a normal distribution.

Because of this tradeoff between frequency resolution, batch size, and time resolution, a single noise environment can be subjected to multiple analyses. As a result of our earlier work and because of the considerations given in Chapter 1, we have focused on time scales of about 200 ms. Shorter intervals do not provide enough information for our statistical tests, while intervals much longer than a few seconds are not of interest for speech coding and enhancement applications.

5.3   False Rejection of the Null Hypothesis

Since we are ..king a large number of tests, we can expect that Z will have large values in some cases by chance: Figure 5.1 shows an analysis performed on simulated white Gaussian noise. In this case, the value of Z exceeded the 1% significance level in 5 out of 832 tests, and if many such analyses were performed on simulated white Gaussian noise we should expect values of Z this large in 1% of all the tests. It is not necessarily meaningful that Z exceeds the indicated 5% or 1% significance levels in a few cases, unless there is an evident pattern such as a grouping of large values of Z in a narrow range of frequencies, or a repetition of large values of Z at the same frequencies across multiple analyses.

To quantify this, it is necessary to decide how many over-threshold values should be regarded as significant. In a series of M independent tests, the number of values of Z above the critical value for a certain significance level α will be

Figure 5.1: Mann-Whitney analysis for white Gaussian noise

approximately a Poisson random variable with expectation αM. In other words, the probability that exactly k values will exceed the α significance level will be

$$p(k) = e^{-\alpha M}(\alpha M)^k/k!. \qquad (5.5)$$

When evaluating a multi-frequency experiment on a single noise record, if we find that over all frequencies taken as a whole there are m values of Z above a chosen significance threshold, we consider the significance level defined as the probability that m or more values will exceed the threshold under the null hypothesis of stationarity. This probability can be computed directly from the Poisson distribution. We refer to this probability as the "overall significance level" for a set of multiple tests across all frequencies. A value of this probability near zero, then, indicates a result that is unlikely under the stationarity hypothesis. A larger value of this probability indicates a number of over-threshold values that is likely under the stationarity hypothesis, due simply to the large number of frequencies under test.

5.4   Sensitivity of the Test

A small experiment was conducted to determine the effect of two principal parameters (half-batch size and frequency resolution) on the sensitivity of the batched Mann-Whitney test. Our approach was to examine the ability of the

	Half-batch size (ms)			
Frequencies	12	48	192	768
16	> +12	≤ −6	≤ −6	≤ −6
64	—	≤ −6	0	≤ −6
256	—	—	+6	+3

Table 5.1: Speech-to-noise ratio (dB) required for detection of nonstationarity

test to detect nonstationarity in a signal masked by more-or-less stationary noise. The method used was to vary the amount of nonstationary noise (through simple attenuation) and, for each combination of test parameters, determine the minimum nonstationary noise level at which the test was effective.

Since we are interested in signal changes occurring on the time scale of speech nonstationarity, we used an actual speech signal (file [200,220]TOMS.DAM on the speech database) as the nonstationary noise source. For the more-or-less stationary noise source we used a recording of HH-53 helicopter noise (file [200,230]HH53.FLT on the digital noise database). The test files were then created by mixing the two noise sources at several different speech-to-noise (or "nonstationary"-to-"stationary") ratios. The ratios, expressed as ratios of high-energy speech frames to the RMS helicopter noise, ranged from +12 dB to −6 dB in steps of 3 dB.

After performing batched Mann-Whitney analyses of the test files with several different choices of frequency resolution and half-batch length, we examined the results (at different speech-to-noise ratios) for each parameter combination to determine the minimum speech-to-noise ratio needed to show significant nonstationarity. The smaller the ratio required, the more sensitive the test was judged to be. The results are summarized in Table 5.1. Combinations labeled "—" were not tested because the half-batch size would have been less than 6.

Tests conducted with a frequency resolution of 256 points (15.625 Hz spacing) seem to be much less sensitive than tests at coarser frequency resolution using the same half-batch length. This insensitivity may be due in part to the small number of data records in each half-batch. Little discriminatory power was found with a half-batch size of 12 ms (96 samples).

Chapter 6

Nonstationarity Experiments in Aircraft Noise

This chapter presents the results of batched Mann-Whitney analyses performed on the aircraft noise samples described in Chapter 2, using the analysis method described in Chapter 5. For each 5-second noise sample, four batched Mann-Whitney analyses were performed with the parameters:

- 16 frequencies, 192ms half batch (96 frames per half batch)

- 64 frequencies, 192ms half batch (24 frames per half batch)

- 64 frequencies, 48ms half batch (6 frames per half batch)

- 64 frequencies, 768ms half batch (96 frames per half batch)

The following noise database files were analyzed:

1. EC-135 Battle Staff area, tape 21N (EC135B)

2. E-4B Battle Staff area, tape 12N (E4BBS)

3. E-3A console 13, tape 201 (E3AC13)

4. EC-130 ABCCC, tape C (EC130A)

5. HC-130, tape 1A, channel A (HC130A)

6. HC-130, tape 1A, channel B (HC130B)

7. P-3C, tape ND (P3C)

8. F-15A helmet microphone, tape 5-2 **(F15HT05)**

9. F-15A mask microphone, tape 5-1 **(F15MT05)**

10. F-15A cockpit microphone, 1.2 Mach **(F15C33)**

11. F-15A cockpit microphone, 1.3 Mach **(F15C59)**

12. F-16A helmet microphone, tape 6-2 **(F16H06)**

13. F-16A mask microphone, tape 6-1 **(F16M06)**

14. A-10 helmet microphone, tape 8-2 **(A10H25)**

15. A-10 mask microphone, tape 8-1 **(A10M24)**

16. F-4E helmet microphone, tape 11-2 **(F4EH11)**

17. F-4E mask microphone, tape 11-1 **(F4EM11)**

18. F-4E helmet microphone, tape 11-2, low altitude, 500 kt, **(F4EHL)**

19. Tornado pilot position, demist on, tape TOR **(TOR34)**

20. HH-53 cockpit, tape 1AA **(HH53)**

Table 6.1 shows the "overall significance levels," as defined in Section 5.3, for all the analyses. Levels less than 1% are shown in italics, representing analyses in which over all frequencies, large values of Z occur more frequently than would be likely if the noise were stationary.

6.1   Interpretation of Plots

Figures 6.1 through 6.20 present these analyses. Each of the figures shows the results of three analyses of the same noise record; from bottom to top, they are the 192-ms analysis with 250 Hz resolution; the 192-ms analysis with 62.5 Hz resolution; and the 768-ms analysis with 62.5 Hz resolution. Each "X" mark represents one value of Z, for one "batch" and at one frequency. A hollow square at the top edge of the plot represents a value of Z that exceeds 4. The dotted lines ($Z = 1.96$ and $Z = 2.58$) are the critical values of the unit normal distribution for the significance levels 5% and 1%, respectively[1]. For the lower two plots, both of which have 192-ms half-batches, the 5 sec of data is divided into 13 batches, and so there are 13 points plotted for each frequency. For the upper plot, with 768-ms half-batches, the 5 sec of data supplies 3 batches, and there are 3 points plotted for each frequency.

---

[1] As we discussed in Chapter 5, Z is approximately normally distributed with mean 0 and variance 1.

		Half-Batch Length, Frequency Resolution			
Aircraft	Figure	192 ms, 250 Hz	192 ms, 62.5 Hz	48 ms, 62.5 Hz	768 ms, 62.5 Hz
EC-135	6.1	1.96	13.64	86.66	12.87
E-4B	6.2	6.02	59.07	100.00	85.34
E-3A	6.3	34.50	72.41	99.90	12.87
EC-130	6.4	87.51	1.16	100.00	100.00
HC-130	6.5	87.51	91.73	99.90	*0.36*
HC-130	6.6	6.02	59.07	98.80	100.00
P-3C	6.7	87.51	21.71	99.95	4.57
F-15A	6.8	100.00	59.07	100.00	4.57
F-15A	6.9	1.96	4.46	100.00	100.00
F-15A	6.10	61.52	32.38	99.98	*<0.01*
F-15A	6.11	*<0.01*	*<0.01*	95.28	*0.08*
F-16A	6.12	15.76	99.77	100.00	30.17
F-16A	6.13	*<0.01*	59.07	100.00	57.19
F-4E	6.16	1.96	45.20	86.66	2.34
F-4E	6.17	*0.03*	1.16	99.98	*<0.01*
F-4E	6.18	*0.14*	2.34	98.03	*<0.01*
A-10	6.14	100.00	25.53	97.50	100.00
A-10	6.15	*0.42*	11.70	100.00	13.52
Tornado	6.19	1.96	91.73	99.99	1.38
HH-53	6.20	*<0.01*	*0.02*	26.35	*<0.01*

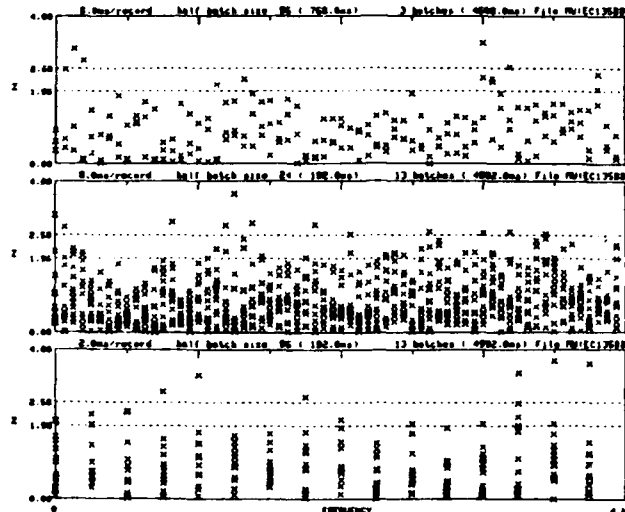Table 6.1: Overall significance levels (%)

Figure 6.1: Mann-Whitney analysis for EC-135 Battle Staff area (**BC135B**)

6.2   Survey by Aircraft Type

6.2.1   Large Jet Aircraft

Noise in these aircraft (E-4B, E-3A, EC-135) is characterized by a continuous spectrum usually lacking strong sinusoidal components. (Both in overall sound level (88 dB) and in general spectral shape, noise recorded in the E-4B Battle Staff area is quite similar to noise recorded in the E-3A, except that the E-3A has a sharper low-frequency peak and some tonal engine noise.) There is no evidence of nonstationarity in the noise analyzed from these aircraft. (See Figures 6.1, 6.2, and 6.3.)

6.2.2   Turboprop Aircraft

Noise in these aircraft (EC-130, HC-130, and P-3C) includes some strong low-frequency sinusoidal components at the first few harmonics of the propeller blade passage rate (typically about 70 Hz). Much of the remainder of the spectrum is rather smooth.[2] There are some large values of Z in the EC-130 analysis with a time scale of 192 ms and a resolution of 62.5 Hz, but the overall significance level (1.16%) is not decisive. Some large values of Z appear between 3000 and 4000 Hz in one of the HC-130 analyses (Figure 6.5), this time with an overall
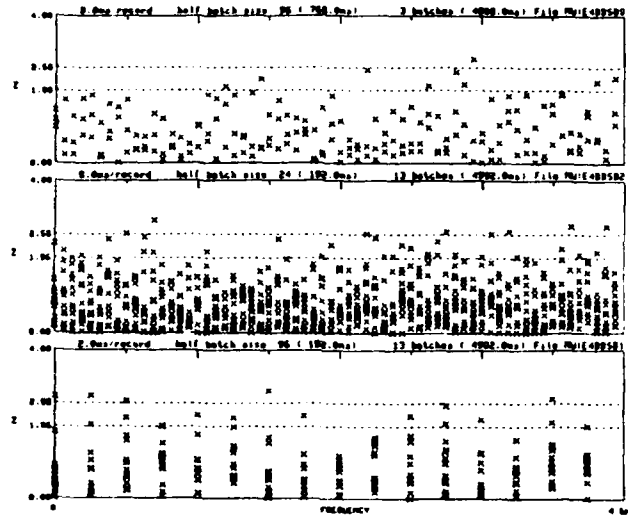
---

[2] See Figures 4.10, 4.11, and 4.12.

