UCID- 17243

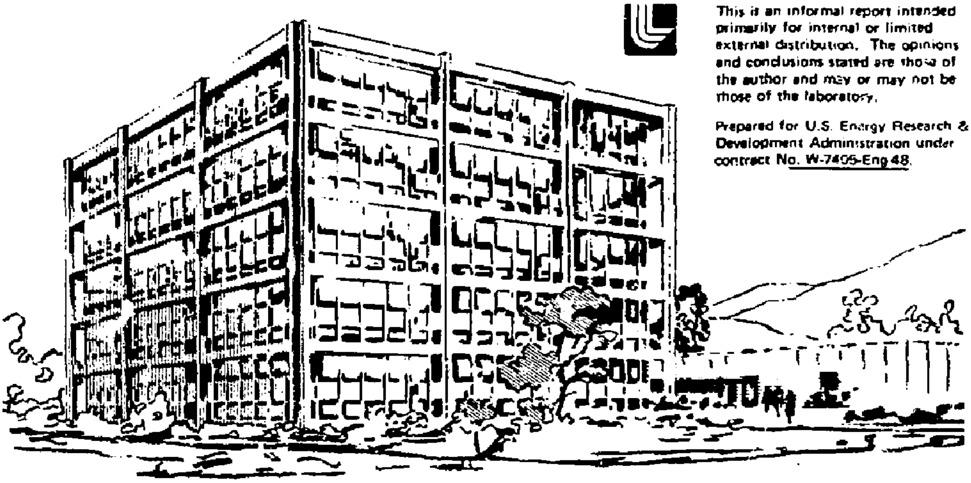
Lawrence Livermore Laboratory

STATISTICAL COMPARISONS OF SAVANNAH RIVER ANEMOMETER DATA APPLIED TO QUALITY CONTROL OF INSTRUMENT NETWORKS

William ... Porch and Marvin H. Dickerson

August 1976

MASTER



Prepared for U.S. Energy Research & Development Administration under contract No. W-7405-Eng-48.

MASTER

UCID-17243

STATISTICAL COMPARISONS OF SAVANNAH RIVER ANEMOMETER DATA APPLIED TO QUALITY CONTROL OF INSTRUMENT NETWORKS*

William M. Porch and Marvin H. Dickerson
Lawrence Livermore Laboratory
Livermore, California 94550

August 1976

INTRODUCTION

Continuous monitoring of extensive meteorological instrument arrays is a requirement in the study of important mesoscale atmospheric phenomena. The phenomena include pollution transport prediction from continuous area sources,[1] or one time releases of toxic materials[2] and wind energy prospecting in areas of topographic enhancement of the wind.[3] This paper will be oriented, however, toward our most immediate need which is associated with the Atmospheric Release Advisory Capability (ARAC).

Within the next several months the ARAC Central Facility (ACF) will begin to collect meteorological data from three sites on a routine basis.[2] During the past year we have been concerned with investigating various data quality control techniques that can be applied to these data to determine if the instruments are operating within their prescribed tolerances. Obviously if the instrument is "frozen" a rather simple check would be sufficient; however, it is those cases where the instrument gradually looses its

*Work performed under the auspices of the U.S. Energy Research & Development Administration under contract No. W-7405-Eng-48.

OT W-7405-ENG-26

calibration that is of greatest concern. We want to analyze the data to determine if the instrument should be checked prior to its normal scheduled maintenance. The method of analysis must be suitable to automation since the number of instruments on line to the ACF in the next few years will preclude other methods. Furthermore, the types of errors that concern us are not easily detected without the use of an analytical approach. It is our goal that the method of analysis can be performed by the ARAC Site Facility (ASF) and/or the ACF.

The problems which can develop with anemometric data systems include the following:

1. Wearing of bearings caused by dust or cold weather effects.[4]
2. Data transmission and recording errors.
3. Imperfect instrument exposure.[5,6]
4. Weather damage caused by lightning, hail and severe winds.
5. Misalignment of the instrument[7] and improper calibration.

