AD-A063 817 AIR FORCE GEOPHYSICS LAB HANSCOM AFB MASS F/G 12/1

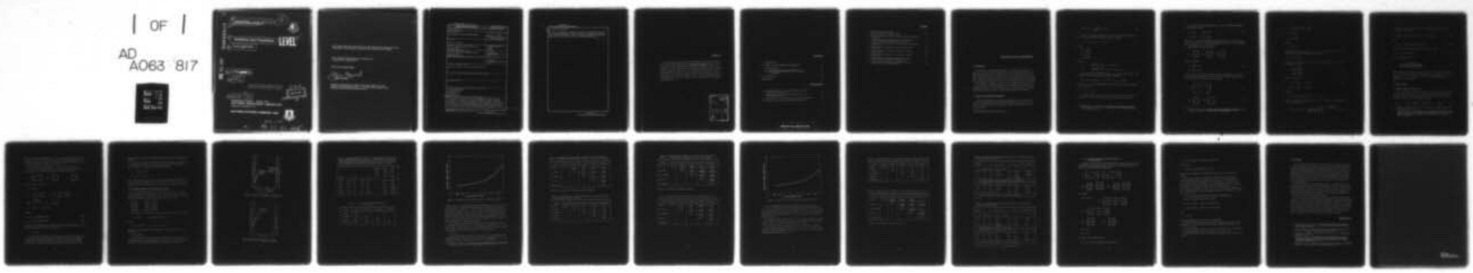
CONDITIONAL JOINT PROBABILITIES.(U)

OCT 78 I I GRINGORTEN

UNCLASSIFIED AFGL-TR-78-0238 NL

1 OF 1

AD A063 817



END

DATE FILMED

3-79

DDC

MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

AFGL-TR-78-0238 AFGL-ERP-643
ENVIRONMENTAL RESEARCH PAPERS, NO. 643

LEVEL II

Conditional Joint Probabilities.

IRVING I. GRINGORTEN

2 Oct [redacted] 78

25p.

6670 09

METEOROLOGY DIVISION PROJECT 6670
AIR FORCE GEOPHYSICS LABORATORY
HANSCOM AFB, MASSACHUSETTS 01731

AIR FORCE SYSTEMS COMMAND, USAF

409 578

This report has been reviewed by the ESD Information Office (OI) and is releasable to the National Technical Information Service (NTIS).

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

Chief Scientist

Qualified requestors may obtain additional copies from the Defense Documentation Center. All others should apply to the National Technical Information Service.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFGL-TR-78-0238	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) CONDITIONAL JOINT PROBABILITIES		5. TYPE OF REPORT & PERIOD COVERED Scientific. Interim.
		6. PERFORMING ORG. REPORT NUMBER ERP No. 643
7. AUTHOR(s) Irving I. Gringorten		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Air Force Geophysics Laboratory (LYD) Hanscom AFB Massachusetts 01731		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 62101F 66700904
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Geophysics Laboratory (LYD) Hanscom AFB Massachusetts 01731		12. REPORT DATE 2 October 1978
		13. NUMBER OF PAGES 25
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Conditional probability
Partial correlation
Joint probability
Predictand
Partial regression

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Formerly, for the solution of the conditional probability of a single predictand, its equivalent normal deviate (END) was obtained, under the assumption of multivariate normality, by linear regression on the END's of the predictors. For the joint probability of two predictands, the approach is to find the two corresponding END's by the same method, but in addition, to find the conditional correlation coefficient between the predictands. Such correlation has proved to be the well-known partial correlation. In a few test

→ next page

DD FORM 1473, 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

79 01 26 003

20. Abstract (Continued)

examples, the conditional correlation has decreased significantly from the more basic unconditional correlation. However, the conditional correlation has remained large enough to make the conditional probabilities significantly higher than the mere product of the two marginal probabilities.

Preface

This report logically follows previous investigations on conditional probability. The first report, appearing in the Monthly Weather Review (July 1972), was written for use with a single predictor; the second, appearing as AFSG No. 354, provided for several predictors that could contribute information on one predictand by linear regression. However, for two predictands occurring jointly, such as the ceiling and the visibility at a target, it does not suffice to produce individual predictions for the conditional probability of each. To meet the additional requirement, this report is based on in-house work, done at the request of Air Weather Service (RCS-7-11).

Contents

Illustrations

Tables

Conditional Joint Probabilities

1. INTRODUCTION

The problem of estimating the probability of concurrent events arises frequently. One of the most persistent problems is the finding of probability wherein both ceiling and visibility at a station jointly will have values above their thresholds. In this study, the same predictors will serve both weather elements, although prior ceilings should be better than prior visibility for the prediction of subsequent ceiling, and prior visibilities should serve more efficiently for subsequent visibility.

The conditional probability of two jointly occurring predicted events, given one or more predictors, is clearly dependent upon the conditional probability of each individual predictand. It is also dependent on the correlation between predictands. Generally, the latter correlation is conditioned upon selection of predictors.