Figure 6.2: Mann-Whitney analysis for E-4B Battle Staff area (**E4BBS**)
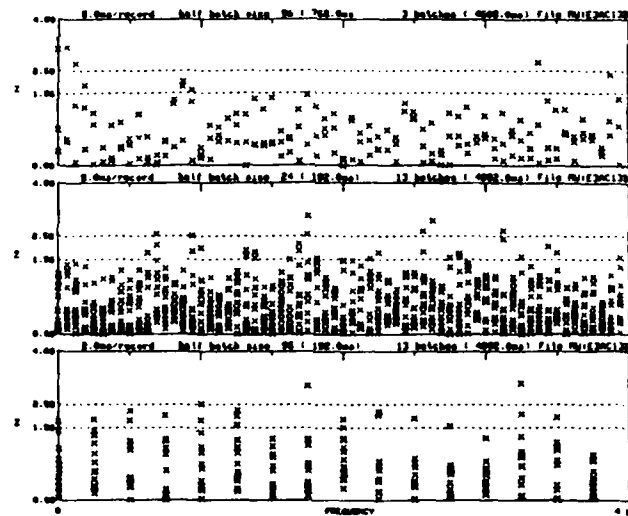

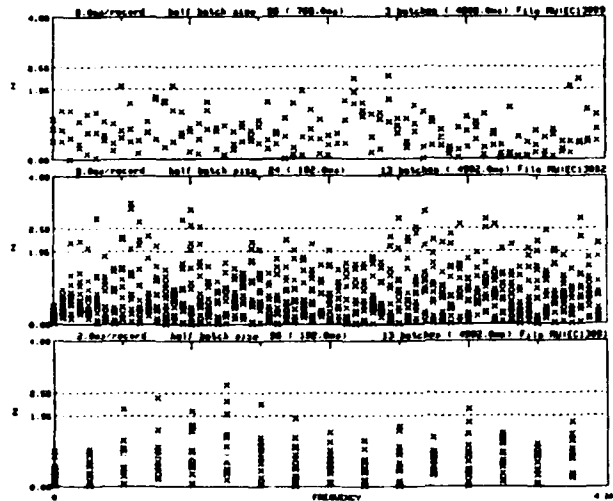Figure 6.3: Mann-Whitney analysis for E-3A console 13 (**E3AC13**)

Figure 6.4: Mann-Whitney analysis for EC-130 ABCCC (**EC130A**)

significance level of 0.36%. Except for these two analyses, there is no evidence of nonstationarity among the turboprop aircraft.

6.2.3   Fighter/Attack Aircraft

These aircraft include the F-15A, F-16A, A-10, F-4E, and Tornado. (See Figures 6.8–6.19.) In some of these aircraft we have found some evidence of nonstationarity in narrow frequency ranges, probably associated with variation of tonal noise sources. The two 1976 F-15A segments, both made at supersonic speed, show apparent variation of a tonal source near 1 kHz (Figure 6.10) and near 3 kHz (Figure 6.11). The 1982 recordings made at lower speed in another F-15A show little evidence of nonstationarity either inside or outside the oxygen mask.

The recording made inside the oxygen mask of an F-16A shows significant nonstationarity across a broad range of frequencies in the analysis made with a frequency resolution of 250 Hz and a time scale of 192 ms. No such nonstationarity is evident outside the mask at the same time, and we conjecture that there may be more-than-usual breath noise or valve noise present.[3]

The A-10 recording made inside the oxygen mask shows some evidence of nonstationarity concentrated in low frequencies. The F-4E in-mask recording,

---

[3]However, this segment of noise, like all the other in-mask segments, was taken from an interval during which the pilot was attempting to hold his breath.
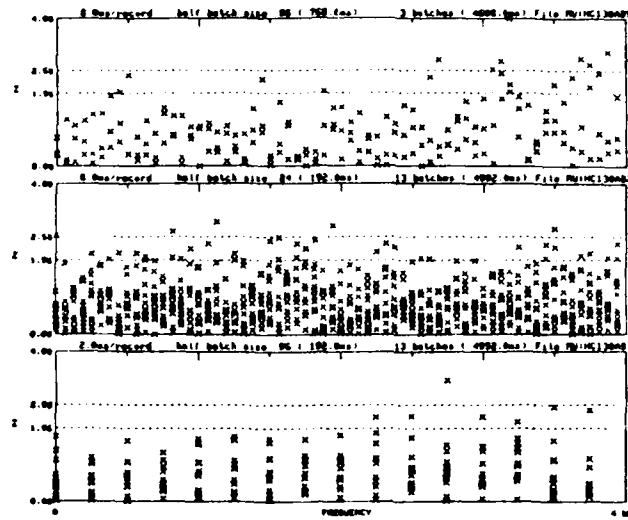
Figure 6.5: Mann-Whitney analysis for HC-130 radio operator position, mic. 1 (HC130A)
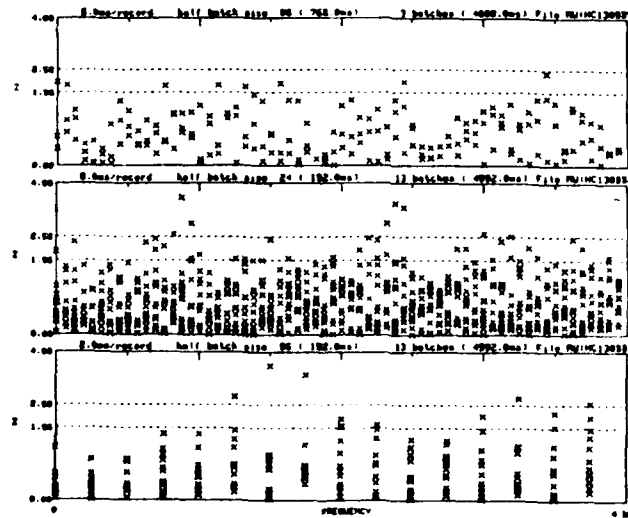

Figure 6.6: Mann-Whitney analysis for HC-130 radio operator position, mic. 2 (HC130B)
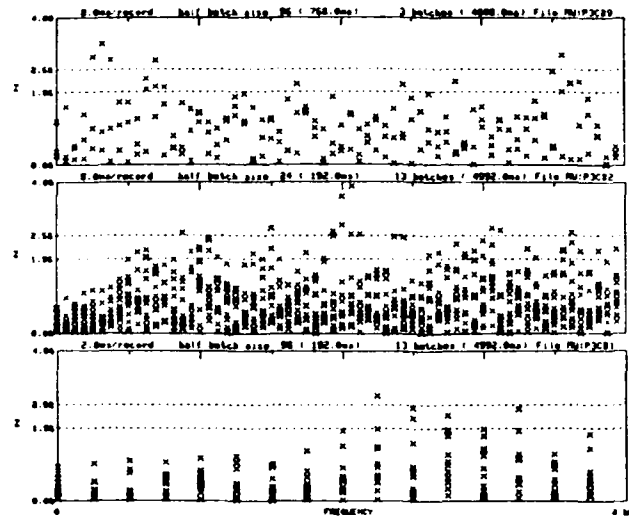
Figure 6.7: Mann-Whitney analysis for P-3C **(P3C)**

in one analysis, shows evidence of nonstationarity, mostly concentrated in the vicinity of 3 kHz. On the other hand, *no nonstationarity is apparent in the* noise analyzed from the Tornado aircraft.

6.2.4   Helicopters

In the HH-53 there is significant evidence of nonstationarity in the frequency range 1300-1700 Hz, and also at frequencies of 2500 Hz and higher (see Figure 6.20).

6.3   An Example of Tone Removal

We conclude with a specific example of the impact of non-stationarity on spectral restoration. In Section 6.2.3 we found non-stationarity in the F-15A noise recorded in supersonic flight, a non-stationarity that appeared as variation in the power levels near 3000 Hz. Figure 6.21 shows a comparison of the input and output of a spectral restoration algorithm (with parameters $c = 1/12$, $\beta = 1$, $\mu = 1$, Hamming window) applied to this noise signal alone[4]. The attenuation