Each of these problems has associated with it data effects which range from simply having no data to very subtle frequency response changes.

In order to investigate several different approaches to these problems, we have used data collected from the Savannah River Plant (SRP) during one of the real time ARAC exercises. Although these data are of limited extent, (15 minute averages for 11 hours) they are sufficient for preliminary studies designed to investigate technique development and comparative computer storage and calculational time requirements. At the time the ACF begins to collect meteorological data from one or more sites we plan to extend our studies to longer period data records using the quality control analysis methods that show promise.

We have analyzed the Savannah River data with both independent and comparitive statistical techniques. The independent techniques calculate the mean, standard deviation, moments about the mean, kurtosis, skewness, probability density distribution, cumulative probability and power spectra. The comparitive techniques include covarience, cross-spectral analysis and two dimensional probability density. At present the calculating and plotting routines for these statistical techniques do not reside in a single code so it is difficult to ascribe independent memory size and computation time accurately. However, given the flexibility of a data system which includes simple and fast running statistics at the instrument end of the data network (ASF) and more sophisticated techniques at the computational end (ACF) a proper balance will be attained. In the sections to follow these techniques will be described in detail and preliminary results presented.

STATISTICAL TECHNIQUES

The simplest and probably the most powerful statistical technique is to plot a time series of the data from all the anemometers along with the calculated mean, standard deviation and a covarience matrix for the instruments in service. A possible modification would be to first adjust the anemometer data by referencing all the data to the same height using equations like the Deacon profile[8]

$$\bar{u} = \frac{u_*}{k(1 - \beta)} \left[\left(\frac{z}{z_o} \right)^{1 - \beta} - 1 \right] \qquad (1)$$

where $\bar{u}$ is the mean wind at height z

u_* is the friction velocity $\sqrt{\text{shear stress/density}}$

k is Von Karmon's constant $\sim .4$

and β is factor which decreases with increasing Richardson number

These simple and basic statistical techniques require little computer storage and execution time and should prove suitable for computation at the ARAC Site Facility (ASF) to check for the gross errors. More elegant statistical techniques (mentioned below) will be better suited to more subtile instrument problems.

The next technique we have applied is built on the properties of probability density.[9] If the wind velocity u is a random variable then

$$\mu_n = \int_{-\infty}^{+\infty} u^n P(u)\, du \qquad (2)$$

where μ_n is the n-th moment of u (the second moment is the standard deviation); P(u) is the probability density function of u. These moments give insight into the symetry and shape of the density function. When a density function is close to normal, the Gram-Charlier representation is convenient for deriving a correction to a normal distribution. Deviation from a normal distribution is given by the coefficients of the series expansion. For simplisity we take those moments about the mean and the time varying velocity in standard measure. P(u) can then be expanded as

$$P(u) = \phi(u)\ \{1 + 1/6\ \mu_3 H_3 + 1/24(\mu_4 - 3)\ H_y + \ldots\} \qquad (3)$$

where $\phi(u) = \frac{1}{\sqrt{2\pi}} \exp(-1/2\ u^2)$

The functions H_n are Hermit polynomials. The first few are given by

$$H_3 = u^3 - 3u \qquad (4)$$

$$H_4 = u^4 - 6u^2 + 3$$

The expansion coefficients μ_n are derived by Equation (1). μ_3 and $\mu_4 - 3$ are the skewness and kurtosis respectively. The skewness is a direct indication of the asymmetry of a density function. Skewness and kurtosis are zero when a density function is normal. The Monin-Obukhov similarity theory implies that the standard deviation (nondimensionalized by u^* or scaling temperature T^*) is a universal function of the stability parameter z/L where L is the Monin-Obukhov length. The statistically higher order terms, skewness and kurtosis, are also expected to be universal functions of the stability parameter z/L because the shape of the density function depends on the feature of the turbulence. Analogously the moment of joint density function can be defined. The first moment is the covarience. The correlation coefficient is the covarience normalized by the standard deviation of each variable component

$$\rho_{ij} = \frac{1}{\sigma_i \sigma_j} \int \int_{-\infty}^{\infty} u_i u_j \, P(u_i, u_j) \, du_i du_j \qquad (5)$$

where $P(u_i, u_j)$ is the joint density function of wind from two towers or heights in our case.