2. FORMULA FOR CONDITIONAL PROBABILITY

Before proceeding further, one takes for granted that each variable has been transformed into the equivalent normal deviate (END) through its cumulative frequency. The variable X is transformed into the END, $x(0, 1)$, through the cumulative probability:

(Received for publication 29 September 1978)

$$P(\geq X) = P(\geq x) = \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} e^{-\xi^2/2}\, d\xi \qquad (1)$$

There is a convenient approximate formula[1] for determining the END (x) in terms of the probability of exceedance, $[P(\geq X)]$, as follows:

$$x = k[t - (a_o + a_1 t)/(1 + b_1 t + b_2 t^2)] \qquad (2)$$

where

$a_o = 2.30753$

$a_1 = 0.27061$

$b_1 = 0.99229$

$b_2 = 0.04481$

and where

$k = 1 \quad , \quad t = \sqrt{\ell n\ 1/p^2} \quad$ when $\quad p = P(\geq X) \leq 1/2$

$k = -1 \quad , \quad t = \sqrt{\ell n\ 1/(1 - p)^2} \quad$ when $\quad p = P(\geq X) > 1/2$

If y_1, y_2 are the predictand END's and $x_1, \cdots, x_n$ the predictor END's, then the problem becomes the finding of conditional probability:

$$P(y_1 \geq y_{1c},\ y_2 \geq y_{2c} | x_1, \cdots, x_n) = P(y_1 \geq y_{1c},\ y_2 \geq y_{2c} | \underline{X})$$

where y_{1c}, y_{2c} are the threshold END-values whose joint probability of exceedance is to be found, and

$$\underline{X} = (X_1, \cdots, X_n) \quad .$$

Correlation coefficients (CC's) are assumed to exist between each pair of variables, with each CC constant throughout the distribution of the variates.

---

1. National Bureau of Standards (1964) Handbook of Mathematical Functions, Applied Mathematics Series, 55, Government Printing Office, Washington, D.C. 20402, pp 932-933.

As in previous work,[2] each predictand is assumed to be linearly dependent on the predictors. Hence,

$$y_1 = a_{11}x_1 + \cdots + a_{1n}x_n + b_1\eta_1 \tag{3}$$

$$y_2 = a_{21}x_1 + \cdots + a_{2n}x_n + b_2\eta_2$$

where η_1, η_2 are normally distributed and random except for their intercorrelation. The a's are partial regression coefficients and the b's are of such magnitude that the normality of y_1, y_2 are preserved.

Introducing the vectors, we write

$$\underline{A}_1 = \begin{pmatrix} a_{11} \\ - \\ a_{1n} \end{pmatrix} \quad \underline{A}_2 = \begin{pmatrix} a_{21} \\ - \\ a_{2n} \end{pmatrix} \quad \underline{X} = \begin{pmatrix} x_1 \\ - \\ x_n \end{pmatrix}$$

Equation (3) becomes

$$y_1 = \underline{A}_1^T \underline{X} + b_1\eta_1 \tag{4}$$

$$y_2 = \underline{A}_2^T \underline{X} + b_2\eta_2$$

Let r_{ij} be the CC between the ith predictor and the jth predictor. Let ρ_{ij} be the CC between the ith predictand and the jth predictor. Then the coefficients (A_1, A_2, b_1, b_2) can be obtained in terms of the matrix:

$$\underline{C} = \begin{pmatrix} r_{11}, & r_{12} & \cdots, & r_{1n} \\ r_{21}, & r_{22} & \cdots, & r_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ r_{n1}, & r_{n2} & \cdots, & r_{nn} \end{pmatrix} \tag{5}$$

$$\underline{R}_1 = \begin{pmatrix} \rho_{11} \\ - \\ \rho_{1n} \end{pmatrix} \quad \underline{R}_2 = \begin{pmatrix} \rho_{21} \\ - \\ \rho_{2n} \end{pmatrix} \tag{6}$$

---

2. Gringorten, Irving I. (1976) Multi-Predictor Conditional Probabilities, AFSG No. 354, AFSC (AFGL), USAF, Hanscom AFB, 24 pp.

In terms of $\underline{C}$, $\underline{R}_1$, $\underline{R}_2$, we have

$$\underline{A}_1 = \underline{C}^{-1}\underline{R}_1 \tag{7}$$

$$\underline{A}_2 = \underline{C}^{-1}\underline{R}_2$$

where

$$\underline{C}^{-1} = \hat{\underline{C}}/|C| \tag{8}$$

where $\hat{\underline{C}}$ is the adjoint matrix of $\underline{C}$, and $|C|$ is the determinant of $\underline{C}$. The adjoint matrix of C is a matrix of cofactors (p_{ij}) as follows:

$$p_{ij} = (-1)^{i+j} \, |M_{ij}| \tag{9}$$

where M_{ij} is the submatrix of order (n-1) obtained by deleting the ith row and the jth column of $\underline{C}$. Also

$$b_1 = \sqrt{1 - \underline{A}_1^T \, \underline{R}_1} \tag{10}$$

$$b_2 = \sqrt{1 - \underline{A}_2^T \, \underline{R}_2}$$

Hence, from Eq. (4)

$$\eta_1 = (y_1 - a_{11}x_1 \cdots - a_{1n}x_n)/b_1 \tag{11}$$

$$\eta_2 = (y_2 - a_{21}x_1 \cdots - a_{2n}x_n)/b_2$$

Clearly,

$$P(y_1 \geq y_{1c}, \; y_2 \geq y_{2c} | \underline{X}) = P(\eta_1 \geq \eta_{1c}, \; \eta_2 \geq \eta_{2c}) \quad .$$