---

[4] These are the same parameters used in the examples of Chapter 4, except that $\mu = 1$. The difference in μ accounts for a difference in the resulting attenuation, which is typically about 11 dB compared to the 4.3 dB found for $\mu = 2$
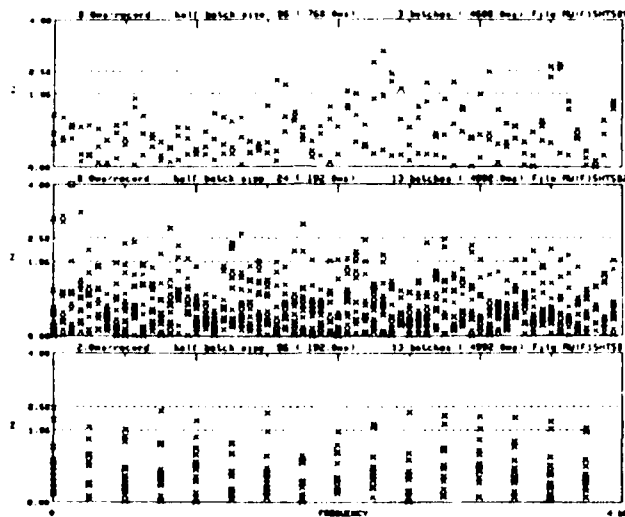
Figure 6.8: Mann-Whitney analysis for F-15A (outside mask) (F15HT6)
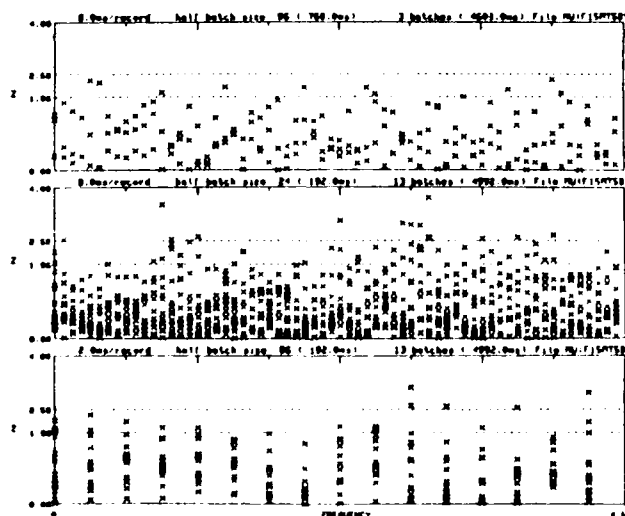
Figure 6.9: Mann-Whitney analysis for F-15A (inside mask) (F15HT6)
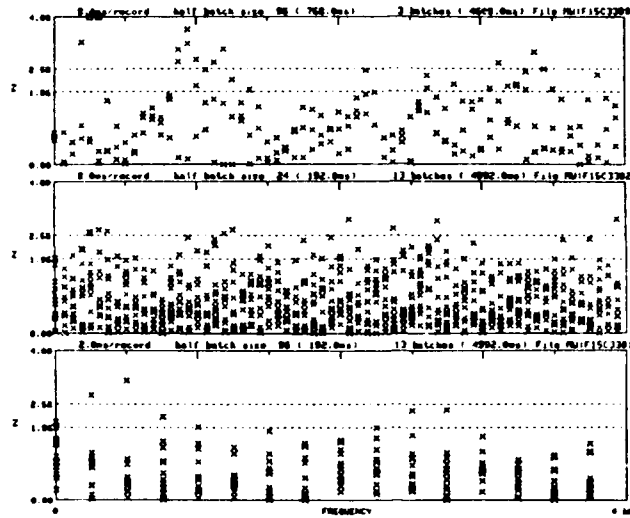
Figure 6.10: Mann-Whitney analysis for F-15A (outside mask) (F15C33)
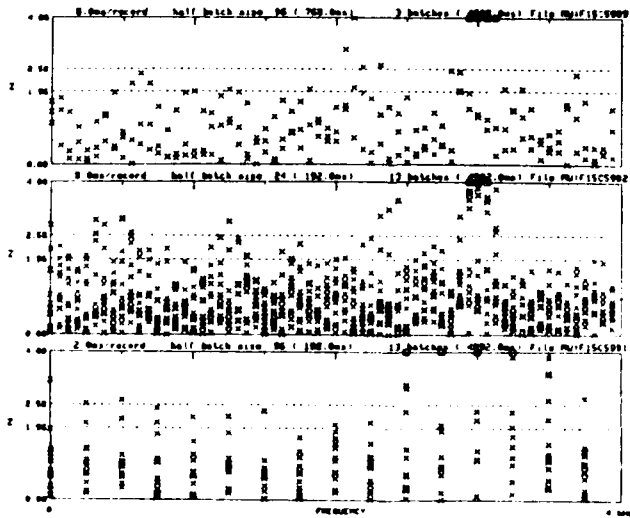
Figure 6.11: Mann-Whitney analysis for F-15A (outside mask) (F15C59)
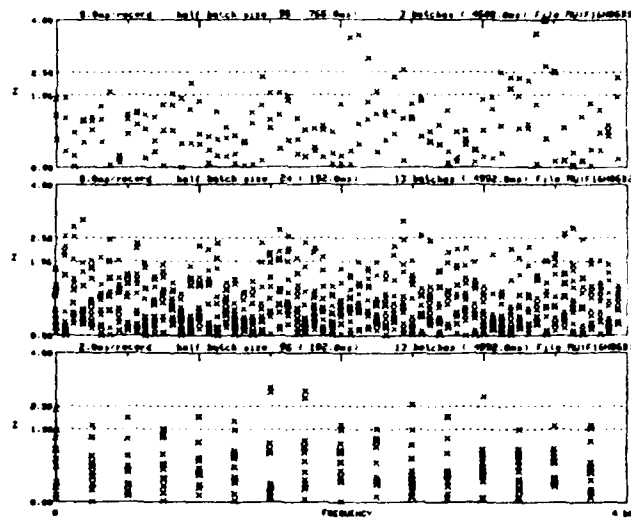
Figure 6.12: Mann-Whitney analysis for F-16A (outside mask) (F16H06)
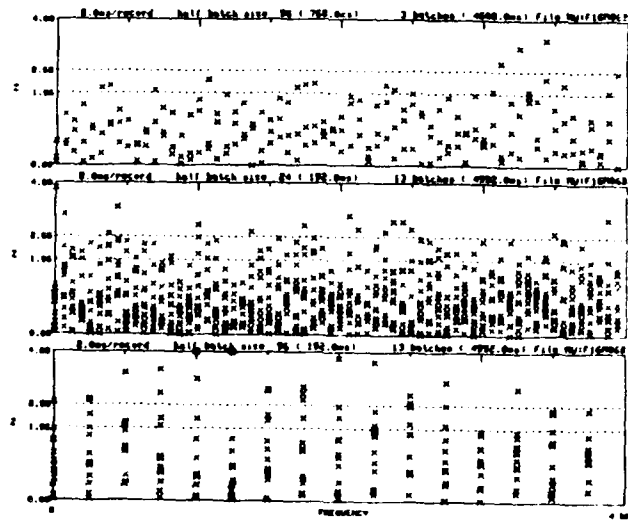
Figure 6.13: Mann-Wh't*ey analysis for F-16A (inside mask) (F16M06)
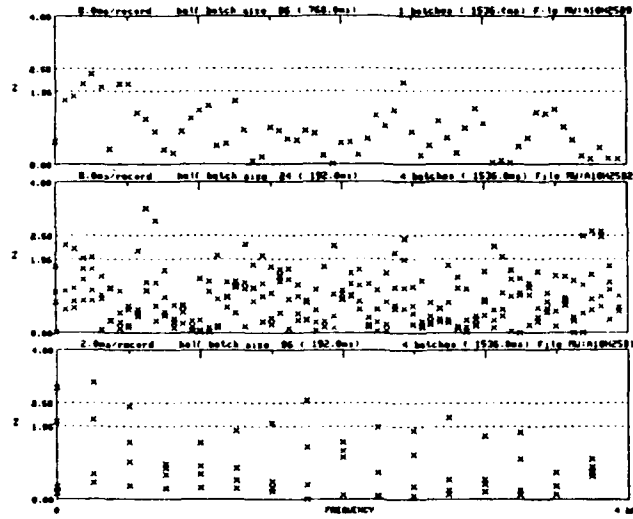
Figure 6.14: Mann-Whitney analysis for A-10 (outside mask) **(A10H25)**
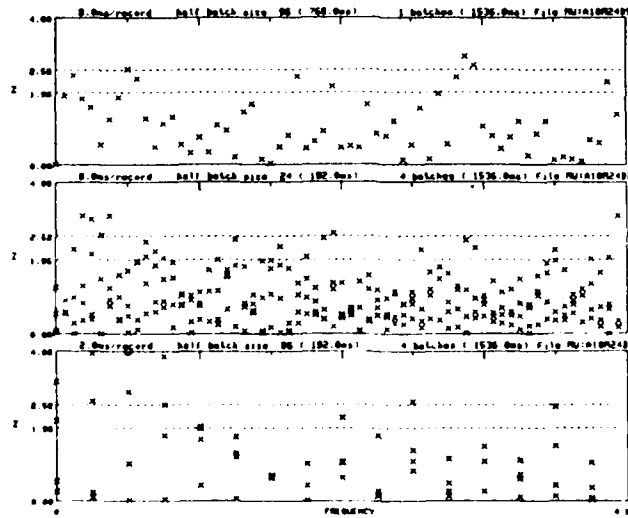
Figure 6.15: Mann-Whitney analysis for A-10 (inside mask) **(A10M24)**
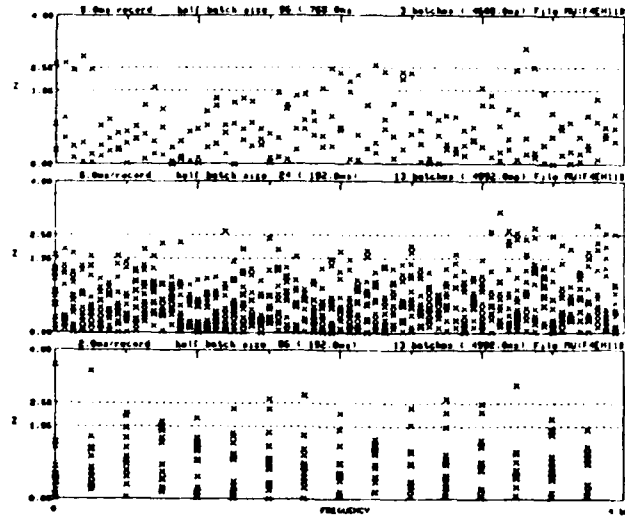
Figure 6.16: Mann-Whitney analysis for F-4E (outside mask) (**F4EH11**)
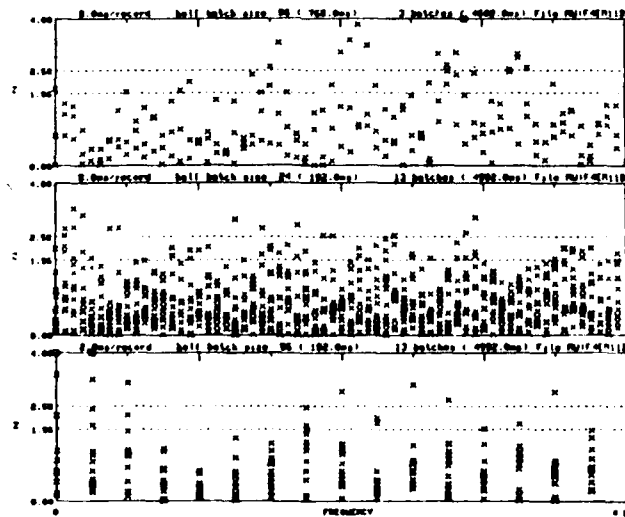
Figure 6.17: Mann-Whitney analysis for F-4E (inside mask) (**F4EM11**)
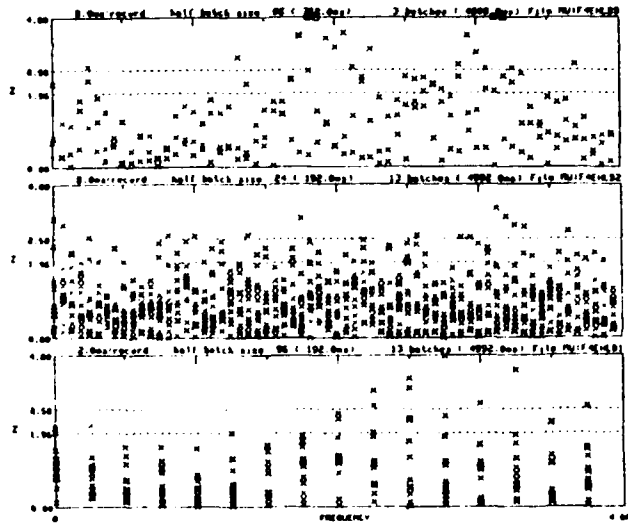
Figure 6.18: Mann-Whitney analysis for F-4E, low alt. (outside mask) **(F4EHL)**
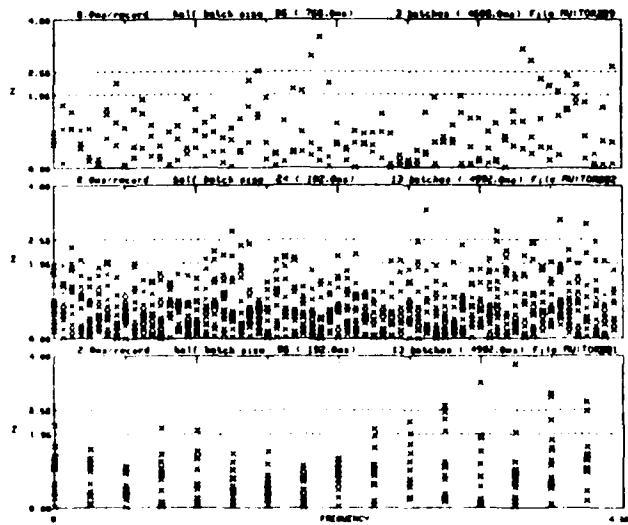
Figure 6.19: Mann-Whitney analysis for Tornado, outside mask **(TOR34)**
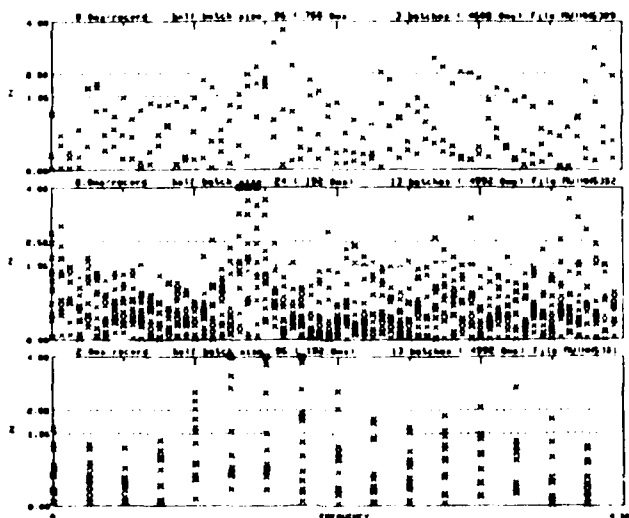
Figure 6.20: Mann-Whitney analysis for HH-53 helicopter (**HH53**)