The next statistical technique applies recent improvements in power spectral analysis developed by O. N. Strand and Chadwick[10] to a technique developed by Burg[11] in 1967. This technique is manditory when the number of data points is limited (as in our case where only 44, 15 minute averages were available). In a continuous mode operating system this would imply faster detection of instrument changes. The basic idea of the maximum entropy method (MEM) is the minimal assumption about the unknown values of a stocastic process in order to fit filters (whitening) to obtain the reciprocal of the squared magnitude of the Fourier transformed filter coefficients. The earlier methods artifically extend the data assuming a zero or periodic signal with specified windows. The computer code developed by

Strand uses a Tapered Entropy Method (TEM) which fits a filter by optimal extension of the covariances to keep the covarient determinant properly bounded (i.e. positive definite).

RESULTS

Figure 1 a) through e) show the time series from 5 anemometers with 44, 15 minute averages. Three of the anemometers, TV10, TV36 and TV304 were on a television tower near the Savannah River Plant (SRP) at heights of 10, 36, and 304 meters. These data were not adjusted for varying height. These data are shown in Figure 1 a), b) and c). Figures 1 d) and e) show analogous data from a network of 30 meter towers (NET F and NET H respectively). Within these data there is something obviously wrong with the signals from TV304 and NET F. NET F wind speeds were almost twice that of corresponding data from the other anemometers and TV 304 seemed to "die" halfway through the data period. The remaining three anemometers and associated data systems appear to have functioned properly during this period. Table 1 shows the result of elementary statistics which can be expressed as single numbers. The overall impression one gets from looking at this table is that unless one carefully assigns more importance to one statistic (such as mean) than others all the instruments could be found to give a reasonable value for at least one statistic. On the other hand, only TV36 seems to have avoided an extreme value in at least one statistic.

Figure 2 a) through e) show the one dimensional probability density for the data shown in Figure 1 a) through e). These plots provide a visual display of the similiarity of the data to a normal distribution. For completeness the numerical statistics from Table 1 are included with the probability density. We can see that with this minimal data set only NET F has the appearance of a normal distribution. Also included in this Figure is the cumulative probability which is the integral of probability density from -∞ to the variable sigma unit. In this representation the probability density is the slope of the cumulative probability. Figures 3 a) through d) show the joint probability density between the stations listed on the ordinate and absissa. The contour shapes should lean to the right for positive correlation and to the left for negative correlation. Perfect symetry implies lack of correlation and a Gaussian distribution. The sharp edges on the distribution are partly due to the contour plotting routine and partly due to a scarcity of data points. Figure 3a) shows the most skewed symetry as a result of the relatively high correlation between the two stations. The symetry shown in Figure 3 c) is actually skewed slightly to the left because of the slight negative correlation between TV10 and the "sick" anemometer at 304 meters. The distribution function is actually split in two due presumably to the fact that the probability density for TV304 is partially for real data and partially associated with noise after the instrument died.

Finally, Figure 4 compares the Maximum Entropy Method (MEM) derived spectral densities for the data from the five stations versus the frequency up to the Nyquist frequency of 1/30 minutes. We can see that the individual character of each station is maintained in spite of a relatively high filter lag of 6, 15 minute averages. The peaks and spectral gaps from the five stations appear closely associated with the obvious exception of TV304 which has almost no energy in the higher frequencies.