Since η_1, η_2 have Gaussian distributions, their joint probability is dependent upon the CC (ρ) between them in

$$P(\eta_1 \geq \eta_{1c}, \; \eta_2 \geq \eta_{2c}) = \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{\eta_{1c}}^{\infty} \int_{\eta_{2c}}^{\infty} e^{-\frac{\xi^2 - 2\rho\xi\eta + \eta^2}{2(1-\rho^2)}} \, d\xi \, d\eta \tag{12}$$

To find ρ, one needs to recognize it as the expected value of the product of η_1 and η_2. From Eq. (11):

$$b_1 b_2 \eta_1 \eta_2 = y_1 y_2 - \Sigma\, a_{1i} x_i y_2 - \Sigma\, a_{2i} x_i y_1 + \Sigma\Sigma\, a_{1i} x_i x_j a_{2j} \quad . \tag{13}$$

Hence

$$b_1 b_2 \rho = \rho_o - \underline{A}_1^T \underline{R}_2 - \underline{A}_2^T \underline{R}_1 + \underline{A}_1^T \underline{C}\underline{A}_2 \tag{14}$$

where ρ_o is the unconditional CC between y_1 and y_2. Finally,

$$\rho = (\rho_o - A_1^T CA_2)/b_1 b_2 \tag{15}$$

or, in purely algebraic terms

$$\rho = \frac{\rho_o - \sum_{i,j} a_{1i} a_{2j} r_{ij}}{\sqrt{1 - \sum_i a_{1i}\rho_{1i}}\ \sqrt{1 - \sum_i a_{2i}\rho_{2i}}} \tag{16}$$

Equation (15) or (16) can be recognized as the partial correlation between y_1 and y_2, a result that is a consequence of the absence of observational x-values in the computation of ρ.[3] In this report, therefore, the terms conditional CC and partial CC are interchangeable.

3. SAMPLE STUDIES

3.1 Ceiling and Visibility in Southern Germany

This case concerns probability at Nurnberg (49°30'N, 11°5'E), given observations at Kitzingen (49°42'N, 10°6'E). The distance between stations is 74.21 km. In Model B,[4] the CC's between ceiling or visibility at one station and ceiling or visibility at another station are given by the formula:

$$\rho(s') = \rho_o \cdot \frac{2}{\pi}\left[\sin^{-1}\sqrt{1-\sigma^2} - \sigma\sqrt{1-\sigma^2}\right] \quad , \qquad \sigma = \frac{s'}{128r} \tag{17}$$

3. Lawrance, A.I. (1976) On conditional and partial correlation, The American Statistician 30:146-149.

4. Gringorten, Irving I. (1978) Modelling for the Probability Distribution of Ceiling and Visibility at a Point, Along a Line of Travel, or in an Area, in the Southern Region of West Germany, Unpublished Report, submitted to Hq. AWS, 26 April 1978.

where s' is the distance between stations, r is the parameter known as the scale distance, and ρ_o is a correction factor, mainly due to observational errors, equal to 0.85 for ceiling versus ceiling, and 0.85 for visibility versus visibility. For ceiling versus visibility, the value of ρ_o is 0.17.

For the German mountain stations, r = 6.24 km in winter. Hence, for s' = 74.21 km, Eq. (17) gives the January CC's that compose the matrix C and vectors R_H and R_V for ceiling height (H) and visibility (V) as follows:

$$\underline{C} = \begin{pmatrix} 0.85 & 0.17 \\ 0.17 & 0.85 \end{pmatrix} \qquad \underline{R}_H = \begin{pmatrix} 0.75 \\ 0.15 \end{pmatrix} \qquad \underline{R}_V = \begin{pmatrix} 0.15 \\ 0.75 \end{pmatrix}$$

From Eqs. (7) to (11), we find

$$|C| = 0.9711$$

$$\hat{C} = \begin{pmatrix} 0.85 & -0.17 \\ -0.17 & 0.85 \end{pmatrix} \qquad C^{-1} = \begin{pmatrix} 1.2255 & -0.2451 \\ -0.2451 & 1.2255 \end{pmatrix}$$

$$\underline{A}_H = \begin{pmatrix} 0.8824 \\ 0 \end{pmatrix} \qquad \underline{A}_V = \begin{pmatrix} 0 \\ 0.8824 \end{pmatrix}$$

$$b_H = 0.582 \quad , \quad b_V = 0.582$$

Finally,

$$\eta_H = (y_H - 0.8824\, x_H)/0.582 \tag{18a}$$

$$\eta_V = (y_V - 0.8824\, x_V)/0.582 \tag{18b}$$

Fortuitously, ceiling depends on ceiling predictor only; visibility depends on visibility predictor only. From Eq. (15) or Eq. (16)

$$\rho = 0.118$$

To find the joint probability, Eqs. (18a and b) must be used; the values of x_H and x_V corresponding to the observations of ceiling and visibility at Kitzingen must be known, and the threshold values of y_H and y_V at Nurnberg be selected. For this report, the END's are found through the cumulative frequencies of the ceilings and the visibilities at the two stations, Kitzingen and Nurnberg (Figures 1 and 2).