of the variable tonelike noise near 3 kHz is about 12 db, almost the same as the attenuation found at other frequencies. Figure 6.22 shows a similar comparison, but with a different, synthetic input. The synthetic noise is real noise recorded in an F-15A (in subsonic flight), to which a pure sinusoid at 3 kHz has been added. In this case, the spectral restoration algorithm attenuates the pure sinusoid by about 23 dB, compared to the 11-dB attenuation at other frequencies.
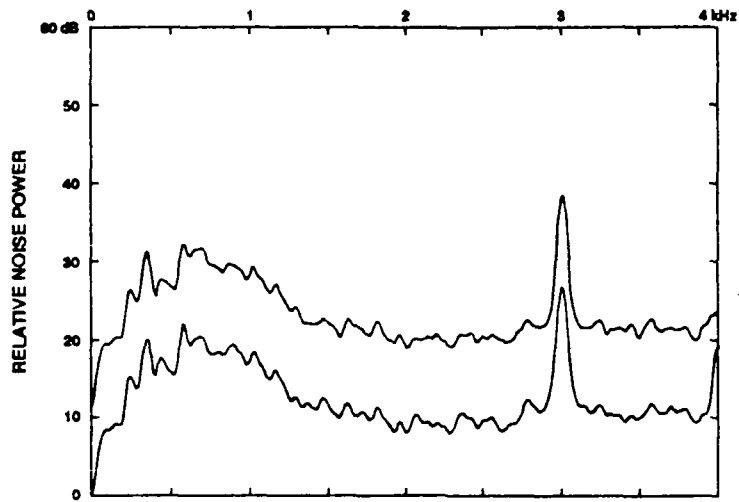
Figure 6.21: Real F-15A noise before and after enhancement **(F15C59)**
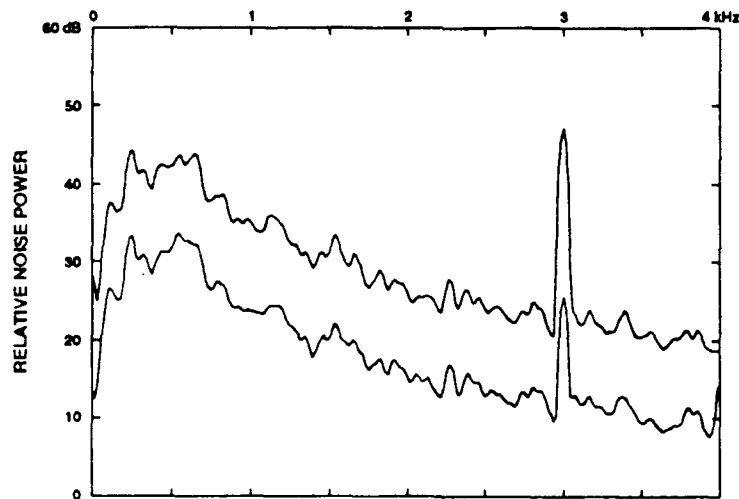
Figure 6.22: Synthetic "F-15A noise" before and after enhancement **(F15HT5S)**

Chapter 7

Summary

In this study of short-term noise variation in Air Force platforms, we have followed two avenues of investigation. In the first, we applied quantitative measures of variation to individual noise recordings, and compared the results across various aircraft. In the second, we applied non-parametric hypothesis tests to search for nonstationarity in the same noise recordings. In both efforts, we have gone on the hypothesis that aircraft noise may be the sum of some noise that is essentially stationary and some noise that is nonstationary but only affects some parts of the 0–4 kHz vocoder range.

We devised two simple frequency-dependent measures of short-term variation: the standard-deviation-to-mean ratio and the residual RMS prediction error, both applied to short-term power spectrum estimates. Each of these measures gives a number at each frequency, and is intended to isolate narrow-band nonstationarity. For white Gaussian noise, we obtained the expected value of the standard-deviation-to-mean ratio; this value can be used as a guide to interpreting values for real-world noise. The RMS prediction error measurement, which is motivated by a very simple model of spectral restoration, measures the discrepancies between single-frame STFT magnitudes and their short-term estimators based on the recent past. Both of these simple quantitative measures of variation showed distinctively low variation at low frequencies in turboprop aircraft, but seemed to be too variable to draw more precise conclusions.

We analyzed theoretically the behavior of a broad class of spectral restoration algorithms for the special case of noise-only inputs, and used the performance of such algorithms as a gauge to locate differences between aircraft types. Using noise-only performance as a criterion, we found that spectral restoration had superior performance in removing propeller noise in turboprop aircraft, and in removing tonal noise in one particular recording from an E-3A, but that generally the performance of spectral restoration was nearly the same as that predicted theoretically for white Gaussian noise. This was true across all frequencies, and applied to time-varying tonal noise as well as noise whose spectrum is smoother.

To test for nonstationarity, we used the nonparametric Mar n-Whitney statistic in an experimental design that compared batches of short-term power spectrum estimates over adjacent 192-ms (or 768-ms) intervals. We found little or no evidence of nonstationarity in the noise recordings from large jet or turboprop aircraft with wing-mounted engines. In fighter aircraft noise recordings, the picture was less uniform. In some fighter aircraft there was strong evidence of nonstationarity that appeared to be confined to more or less narrow frequency ranges. Finally, in the helicopter noise recordings we studied, we again found evidence of nonstationarity concentrated at certain frequencies but leaving substantial parts of the spectrum unaffected.

Finally we return to the question: does nonstationarity limit the performance of spectral restoration in these aircraft? Table 7.1 presents the nonstationarities found (in the left-hand column) and the corresponding peaks or valleys, if any, of the spectral-restoration attenuation figure (in the right-hand column). A dash in the right hand column means that the attenuation was close to the "Gaussian" 4.3 dB figure all across the frequency range. In only a single case, at a single frequency, was a finding of nonstationarity coupled with a significant drop in attenuation below 4.3 dB. In two other cases, nonstationary that is narrowly confined in frequency results in poorer noise attenuation than would be expected for truly tonal noise, but no worse attenuation than is normal for broadband noise.

Aside from these cases, the nonstationarity that we found did not have any apparent effect on the attenuation achieved by spectral restoration. If spectral restoration can perform as well against a nonstationary noise source as it does against white Gaussian noise, then it cannot be said that the nonstationarity itself is the culprit. Therefore we conclude that the kinds of nonstationarity that we found in real aircraft did not degrade the performance of spectral restoration, as measured against noise-only inputs.

Aircraft	Non-stationarity found	Spectral restoration effects
E-4B	—	—
E-3A	—	—
EC-135	Marginal, 250Hz/192ms analysis only	Some extra attenuation of 'tones', 3400 Hz and 3750 Hz
EC-130	—	Greatly increased attenuation at low frequencies and at 750 Hz spectral peak
P-3C	Mild indication at 2050 Hz	Greatly increased attenuation at low frequencies; reduced attenuation at 2050 Hz
HC-130	3000-4000 Hz, one analysis only, one microphone only	Less dramatically increased attenuation at low frequencies
F-15A	Confined to narrow bands, and only in supersonic and in-mask recordings	—
F-16A	Corrupted recording?	—
A-10A	Low frequencies	—
F-4E	In mask, one analysis only, narrow band at 3000 Hz	—
Tornado	—	—
HH-53	Narrow band near 1500 Hz	Increased attenuation near 80 Hz only, decreased attenuation near tone at 1500 Hz

Table 7.1: Comparison of Mann-Whitney nonstationarity findings with irregularities in spectral restoration attenuation, noise only

Appendix A

Detailed Description of Noise Records Used

The digitized noise recordings used in this study were all taken from the Digital Acoustic Noise Database, a collection of files that represent a subset of the entire RADC/EEV Acoustic Noise Database. This appendix supplements the information given in Chapter 2 by describing the aircraft on which those recordings were made, and the circumstances of recording. For clarity, all the digital noise file names are given in a distinctive type face (for example: E3AC13). Where figures in this report are based on data from particular noise files, these files are identified in the figure captions.

The aircraft information in the following pages is taken primarily from the annual volumes of Jane's *Aircraft of the World*. In a few cases, where a military aircraft type resembles a commercial aircraft type, we have given information on the commercial version where none was available for the military version.

E-4B

The E-4B Advanced Airborne Command Post is based on the commercial Boeing 747 airframe and is the successor to the EC-135 as a strategic command and control platform. Recordings were made [37] in three areas of the E-4B: the Battle Staff work area, near the middle of the aircraft; the briefing room, just forward of the Battle Staff area; and the National Command Authority (NCA) compartment. During the recordings, the aircraft was in normal, level flight.

The file **E4BBS** was digitized from the recording made in the Battle Staff area, designated field tape 12N.

E-4B National Emergency Airborne Command Post	
Reference	Jane's 82-83 p. 333(747-200B), p. 339(E-4B)
Manufacturer	Boeing (modified commercial 747-200B)
Primary Mission	Strategic Command & Control
Power plant	Four General Electric CF65-50E turbofan engines with 52500lb thrust each
Length	231'4" (70.51 m)
Height	63'5" (19.33 m)
Wingspan	195'8" (59.64 m)
Max T.O. Weight	800000lb (362874 kg)
Airspeed	523 kt (969 km/hr; 602 mph) [max, level, 747]
Mission Endurance	72hr
Cruise Altitude	45000' (13700 m) [747]
Crew	94
Flight Profile	cruise, level

EC-135

The EC-135 was a modified version of the KC-135 tanker. As such it was similar to the commercial Boeing 707. Although the EC-135 was equipped for in-flight refueling of other aircraft, its primary function was command and control. The EC-135 was a predecessor of the E-4B as a strategic command and control platform. Noise recordings [37] were made in July 1982 at the Radio Operator's compartment and in the Battle Staff work area.

The file **EC135B** was digitized from the analog recording made in the Battle Staff area, designated field tape 21N.

EC-135 SAC Looking Glass aircraft	
Reference	Jane's 67-68 p. 309
Manufacturer	Boeing (commercial 707)
Primary Mission	EC-135 Command & Control
Powerplant	Four Pratt & Whitney TF33 turbofans with 18000lb thrust each.
Length	136'3" (41.53 m)
Height	38'4" (11.68 m)
Wingspan	130'10" (39.88 m)
Max T.O. Weight	297000lb (134715 kg)
Airspeed	466 kt (530 mph) [max, level @30000']
Total Mission Time	5.5 hr
Service Ceiling	36000' (11000 m) [707-320B]
Flight Profile	cruise, level flight

E-3A

The E-3A (AWACS) carries a surveillance radar and a crew of radar operators who track hostile targets and control fighter aircraft. The E-3A shares the same basic airframe used in the EC-135 and the commercial Boeing 707. Recordings [34] were made in 1982 at Consoles 4 (Senior Director), 10 (Air Surveillance Technician), 13 (Weapons Director), 25, and 30, among others. The recordings were made during a training mission while operators were present and speaking. The aircraft was in its surveillance orbit.