CONCLUSIONS

We have established several viable options for the initial statistical analysis package for assessment of meteorological instrument performance. Each of the statistical analysis and plotting routines included in this paper require no additional storage or speed than is available for analysis at an ACF. However, the data we analyzed included only 44, 15 minute averages. For a longer data record, which would improve almost all the statistics we used, storage and computer time may be a problem at an ACF. This is especially true for the maximum entropy spectral analysis and the joint probability density display whose calculation time and storage increase as the square of the number of data points. In the future we hope to continue this data analysis with instruments through the process of deterioration to better quantify the relationships between the numerical and graphical statistics and potential instrument problems.

As one or more techniques emerge, we will incorporate them in the ARAC system as part of a standard check on the quality of the data that are received from the sites. In addition these techniques can be applied to other projects that are concerned with the reliability of meteorological measurements.

ACKNOWLEDGMENT

The authors would like to thank O. N. Strand for his cooperation in giving us his MEM spectral analysis code and a special thanks to Keith Grant whose ability to talk to computers has put the authors out of innumerable numerical quicksands.

BIBLIOGRAPHY

1. MacCracken, M. C., "User's Guide to the LIRAQ Model: An Air Pollution Model for the San Francisco Bay Area," Lawrence Livermore Laboratory Report UCRL-51983 (1975).

2. Dickerson, M. H. and R. C. Orphan, "Atmospheric Release Advisory Capability," Nuclear Safety, 17, p. 281 (1976).

3. Knox, J. B., Hardy, D. M., Sherman, C. A., and T. J. Sullivan, "Status Report: Lawrence Livermore Laboratory Wind Energy Studies," Lawrence Livermore Laboratory Report UCID-17157-1 (1976).

4. Alexeiev, Ju. K., Dalrymple, P. C. and H. Gerger, "Instrument and Observing Problems in Cold Climates," World Meteorological Organization, Tech. Note #135 WMO-#384 (1974).

5. Angell, J. K. and A. B. Bernstein, "Evidence for a Reduction in Wind Speed on the Upwind Side of a Tower," Journal of Applied Meteorology, 15, p. 186 (1976).

6. Wieringa, J., "An Objective Exposure Correction Method for Average Wind Speeds Measured at a Sheltered Location," Quart. J. R. Met. Soc., 102, P. 241 (1976).

7. Pendergast, M. M., "A Cautionary Note Concerning Aerodynamic Flying of Bivane Wind Direction Indicators," Journal of Applied Meteorology, 14, p. 626 (1975).

8. Deacon, E. L., "Vertical Diffusion in the Lowest Layers of the Atmosphere," Quart. J. R. Met. Soc., 75, p. 89 (1949).

9. Hayashi, M., "A Preliminary Study on the Statistical Characters of the Atmospheric Turbulence," J. of the Met. Soc. of Japan, 52, p. 400 (1974).

10. Strand, O. H. and R. B. Chadwick, "A New Procedure for Computing and Analyzing Whitening Filters and Maximum Entropy Spectral Estimates," paper presented at the American Geophysical Union Annual Meeting, San Francisco, CA, December (1974).

11. Burg, J. P., "Maximum Entropy Spectral Analysis," paper presented at the 37th Annual Meeting, Society of Exploration Geophysicists, Oklahoma City, Oklahoma, October (1967).