Suppose that the ceiling and visibility at the predictor station, Kitzingen, are known to be 300 meters and 1 km, respectively. Then, from Figures 1 and 2, $x_H = -1.00$ and $x_V = -1.60$. Therefore Eq. (18) becomes

$$\eta_H = 1.719\, y_H + 1.339$$

$$\eta_V = 1.719\, y_V + 2.427$$

Table 1 gives values of y_H and y_V for several threshold values of ceiling and visibility at Nurnberg, the predictand station. The entries for joint conditional probability were found by numerical solution of Eq. (12) on a Hewlett-Packard 9810A.

3.2 Joint Conditional Probability of 24-hour Rainfall in New England

The selected example is that of the joint probability of 24-hour rainfall at Boston, Massachusetts, and Portland, Maine. The question concerns whether or not there has been rain, and how much, at the more westerly stations: Providence, Hartford, and Burlington. The five stations are located as follows:

Boston	$42^{o}22'$N	$71^{o}02'$W
Portland	$43^{o}39'$N	$70^{o}19'$W
Providence	$41^{o}44'$N	$71^{o}26'$W
Hartford	$41^{o}56'$N	$72^{o}41'$W
Burlington	$44^{o}28'$N	$73^{o}09'$W

From the formula for distance s' (kilometers) between the ith and jth stations:

$$s' = 111.12\, b \qquad \text{(b in degrees and tenths)}$$

where

$$\cos b = \sin \phi_i \sin \phi_j + \cos \phi_i \cos \phi_j \cos (\lambda_i - \lambda_j)$$

and where ϕ represents latitude and λ longitude. For distances between stations, see Table 2.

3.2.1 JANUARY

Previous records gave the probability distribution of 24-hour precipitation in New England (January), producing the curve of rainfall versus END (Figure 3). The 26-year record (1952-1977) yielded an average frequency of 0.3801 for

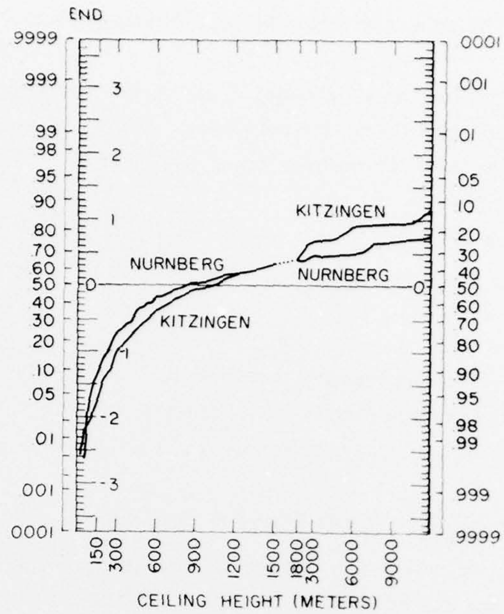


Figure 1. Cumulative Frequency of Ceiling Height (January), 0900 L, at Kitzingen and Nurnberg

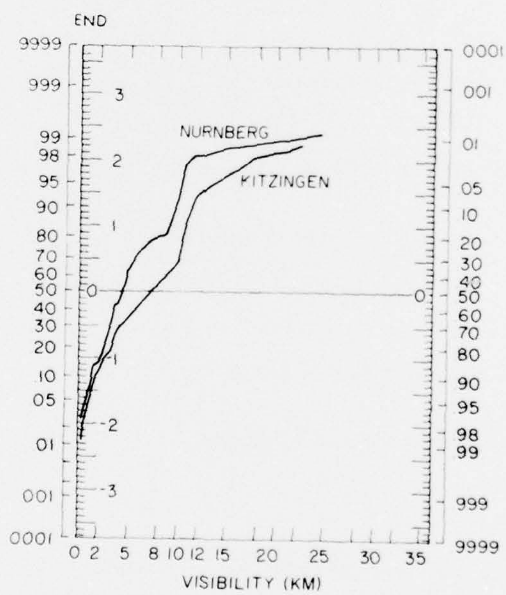


Figure 2. Cumulative Frequency of Visibility (January), 0900 L, at Kitzingen and Nurnberg