The file **E3AC13** was digitized from the recording made at Console 13, designated field tape 201.

E-3A Sentry	
Reference	Jane's 80-81 p. 298
Manufacturer	Boeing (based on commercial 707-320B)
Primary Mission	Airborne Warning and Control System (AWACS)
Powerplant	E-3A Four Pratt & Whitney TF33-PW-100 Turbofans with 21000lb thrust each.
Length	152'11" (46.61 m)
Height	41'4" (12.60 m)
Wingspan	145'9" (44.42 m)
Max T.O. Weight	325000lb (147400 kg)
Airspeed	460 kt (853 km/hr; 530 mph) [max, level]
Endurance on Station	6 hr
Service Ceiling	> 29000' (8850 m)
Crew	4 aircrew + 13 AWACS Specialists
Flight Profile	surveillance orbit

EC-130 and HC-130

The EC-130 is a multi-engine turboprop aircraft, a version of the C-130 equipped for the command and control function. An Airborne Battlefield Command & Control Center (ABCCC AN/USC-15) was installed in the EC-130 when noise recordings [35] were made. The recordings were made at seat #1, the communicator's console, and another location in the ABCCC unit. The file EC130B was digitized from the latter recording, designated as tape C.

The HC-130 is a search and rescue variant of the same basic airframe. Noise recordings [35] were made in 1984. The files HC130A and HC130B were digitized from the two channels of the recording, designated field tape 1A.

C-130H Hercules	
Reference	Jane's 84-85 p. 435
Manufacturer	Lockheed
Primary Mission	Military Transport
	EC-130 Command & Control
	HC-130 Search & Rescue
Powerplant	Four Allison T56-A-15 Turboprop
	rated at 4568 ehp each. Four Hamilton
	Standard four-bladed constant speed
	propellers of 13'6".
Length	97'9" (29.79 m)
Height	38'4.5" (11.66 m)
Wingspan	132'7" (40.41 m)
Operating Weight	76469lb empty (34686 kg)
Cruising Airspeed	295 kt (547 km/hr; 340 mph) (547 km/hr)
Maximum Airspeed	325 kt (602 km/hr; 374 mph)
Range (max payload)	2046 nm (37913 km; 2356 mi)
Service Ceiling	33000' (10000 m)
Crew	4 + ABCCC Staff of 12 [EC-130]
Tail Number	FC-130 1836 7th ACCS
Flight Profile	level flight

P-3C

The P-3C is a long-range anti-submarine patrol aircraft, developed from the commercial Lockheed Electra and used by the U. S. Navy. Noise recordings [36] were made in 1978 at the NAVSEA position.

File P3C was digitized from this recording, designated field tape ND.

P-3C Orion	
Reference	Jane's 75-76 p. 432
Manufacturer	Lockheed (based on commercial Electra)
Primary Mission	Naval Anti-Submarine Warfare
Powerplant	Four Allison T56-A-14 turboprops with 4910 hp each. Hamilton-Standard 54H60 four-bladed constant speed propellers of 13'6".
Length	116'10" (35.61 m)
Height	33'8.5" (10.29 m)
Wingspan	99'8" (30.37 m)
Max T.O. Weight	135000lb (61235 kg)
Patrol Airspeed	206 kt (381 km/hr; 237 mph) [@1500']
Airspeed	411 kt (761 km/hr; 473 mph) [max, level @15000']
Mission Radius	2070 nm (3835 km; 2383 m) [max - no time on station]
Service Ceiling	28300' (8600 m)
Crew	10
Flight Profile	350 kt, 25000'

F-15

The F-15A is a twin-engine single-seat air-superiority fighter. Recordings were made by H. Hille aboard an F-15A in 1976. Unfortunately, we have no absolute calibration for these recordings and so we can say nothing about the absolute noise levels. The 1976 recordings include short segments of noise during level flight at 6000, 25000, and 40000 ft; during a climb from 10000 ft to 25000 ft; and during a climb from 25000 to 40000 ft.

The files **F15C33** (level flight, 6000 ft altitude, Mach 1.2), **F15C418** (climb from 15000 ft to 20000 ft, Mach 0.9), **F15C59** (level flight 25000 ft altitude, Mach 1.3), and **F15C68** (climbing to 40000 ft. at Mach 0.88, mil power, some speech present) were digitized from the 1976 tape, designated field tape HHNT.

Later, recordings [25] were made aboard an F-15A in January 1981. The 1981 recordings were made with microphones located both inside the pilot's oxygen mask and on the pilot's helmet.

The files **F15HT5** and **F15MT5** were digitized from field tapes 5-1 and 5-2, and represent simultaneous recordings inside and outside the oxygen mask at a time when the pilot was holding his breath. Files **F15HT10** and **F15MT10** were digitized from field tapes 10-1 and 10-2, but are not used, because of apparent saturation in the original field tape 10-1.

F-15A Eagle	
Reference	Jane's 81-82 p. 403
Manufacturer	McDonnell Douglas
Primary Mission	All-Weather Air-Superiority Fighter
Powerplant	Two Pratt & Whitney F100-PW-100 turbofans with 23930lb thrust each (afterburner at takeoff).
Length	63'9" (19.43 m)
Height	18'5.5" (5.63 m)
Wingspan	42'9.75" (13.05 m)
Max T.O. Weight	56000lb (25401 kg)
Airspeed	> Mach 2.5 [max, level]
Service Ceiling	60000' (18300 m)
Absolute Ceiling	100000' (30500 m)
Crew	Pilot only [F-15B two-seat version]
Tail Number	WA111 and other
Flight Profile	varied

F-16

The F-16A is a highly maneuverable lightweight fighter. Recordings [25] were made aboard an F-16A in January 1981. These recordings were made with microphones located both inside the pilot's oxygen mask and on the pilot's helmet.

The files **F16H06** and **F16M06** were digitized from field tapes 6-1 and 6-2, and represent simultaneous recordings inside and outside the oxygen mask at a time when the pilot was holding his breath.

F-16A Fighting Falcon	
Reference	Jane's 81-82 p. 361
Manufacturer	General Dynamics
Primary Mission	Lightweight Combat Fighter
Powerplant	One Pratt & Whitney F100-PW-200 turbofan with 25000lb thrust.
Length	49'4" (15.03 m)
Height	16'8.5" (5.09 m)
Wingspan	31'("9.45 m)
Max T.O. Weight	23810lb (10800 kg)
Airspeed	> Mach 2 [max, level @40000']
Combat Radius	> 500 nm (925 km; 575 mi)
Service Ceiling	> 50000' (15200 m)
Crew	*Pilot Only* [F-16B *two*-seat version]
Tail Number	WA79-336
Flight Profile	380 kt, 15000'

F-4

The F-4 was originally developed as an attack fighter for the US Navy. The Air Force has used it for air defense, close air support, reconnaissance, and electronic countermeasures. Recordings [25] were made aboard an F-4E in January 1981. These recordings were made with microphones located both inside the pilot's oxygen mask and on the pilot's helmet.

The files **F4EH11** and **F4EM11** were digitized from field tapes 11-1 and 11-2, and represent simultaneous recordings inside and outside the oxygen mask at a time when the pilot was holding his breath. File **F4EHL** was digitized from field tape 11-1, during a high-speed low-altitude run (500 kt, 300' above ground level).

F-4E Phantom II	
Reference	Jane's 78-79 p. 373
Manufacturer	McDonnell Douglas
Primary Mission	All-Weather Fighter
Powerplant	Two General Electric J79-GE-17 Turbojet engines rated at 17900lb thrust each (afterburner at takeoff).
Length	63' (19.2 m)
Height	16'5.5" (5.02 m)
Wingspan	38'7.5" (11.77 m)
Combat Weight	41487lb (18818 kg)
Max T.O. Weight	51795lb (28030 kg)
Airspeed	Mach 2.24 [max, level]
Ferry Range	1718 nm (3184 km; 1978 mi)
Combat Radius	618 nm (1145 km; 712 mi) [interdiction]
Combat Ceiling	54400' (16600 m)
Crew	2
Tail Number	WA72-140
Flight Profile	350 kt, 7000'; 500 kt, 300' agl

A-10A

The A-10A is a heavily armored ground attack aircraft that can carry 5450 kg of externally-mounted munitions, in addition to its nose-mounted 30-mm Gatling gun. Recordings [25] were made aboard an A-10A in January 1981. These recordings were made with microphones located both inside the pilot's oxygen mask and on the pilot's helmet.

The files A10H25 and A10M24 were digitized from field tapes 8-1 and 8-2, and represent simultaneous recordings inside and outside the oxygen mask during a 2-second period when the pilot was holding his breath. At this time, the aircraft was at an altitude of 1000 ft and an airspeed of 340 kt. Less reliable files, A10H09 and A10MC8, were made during a 5-second period when the pilot was attempting to hold his breath but seemed to be making some sort of audible sound. At this time, the aircraft was at an altitude of 12000 ft and an airspeed of 260 kt.

A-10A Thunderbolt II ('Warthog')	
Reference	Jane's 81-82 p. 351
Manufacturer	Fairchild Republic Co.
Primary Mission	Sustained Close Air Support
Powerplant	Two General Electric TF34-GE-100 turbofan engines rated at 9065lb thrust each.
Length	53'4" (16.26 m)
Height	14'8" (4.47 m)
Wingspan	57'6" (17.53 m)
Operating Weight	25000lb empty (11320 kg)
Max T.O. Weight	50000lb (22680 kg)
Airspeed	381 kt (706 km/hr; 439 mph) [max, level @ S/L]
Combat Radius	252 nm (463 km; 288 mi) [close air support] 540 nm (1000 km; 620 mi) [deep strike]
Cruise Altitude	5000' (1500 m)
Crew	Pilot Only
Tail Number	WA168
Flight Profile	Low Altitude Interdiction Mission

Tornado

There are two variants of the European Tornado aircraft: the Interdictor Strike Aircraft (RAF GR1), and the Air Defense Variant (RAF F2). These aircraft are roughly comparable to the US F-15 in size and weight, and are widely used in NATO air forces. Noise recordings were made aboard a Tornado and provided to RADC/EEV in PCM form by the UK Royal Signals and Radar Establishment. Files **TOR13**, **TOR21**, and **TOR22** (all at 250 ft above ground level), and files **TOR33** and **TOR34** (both at 110 ft above ground level) were digitized from the tape segments designated TOR:13, TOR:21, TOR:22, TOR:33, and TOR:34.

RAF Tornado GR1 (IDS) or F2 (ADV)	
Reference	Jane's 88-89 p129
Manufacturer	Panavia
Primary Mission	IDS All-Weather Multipurpose Combat Aircraft
	ADV Air Defense Interceptor
Powerplant	Two Turbo-Union RB199-34R Mk103 Turbofan
	engines with 16800lb thrust each. [afterburner]
Length	IDS: 54'10" (16.72 m)
	ADV: 59'3" (18.06 m)
Height	19'6" (5.95 m)
Wingspan	45'7" (13.19 m) [variable geometry fully spread]
Max T.O. Weight	IDS w/externals: 60000lb (27215 kg)
	ADV: 61700lb (27986 kg)
Airspeed	> Mach 2.2 [IDS maximum, level]
Low Level Airspeed	1480 kt
Combat Radius	750 nm (1390 km; 863 mi) [IDS]
Intercept Radius	300 nm (556 km; 345 mi) [ADV supersonic]
Intercept Radius	1000 nm (1853 km; 1151 mi) [ADV subsonic]
Operating Ceiling	70000' (21300 m) [ADV]
Crew	2
Flight Profile	420-550 kt, 110'-250' agl

HH-53 helicopter

The HH-53 is a search and rescue helicopter with main and tail rotors powered by a turbine engine. A noise recording was made in 1984 [35] aboard an HH-53 helicopter in flight. The microphones were positioned at the rear bulkhead of the pilot's compartment.

File **HH53** was digitized from this tape, designated field tape 1AA.

HH-53[C] Super Jolly	
Reference	Jane's 74-75 p. 457
Manufacturer	Sikorsky S-65
Primary Mission	Heavy Assault Transport
	USAF HH-53 Search & Rescue
Powerplant	Two General Electric T64-GE-7 turboshaft engines rated at 3925 ehp each.
Length (fuselage)	67'2" (20.47 m)
Height (fuselage)	17'1.5" (5.22 m)
Main Rotor Diam.	72'3" (22.02 m)
Tail Rotor Diam.	16' (4.88 m)
Mission T.O. Weight	38238lb (17344 kg)
Airspeed	170 kt (315 km/hr; 196 mph) [max, level]
Airspeed (Cruise)	150 kt (278 km/hr; 173 mph)
Range	468 nm (869 km; 540 mi)
Service Ceiling	20400' (6200 m)
Crew	3 aircrew + 24 stretchers + 4 attendants
Flight Profile	100' agl

Appendix B

New Analyses of Long-Term Noise Characteristics

In an earlier report [35, pp. 38-61] we described long-term characteristics of noise aboard many of the aircraft included in the present study. Since that time, further noise recordings have been added to the RADC/EEV Acoustic Noise Database, involving aircraft not reported on in [35]. Except for one instance, these analyses are all based on the 1981 recordings made under the supervision of Miller *et al.* of Bolt Beranek and Newman Inc. and reported in [25]. The exception is the Tornado aircraft, for which we used recordings made by the UK Royal Signals and Radar Establishment. All these recordings are listed in Table 2.1, along with the other noise recordings used in this study.

B.1   The F-15A Fighter

In the earlier report [35] we included a discussion of long-term characteristics of F-15A noise, based on recordings made in 1976 by H. Hille of the USAF Armstrong Aeromedical Research Laboratory. Since that report, we have analyzed some of the 1981 F-15A noise recordings [25]. Aside from significant instrumentation and calibration differences, it should be noted that the 1976 recordings included several segments made during high-speed (supersonic) and high-power flight, and that none of the 1976 recordings were made with a microphone inside the oxygen mask.

In the case of outside-mask noise, analyses of the 1981 recordings are in general agreement with those reported previously. Both sources show a large amount of engine noise at frequencies above 5 kHz, outside the range normally

significant for narrowband voice processing. A strong engine-noise tone, usually
near 3 kHz, was found in some of the supersonic segments of the 1976 recordings,
but was not found in the later recordings. Matched recordings, made simultane-
ously inside and outside the oxygen mask while the pilot held his breath, were
analyzed. Power spectrum estimates, obtained by the averaged-periodogram
method with a Hamming window, and plotted on an arbitrary scale of 0 to 60
dB, are shown in Figures B.1 and B.2. The recording made outside the mask
shows a concentration of noise below 1 kHz, although, as we pointed out in [35],
the power spectral density above 700 Hz does not fall off as fast as does the
power spectral density of typical speech. The recording made inside the mask
shows the typical low-pass effect of the mask beginning at about 600 Hz, as
described in Section 2.4.

B.2   The F-16A Fighter

Matched recordings, made simultaneously inside and outside the oxygen mask
while the pilot held his breath, were analyzed. These recordings were made at
an altitude of 15000 ft and an indicated airspeed of 380 kt. Power spectrum
estimates, obtained by the averaged-periodogram method with a Hamming win-
dow and plotted on an arbitrary scale of 0 to 60 dB, are shown in Figures B.3
and B.4. In the recording made outside the mask, the highest noise levels ex-
tend almost to 2 kHz. The recording made inside the mask shows the typical
low-pass effect of the mask beginning at about 600 Hz, as described in Section
2.4.

B.3   The A-10 Ground Attack Plane

The Digital Noise Database includes a matched pair of recordings made during
a normal pause in the pilot's breathing. The aircraft was at an altitude of 1000
ft and an indicated airspeed of 340 kt. Power spectrum estimates, obtained by
the averaged-periodogram method with a Hamming window and plotted on an
arbitrary scale of 0 to 60 dB, are shown in Figures B.5 and B.6.

In the recordings made outside the mask, the bulk of the noise power is
concentrated below 400 Hz. The recordings made inside the mask show the
typical low-pass effect of the mask beginning at about 600 Hz, as described in
Section 2.4.

B.4   The F-4E Fighter

Matched recordings, made simultaneously inside and outside the oxygen mask
while the pilot held his breath, were analyzed. These recordings were made at
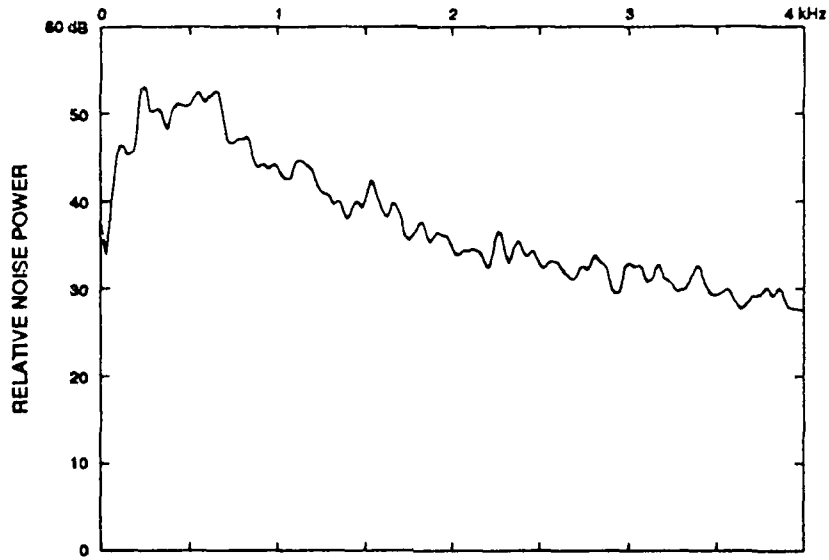an altitude of 15000 ft and an indicated airspeed of 380 kt. We have also added

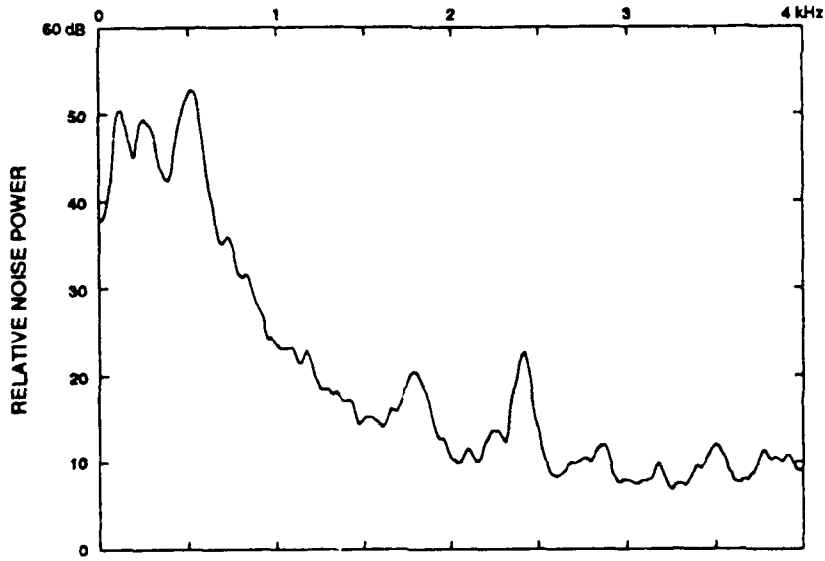Figure B.1: Power spectrum estimate for F-15A (outside mask) (**F15HT5**)

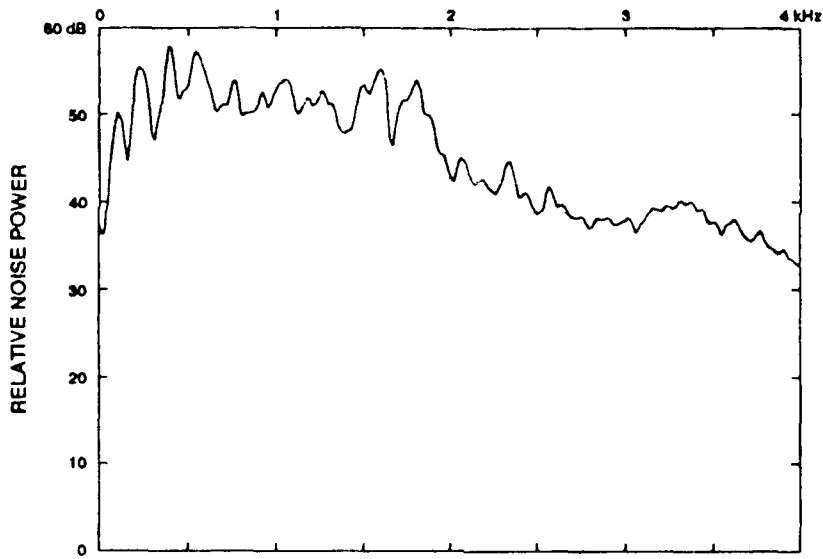Figure B.2: Power spectrum estimate for F-15A (inside mask) (**F15MT5**)

Figure B.3: Power spectrum estimate for F-16A (outside mask) **(F16H06)**
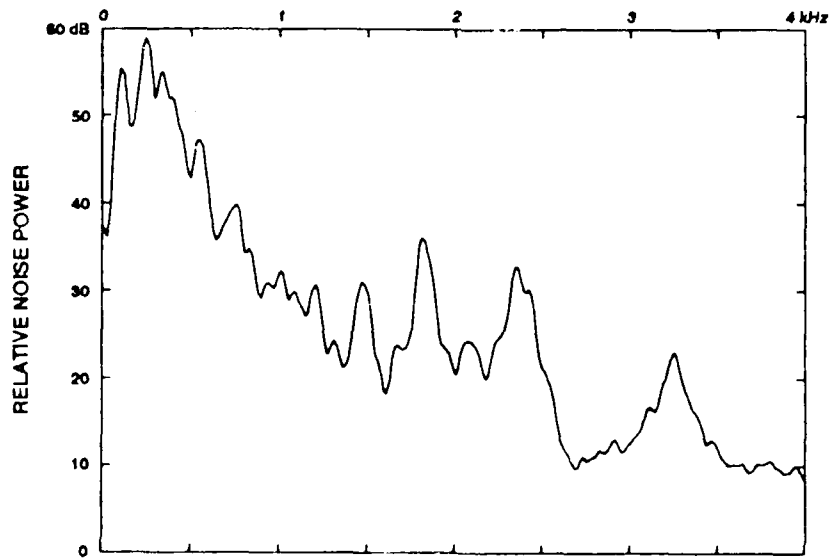