Table 1. Threshold Ceiling and Visibility. Corresponding values (y_H, y_V), unconditional probabilities [$P(\geq y_H)$, $P(\geq y_V)$], resulting values (η_H, η_V) for the specific predictors at Kitzingen (ceiling 300 m, visibility 1 km), and the conditional probabilities [$P(\geq \eta_H)$, $P(\geq \eta_V)$] (Estimates are shown of the joint conditional probabilities when the conditional correlation coefficient is 0.118)

					Visibility	≥1 km	≥5 km
					y_V	-1.50	0.31
					$P(\geq y_V)$	0.933	0.378
					η_V	-0.15	2.96
Ceiling Ht	y_H	$P(\geq y_H)$	η_H	$P(\geq \eta_H)$	$P(\geq \eta_V)$	0.560	0.0015
≥150 m	-1.40	0.92	-1.07	0.86	...	0.489	0.0015
≥300 m	-0.72	0.76	0.10	0.46	...	0.275	0.0010
≥600 m	-0.18	0.57	1.03	0.15	...	0.095	0.00041
≥1500 m	0.35	0.36	1.94	0.026	...	0.017	0.000084
≥9000 m	0.94	0.17	2.95	0.0016	...	0.0012	0.000007

Table 2. Distances Between New England Stations (in kilometers)

Station	Boston	Portland	Providence	Hartford
Portland	154.1	...	...	...
Providence	77.8	231.7	...	...
Hartford	144.3	271.3	105.9	...
Burlington	289.1	243.7	334.1	284.1

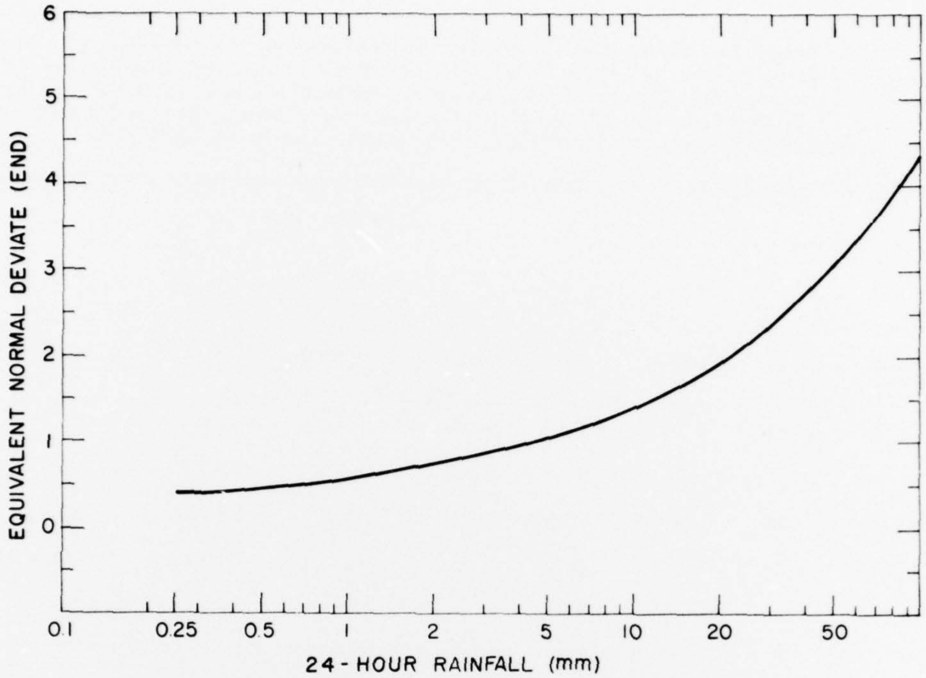


Figure 3. Transformation of New England 24-hour Precipitation (January)

rainfall ≥0.25 mm. The frequencies and joint frequencies of rain ≥0.25 mm were as shown (Table 3). The distances between stations, coupled with a previously determined[5] scale distance of r = 9.45 km for 24-hour precipitation gave CC's between stations (Table 4). These, in turn, together with the marginal frequencies (Table 3), as used in Eq. (12), gave the estimates of joint probabilities. On the whole, the estimates are low by an average of 0.02, suggesting that the parameter size (r = 9.45 km) could have been chosen somewhat larger, at least for winter precipitation equal to or greater than 0.25 mm.

For precipitation ≥21.84 mm, the marginal and joint frequencies were as shown (Table 5). If, again, the CC's, as estimated from Model B with r = 9.45 km, are used in Eq. (12) the estimates of joint probability become the figures in parentheses (Table 5). This time the average bias is less than 0.001.

3.2.2 JULY

Previous records gave the probability distribution of 24-hour rainfall in July as shown (Figure 4). The average frequency is 0.305 of rainfall ≥0.25 mm. The marginal and joint frequencies ≥0.25 mm were as shown (Table 6).

5. Gringorten, Irving I. (1976) Areal Coverage Estimates by Stochastic Modelling, ERP No. 573, AFSC (AFGL), USAF, Hanscom AFB, 56 pp.