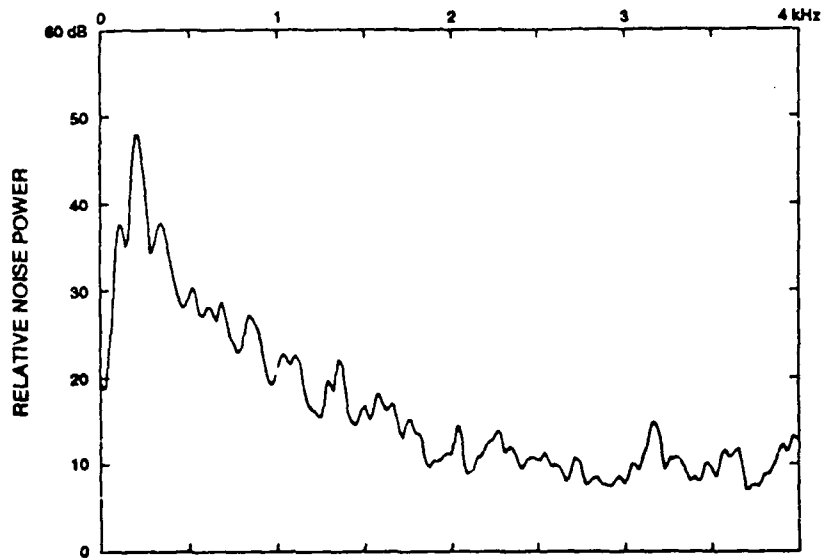Figure B.4: Power spectrum estimate for F-16A (inside mask) **(F16M06)**

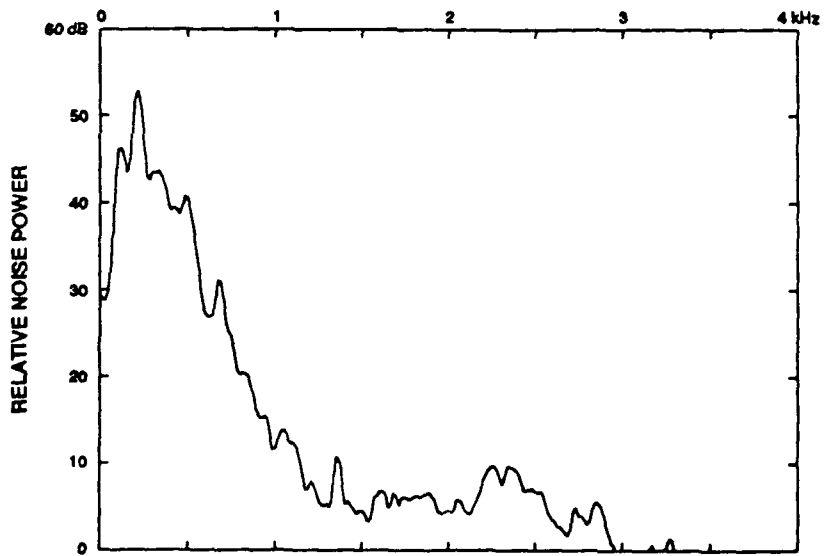Figure B.5: Power spectrum estimate for A-10 (outside mask) (**A10H25**)

Figure B.6: Power spectrum estimate for A-10 (inside mask) (**A10M24**)

one 5-second noise sample from the helmet microphone only, during a high-speed low-altitude run. Power spectrum estimates, obtained by the averaged-periodogram method with a Hamming window and plotted on an arbitrary scale of 0 to 60 dB, are shown in Figures B.7, B.8, and B.9. The recording made outside the mask shows a power spectral distribution quite similar to that observed in the F-15A recording (Figure B.1), except for a little less energy in the 500-700 Hz range. The recording made inside the mask shows a low-pass effect of the mask apparently beginning at about 400 Hz, a lower frequency than observed in other noise recordings.

B.5   The Tornado Fighter

We have added to the RADC/EEV Acoustic Noise Database several recordings made by the UK Royal Signals and Radar Establishment aboard a Tornado fighter aircraft. None of these recordings were made inside the helmet/mask units normally used aboard the aircraft. All recordings were made at an altitude of 250 ft above ground level.

Power spectrum estimates, obtained by the averaged-periodogram method with a Hamming window and plotted on an arbitrary scale of 0 to 60 dB, are shown in Figures B.10-B.13. Figures B.10, B.11, and B.12 were made at the pilot's position, while Figure B.13 was made at the navigator's position. The highest sound level, approximately 115 dB, was observed when the cabin demister was turned on (Figure B.12).

B.6   Comparisons

In the recordings made outside the oxygen masks, noise power spectra in the newly analyzed aircraft showed a general similarity to those measured earlier from the 1976 F-15A recordings. Inside the oxygen masks, there was an attenuation effect, described more fully in Section 2.4. The result of this attenuation was that, while the pilot was holding his breath, the in-mask noise power spectral density was somewhat similar *in shape* to what we have found in aircraft like the E-3A and E-4B, except for the resonance-like peaks we have commented on in Section 2.4.
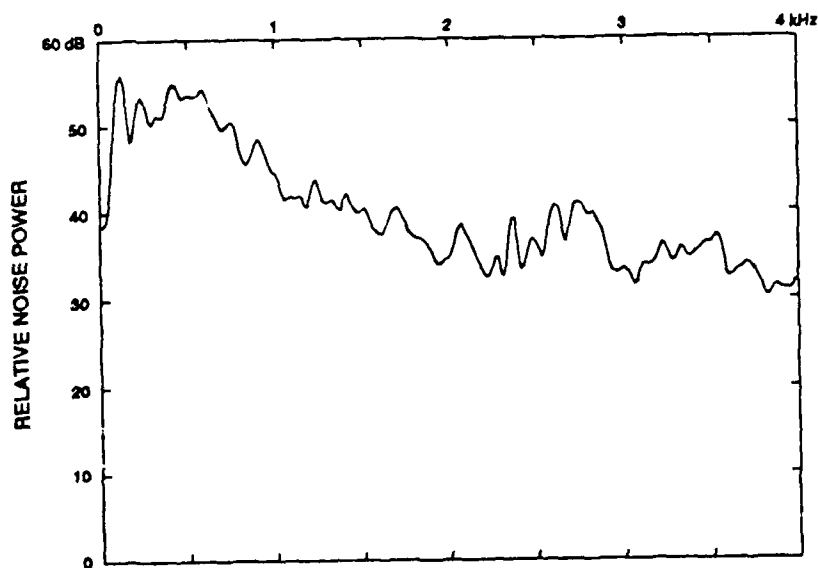
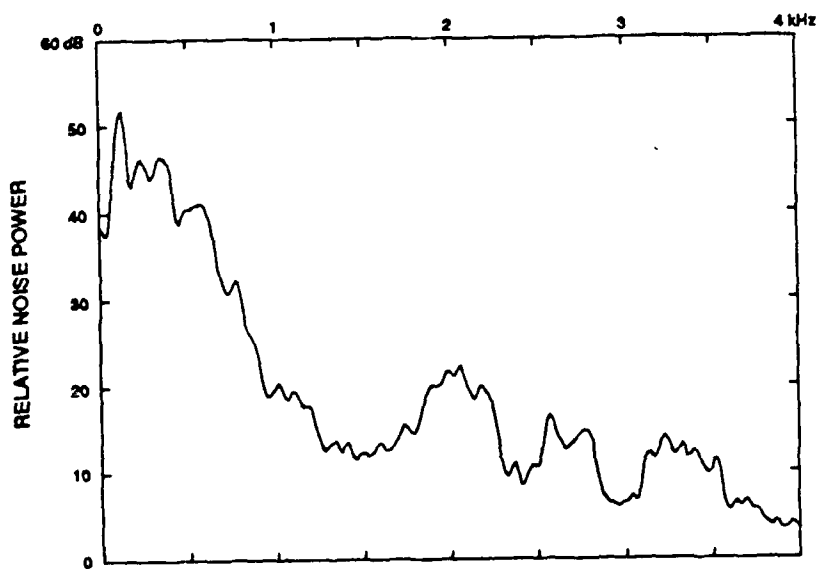rigure B.7: Power spectrum estimate for F-4E (outside mask) (F4EH11)

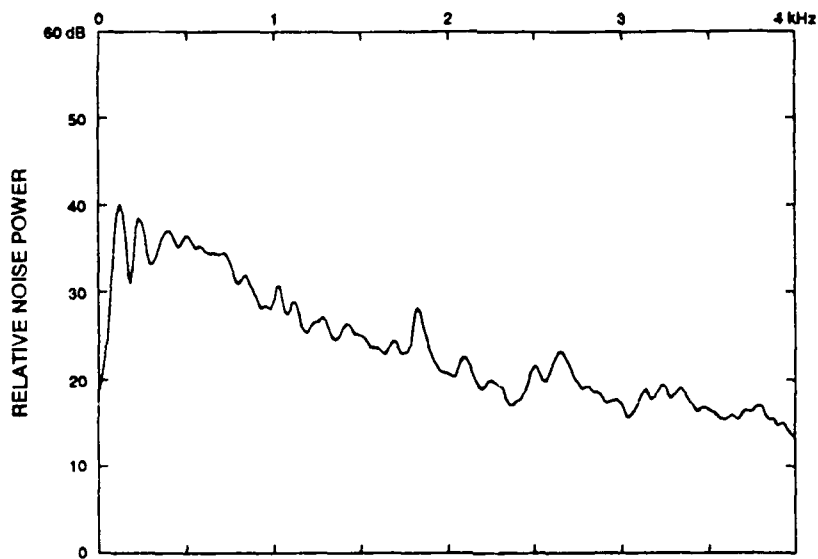Figure B.8: Power spectrum estimate for F-4E (inside mask) (F4EH11)

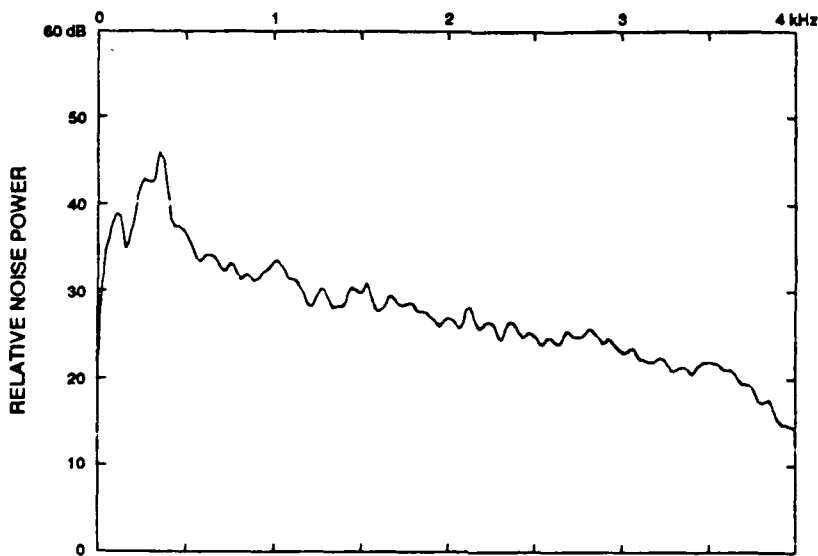Figure B.9: Power spectrum estimate for F-4E (outside mask, low altitude) (F4EHL)