Table 3. Marginal and Joint Frequencies of 24-hour January Precipitation ≥ 0.25 mm.* (Marginal or single-station frequencies are shown on the diagonal)

Station	Boston	Portland	Providence	Hartford	Burlington
Boston	0.385	0.300 (0.28)	0.328 (0.31)	0.318 (0.29)	0.284 (0.28)
Portland	...	0.354	0.279 (0.25)	0.275 (0.24)	0.284 (0.27)
Providence	...	...	0.367	0.315 (0.29)	0.259 (0.26)
Hartford	...	...	...	0.364	0.273 (0.27)
Burlington	...	...	...	...	0.438

*Numbers in parentheses are Model B probability estimates.

Table 4. Model B Estimates of Correlation Coefficients of 24-hour January Precipitation Between Pairs of Stations

Station	Boston	Portland	Providence	Hartford	Burlington
Boston	1.000	0.838	0.918	0.848	0.698
Portland	...	1.000	0.758	0.717	0.745
Providence	...	...	1.000	0.889	0.652
Hartford	...	...	...	1.000	0.704
Burlington	...	...	...	...	1.000

Table 5. Marginal and Joint Frequencies of 24-hour January Precipitation ≥21.84 mm.* (Single-station frequencies are shown on the diagonal)

Station	Boston	Portland	Providence	Hartford	Burlington
Boston	0.0347	0.0099 (0.011)	0.0236 (0.020)	0.0136 (0.011)	0.00124 (0.0020)
Portland	...	0.0298	0.0136 (0.012)	0.00868 (0.0068)	0.00124 (0.00248)
Providence	...	...	0.0422	0.0136 (0.014)	0.00248 (0.0019)
Hartford	...	...	...	0.0223	0.00248 (0.0016)
Burlington	...	...	...	...	0.0037

*Numbers in parentheses are Model B estimates.

Table 6. Marginal and Joint Frequencies of 24-hour July Rainfall ≥0.25 mm

Station	Boston	Portland	Providence	Hartford	Burlington
Boston	0.292	0.196 (0.19)	0.208 (0.23)	0.220 (0.20)	0.177 (0.18)
Portland	...	0.303	0.173 (0.17)	0.199 (0.16)	0.213 (0.19)
Providence	...	...	0.324	0.205 (0.23)	0.159 (0.18)
Hartford	...	...	...	0.324	0.197 (0.19)
Burlington	...	...	...	...	0.391

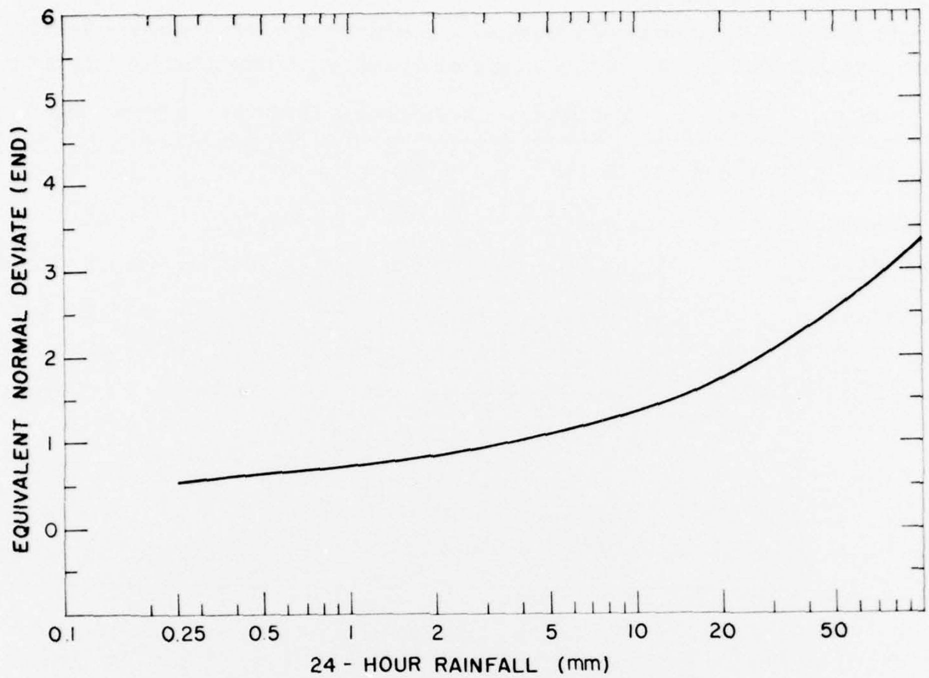


Figure 4. Transformation of New England 24-hour Rainfall (July)

The previously determined[5] scale distance for July 24-hour rainfall is r = 5.19 km. Coupled with the distances between stations as determined above, the CC's between stations are as estimated (Table 7). The CC's, in turn, together with the marginal frequencies in Table 6, when used in Eq. (12), yielded the joint probabilities in parentheses (Table 6). The latter show a bias of only 0.003 or 3/10 of 1 percent.

For rainfall ≥27.94 mm (END ≥2.0), the marginal and joint frequencies were as shown (Table 8). The Model B estimates of joint probability are the bracketed figures. They show a bias of -0.001, and an RMSE of 0.001.

This supports our reliance upon Model B. The scale distances of Model B were used to estimate the CC's that were used in Eqs. (4) to (15) in order to find the partial regression coefficients and other parameters (b_1, - b_2), and finally the conditional probabilities (Section 3.2.3).

Table 7. Model B Estimates of Correlation Coefficients of 24-hour July Rainfall

Station	Boston	Portland	Providence	Hartford	Burlington
Boston	1.000	0.707	0.851	0.725	0.464
Portland	...	1.000	0.565	0.494	0.543
Providence	...	...	1.000	0.798	0.387
Hartford	...	...	...	1.000	0.472
Burlington	...	...	...	...	1.000

Table 8. Marginal and Joint Frequencies of 24-hour July Rainfall ≥27.94 mm*

Station	Boston	Portland	Providence	Hartford	Burlington
Boston	0.0174	0.000775 (0.0031)	0.00574 (0.0064)	0.00326 (0.0047)	... (0.0015)
Portland	...	0.0149	0.00078 (0.0027)	0.00326 (0.0023)	0.000775 (0.0017)
Providence	...	...	0.0223	0.00574 (0.0069)	... (0.0014)
Hartford	...	...	...	0.0236	0.00202 (0.0019)
Burlington	....	...	...	...	0.0136

*Numbers in parentheses are Model B estimates.

Table 9. Marginal and Joint Conditional Probabilities of 24-hour Precipitation at Boston and Portland (January)

Amount (mm)	y_T	η_B	$P(y_B \geq y_T \mid X)$	η_P	$P(y_P \geq y_T \mid X)$	$P(\text{Joint} \mid X)$
On 1 Jan 1952 $x_1 = x_2 = x_3 = -0.44$*						
≥0.25	0.4	2.23	0.013	1.52	0.064	0.0039
≥4.2	1.0	3.84	0.000062	2.59	0.0048	0.0000083
≥21.8	2.0	6.52	...	4.37	0.000006	...
Actual precipitation: None at Boston, 0.25 mm at Portland						
On 3 Jan 1953, $x_1 = 2.24$ $x_2 = 1.97$, $x_3 = 1.51$*						
≥0.25	0.4	-4.46	0.9999959	-2.79	0.9974	0.9974
≥4.2	1.0	-2.85	0.9978	-1.72	0.957	0.955
≥21.8	2.0	-0.17	0.57	0.07	0.47	0.342
Actual precipitation: 30.5 mm at Boston and at Portland						

*See text

Table 10. Marginal and Joint Conditional Probabilities of 24-hour Rainfall at Boston and Portland (July)

Amount (mm)	y_T	η_B	$P(y_B \geq y_T \mid X)$	η_P	$P(y_P \geq Y_T \mid X)$	$P(\text{Joint} \mid X)$
On 1 July 1952 $x_1 = x_2 = x_3 = -0.40$*						
≥0.25	0.51	1.53	0.063	1.10	0.14	0.026
≥3.6	1.0	2.51	0.006	1.76	0.039	0.001
≥27.9	2.0	4.50	0.000003	3.10	0.0010	negl
Actual rainfall: None at Boston or Portland						
On 2 July 1953, $x_1 = 0.70$, $x_2 = 0.81$, $x_3 = 1.09$*						
≥0.25	0.51	-0.439	0.68	-0.26	0.60	0.47
≥3.6	1.0	0.535	0.30	0.40	0.34	0.16
≥27.9	2.0	2.52	0.006	1.74	0.041	0.0015
Actual rainfall: None at Boston, 1.0 mm at Portland						

*For details on the values of x_1, x_2, x_3, see discussion in text.

3.2.3 JANUARY PARAMETERS AND CONDITIONAL PROBABILITIES

Using B, P as subscripts for Boston and Portland, respectively, and 1, 2, 3 for Providence, Hartford, and Burlington respectively, one finds that the solutions of Eqs. (5) to (16) produce the following results:

$$\underline{C} = \begin{pmatrix} 1 & r_{12} & r_{13} \\ r_{21} & 1 & r_{23} \\ r_{31} & r_{32} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0.889 & 0.652 \\ 0.889 & 1 & 0.704 \\ 0.652 & 0.704 & 1 \end{pmatrix}$$

$$\underline{R}_B = \begin{pmatrix} \rho_{B1} \\ \rho_{B2} \\ \rho_{B3} \end{pmatrix} = \begin{pmatrix} 0.918 \\ 0.848 \\ 0.698 \end{pmatrix} \qquad \underline{R}_P = \begin{pmatrix} \rho_{P1} \\ \rho_{P2} \\ \rho_{P3} \end{pmatrix} = \begin{pmatrix} 0.758 \\ 0.717 \\ 0.745 \end{pmatrix}$$

from which

$$|C| = 0.1051 \qquad \hat{\underline{C}} = \begin{pmatrix} 0.5044 & -0.4300 & -0.0261 \\ -0.4300 & 0.5749 & -0.1244 \\ -0.0261 & -0.1244 & 0.2097 \end{pmatrix}$$