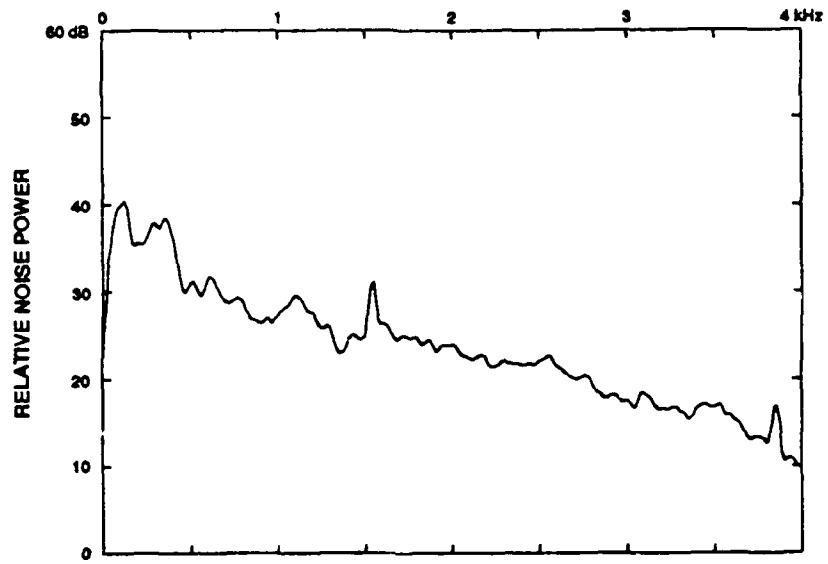Figure B.10: Power spectrum estimate for Tornado (pilot pos., 550 kt) (TOR13)

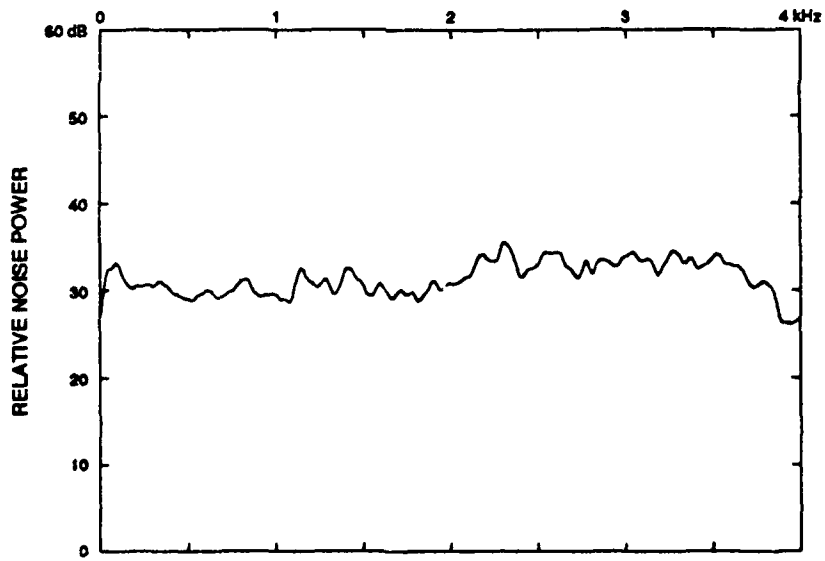Figure B.11: Power spectrum estimate for Tornado (pilot pos., 420 kt) (TOR21)

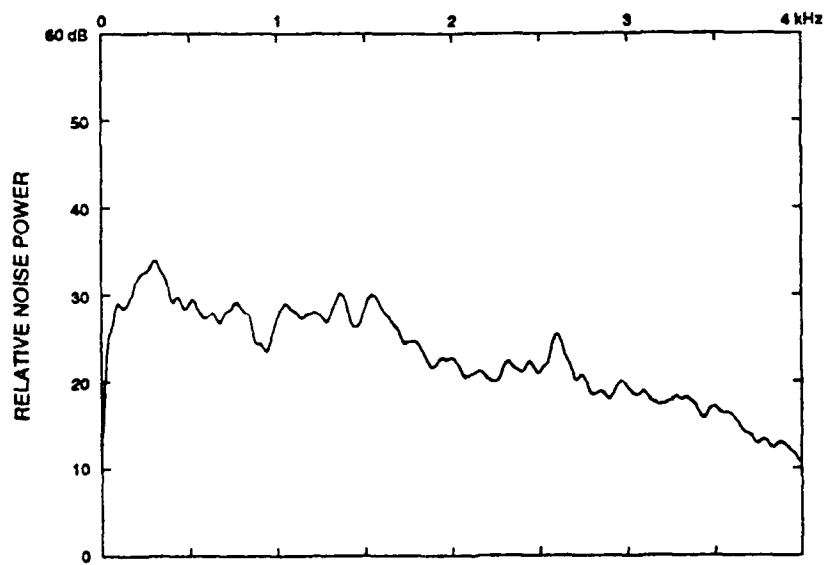Figure B.12: Power spectrum estimate for Tornado (pilot pos., 480 kt, demist on) (TOR34)

Figure B.13: Power spectrum estimate for Tornado (navigator pos., 420 kt) (TOR22)

Bibliography

[1] M. S. Ahmed, "Comparison of Noisy Speech Enhancement Algorithms in Terms of LPC Perturbation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-37, pp. 121-125 (1989).

[2] Jont B. Allen, "Short Term Spectral Analysis, Synthesis, and Modification by Discrete Fourier Transform," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-25, pp. 235-238 (1977).

[3] T. W. Anderson, *The Statistical Analysis of Time Series*, John Wiley & Sons, New York, 1971.

[4] E. Aschkenazy and M. R. Weiss, "Wideband Speech Enhancement Addition," USAF Rome Air Development Center Report RADC-TR-81-53, May 1981.

[5] E. Aschkenazy and M. R. Weiss, "Computerized Audio Proc∪ssor," USAF Rome Air Development Center Report RADC-TR-83-109, May 1983.

[6] E. Aschkenazy and M. R. Weiss, "Multichannel Advanced Speech Enhancement Development," USAF Rome Air Development Center Report RADC-TR-86-244, December 1986.

[7] M. Berouti, R. Schwartz, and J. Makhoul, "Enhancement of Speech Corrupted by Acoustic Noise," *Proc. IEEE ICASSP 1979*, pp. 208-211, IEEE Press, New York, 1979.

[8] S. F. Boll, "Suppression of Acoustic Noise in Speech Using Spectral Subtraction," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-27, pp. 113-120 (1979).

[9] J. N. Cole, *USAF Bioenvironmental Noise Data Handbook, Volume I: Organizaton, Content and Applications*, USAF AAMRL Technical Report AMRL-TR-75-50.

[10] P. Darlington, P. D. Wheeler, and G. A. Powell, "Adaptive Noise Reduction in Aircraft Communication Systems," *Proc. IEEE ICASSP 1985*, pp. 716-719, IEEE Press, New York, 1985.

[11] P. Darlington, "Static and Dynamic Noise Transmission Over an Oxygen Mask," *Proc. IEEE ICASSP 1988,* pp. 2308-2311, IEEE Press, New York, 1988.

[12] Y. Ephraim and D. Malah, "Speech Enhancement Using a Minimum Mean-Square Error Short-Time Spectral Amplitude Estimator," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-32, pp. 1109-1121 (1984).

[13] J. L. Flanagan, *Speech Analysis Synthesis and Perception,* Second Edition, Springer, New York, 1972.

[14] J. D. Gibbons, *Nonparametric Statistical Inference,* Second Edition, Marcel Dekker, New York, 1985.

[15] D. W. Griffin and J. S. Lim, "Signal Estimation from Modified Short-Time Fourier Transform," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-32, pp. 236-242 (1984).

[16] F. J. Harris, "On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform," *Proc. IEEE,* vol. 66, pp. 51-83 (1978).

[17] W. A. Harrison, J. S. Lim, and E. Singer, "A New Application of Adaptive Noise Cancellation," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-34, p. 21 (1986).

[18] G. M. Jenkins and D. G. Watts, *Spectral Analysis and Its Applications,* Holden-Day, San Francisco, 1968.

[19] G. S. Kang and S. S. Everett, "Improvement of the Narrowband Linear Predictive Vocoder, Part I- -Analysis Improvements," Naval Research Laboratory Report 8645, December 27, 1982.

[20] G. S. Kang and L. J. Fransen, "Quality Improvement of LPC-Processed Noisy Speech by Using Spectral Subtraction," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-37, pp. 939-942 (1989).

[21] S. M. Kay, *Modern Spectral Estimation,* Prentice Hall, Englewood Cliffs, NJ, 1988..

[22] L. H. Koopmans, *The Spectral Analysis of Time Series,* Academic Press, New York, 1974.

[23] J. S. Lim and A. V. Oppenheim, "Enhancement and Bandwidth Compression of Noisy Speech," *Proc. IEEE,* vol. 67, pp. 1586-1604 (1979).

[24] R. J. McAulay and M. L. Malpass, "Speech Enhancement Using a Soft-Decision Noise Suppression Filter," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-28, pp. 137-145 (1980).

[25] R. L. Miller, R. D. Bruce, F. N. Iacovino, and A. W. F. Huggins, "Simulation of Cockpit Noise Environments in Four Tactical Aircraft for the Purpose of Testing Speech Intelligibility," Bolt Beranek and Newman Inc. Report No. 4465, September 1981.

[26] G. M. Noether, *Elements of Nonparametric Statistics*, John Wiley & Sons, New York, 1967.

[27] J. E. Porter and S. F. Boll, "Optimal Estimators for Spectral Restoration of Noisy Speech," *Proc. IEEE ICASSP 1984*, Paper 18A.2, IEEE Press, New York, 1984.

[28] G. A. Powell, P. Darlington, and P. D. Wheeler, "Practical Adaptive Noise Reduction in the Aircraft Cockpit Environment," *Proc. IEEE ICASSP 1987*, pp. 173-176, IEEE Press, New York, 1987.

[29] L. R. Rabiner and R. W. Schafer, *Digital Processing of Speech Signals*, Prentice Hall, Englewood Cliffs, NJ, 1978.

[30] J. J. Rodriguez, "Adaptive Noise Reduction in Aircraft Communication Systems," Lincoln Laboratory Technical Report 756, 20 January 1987.

[31] J. J. Rodriguez and J. S. Lim, "Adaptive Noise Reduction in Aircraft Communication Systems," *Proc. IEEE ICASSP 1987*, pp. 6.1.1-6.1.4, IEEE Press, New York, 1987.

[32] G. F. Sandy and J. E. Parker, "Digital Voice Processor Consortium Final Report," Vol. II, Appendix A, MITRE Report MTR-84W00053-02, March 1984.

[33] E. Singer, "The Effects of Microphones and Facemasks on LPC Vocoder Performance," *Proc. IEEE ICASSP 1981*, pp. 1066-1069, IEEE Press, New York, 1981. (Also Lincoln Laboratory Technical Report 584, 21 September 1981.)

[34] C. P. Smith, "An Evaluation of Intelligibility Performance of the DoD Standard LPC-10 Narrowband Vocoder Algorithm in an E3A Aircraft Cabin Acoustic Noise Environment," USAF Rome Air Development Center Report RADC-TR-80-225, June 1980.

[35] J. D. Tardelli, C. M. Walter, J. T. Sims, P. A. La Follette, and P. D. Gatewood, "Research and Development for Digital Voice Processing," USAF Rome Air Development Center Report RADC-TR-86-171, October 1986.

[36] C. F. Teacher and H. F. Watkins, "ANDVT Microphone and Audio System Study," Ketron, Inc., Report No. 1159, August 1978.

[37] C. F. Teacher, "Secure Voice Evaluations: Characterization of Aircraft Noise and Audio Systems in the EC-135 and E-4B Aircraft," Ketron, Inc., Report No. 2709-2, 1983.

[38] M. R. Weiss and E. Aschkenazy, "The Speech Enhancement Advanced Development Model," USAF Rome Air Development Center Report RADC-TR-78-232, November 1978.

[39] "Digital Audio Processor PCM-F1 Operating Instructions," Sony Corporation, 1982.