$$\underline{C}^{-1} = \begin{pmatrix} 4.799 & -4.091 & -0.248 \\ -4.091 & 5.470 & -1.184 \\ -0.248 & -1.184 & 1.995 \end{pmatrix}$$

$$\underline{A}_B = \begin{pmatrix} 0.763 \\ 0.0566 \\ 0.161 \end{pmatrix} \qquad \underline{A}_P = \begin{pmatrix} 0.520 \\ -0.0611 \\ 0.449 \end{pmatrix}$$

$$b_B = 0.373$$

$$b_P = 0.561$$

From Eqs. (11) the END's become

$$\eta_B = (y_B - 0.763x_1 - 0.0566x_2 - 0.161x_3)/0.373$$

$$\eta_P = (y_P - 0.520x_1 - 0.0611x_2 - 0.449x_3)/0.561$$

From Eq. (16)

$$\rho = (0.838 - 0.739)/(0.373)(0.561) = 0.473$$

which makes the conditional CC substantially reduced from the unconditional of 0.838.

When η_B, η_P are found, from the above equations, for several END's of threshold rainfall (y_B, y_P), for specific END's of predictor information (x_1, x_2, x_3), the results are as shown (Table 9). On 1 January 1952, there was no rain at Providence, Hartford or Burlington. Since the probability of no rain at a single station is 0.66, the END was chosen for half this probability (that is, $x_1 = x_2 = x_3 = -0.44$). On 3 January 1953, there was 27.9 mm of precipitation at Providence, 24.1 mm at Hartford, 11.9 mm at Burlington (which Figure 3 converts to $x_1 = 2.24$, $x_2 = 1.97$, $x_3 = 1.51$).

3.2.4 JULY PARAMETERS AND CONDITIONAL PROBABILITIES

In similar procedure, the calculations give:

$$\eta_B = (y_B - 0.746\, x_1 - 0.060\, x_2 - 0.147\, x_3)/0.503$$

$$\eta_P = (y_P - 0.459\, x_1 + 0.058\, x_2 - 0.393\, x_3)/0.746$$

and

$$\rho = 0.469$$

Again the conditional CC is reduced but still significant.

On 1 July 1952, there was no rain at the three predictor stations. Since the probability of no rain is 0.69, the END is taken at one half the probability. Thus $x_1 = x_2 = x_3 = -0.40$.

On 2 July 1953 there were 0.76 mm of rain at Providence, 1.5 mm at Hartford, and 5.1 mm at Burlington. Hence $x_1 = 0.70$, $x_2 = 0.81$, $x_3 = 1.09$.

Table 10 shows the results on the conditional probabilities, similar to those in Table 9.

4. CONCLUSIONS

The conditional joint probability of two events that are conditional upon one or more pieces of predictor information can be related to the conditional probabilities of each event separately, given the conditional CC between them. To find the individual conditional probability of each predictand, the corresponding equivalent normal deviate is determined, using an equation involving the deviates of predictand and predictors. The conditional CC in this work is synonymous with partial correlation coefficient; it is found in terms of the unconditional CC corrected by terms involving CC's between predictand and predictors, and also between predictors and predictors.

While the conditional CC's have proved smaller in the examples cited, it is conceivable that they could have remained unchanged. Consider the first example, ceiling and visibility in the mountain regions of Southern Germany. The effect on the joint conditional probability has been that of estimating only a slight improvement in the joint probability over the product of the two separate probabilities. These estimates cannot be verified by sampling.

In the case of New England rainfall, the viability of Model B was first established by its estimates of joint unconditional probabilities. The Model B estimates of the CC's produced substantially reduced conditional CC's between the predictands. The resulting estimates of joint conditional probability, however, in some instances were substantially higher than the product of the separate conditional probabilities. The most significant result can be seen in the wide variations in estimates of conditional probabilities both separate and joint, depending on the predictor values; that is, depending on what the observers see at the predictor stations.

References

1. National Bureau of Standards (1964) Handbook of Mathematical Functions, Applied Mathematics Series, 55, Government Printing Office, Washington, D.C. 20402, pp 932-933.

2. Gringorten, Irving I. (1976) Multi-Predictor Conditional Probabilities, AFSG No. 354, AFSC (AFGL), USAF, Hanscom AFB, 24 pp.

3. Lawrance, A.I. (1976) On conditional and partial correlation, The American Statistician **30**:146-149.

4. Gringorten, Irving I. (1978) Modelling for the Probability Distribution of Ceiling and Visibility at a Point, Along a Line of Travel, or in an Area, in the Southern Region of West Germany, Unpublished Report, submitted to Hq. AWS, 26 April 1978.

5. Gringorten, Irving I. (1976) Areal Coverage Estimates by Stochastic Modelling, ERP No. 573, AFSC (AFGL), USAF, Hanscom AFB, 56 pp.

Printed by
United States Air Force
Hanscom AFB, Mass. 01731