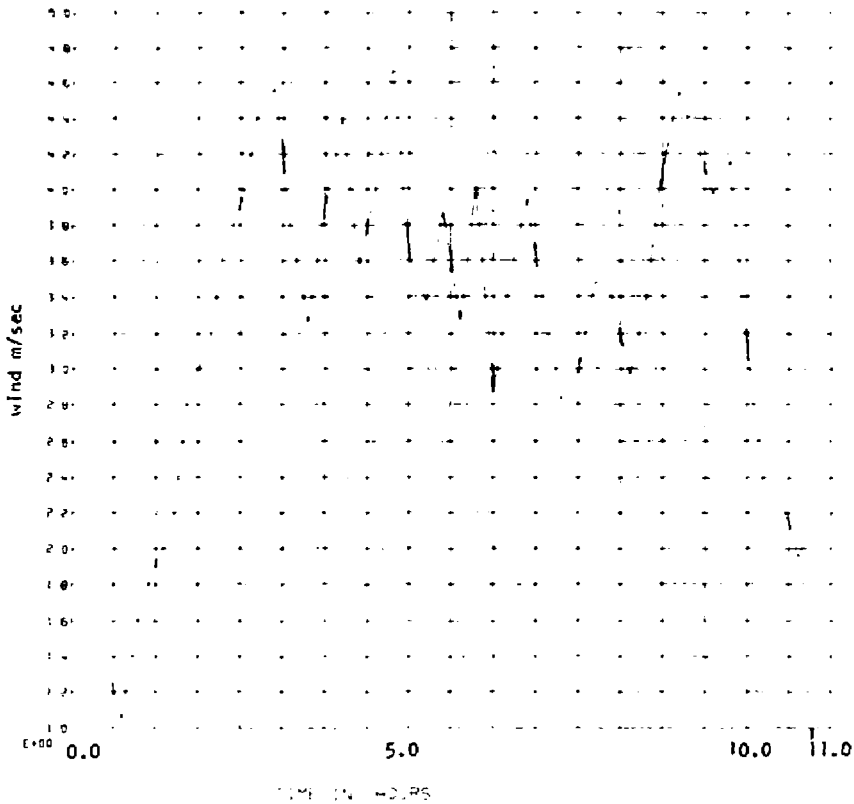


FIGURE 1a)

Time series plot of Savannah River Station TV10 on television tower at a height of 10 meters.

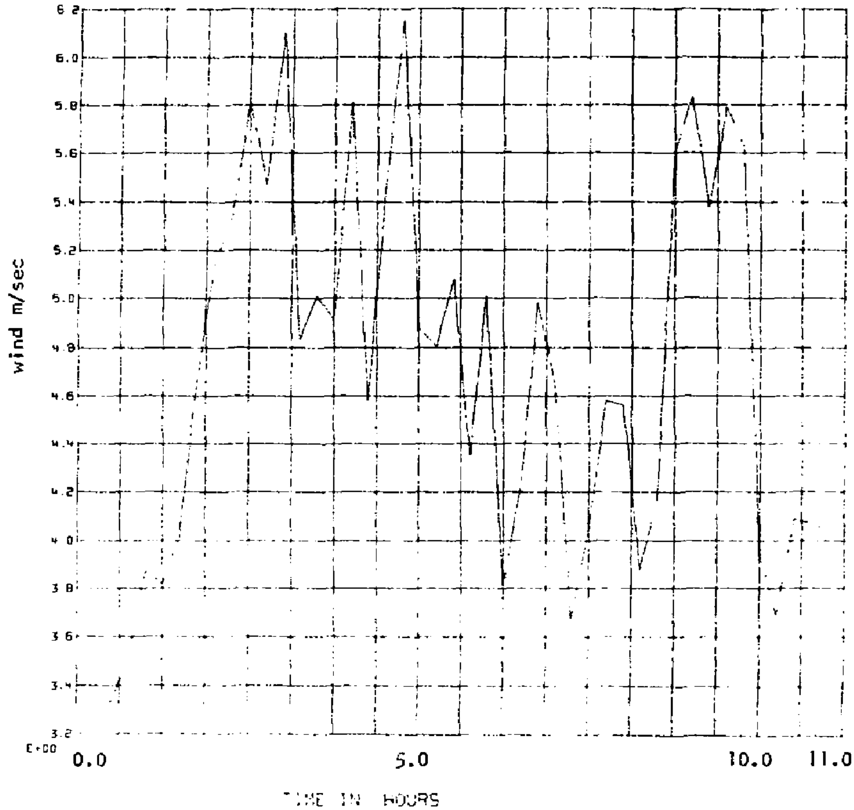


FIGURE 1b)

Time series plot of Savannah River Station TV36 on television tower at a height of 36 meters.

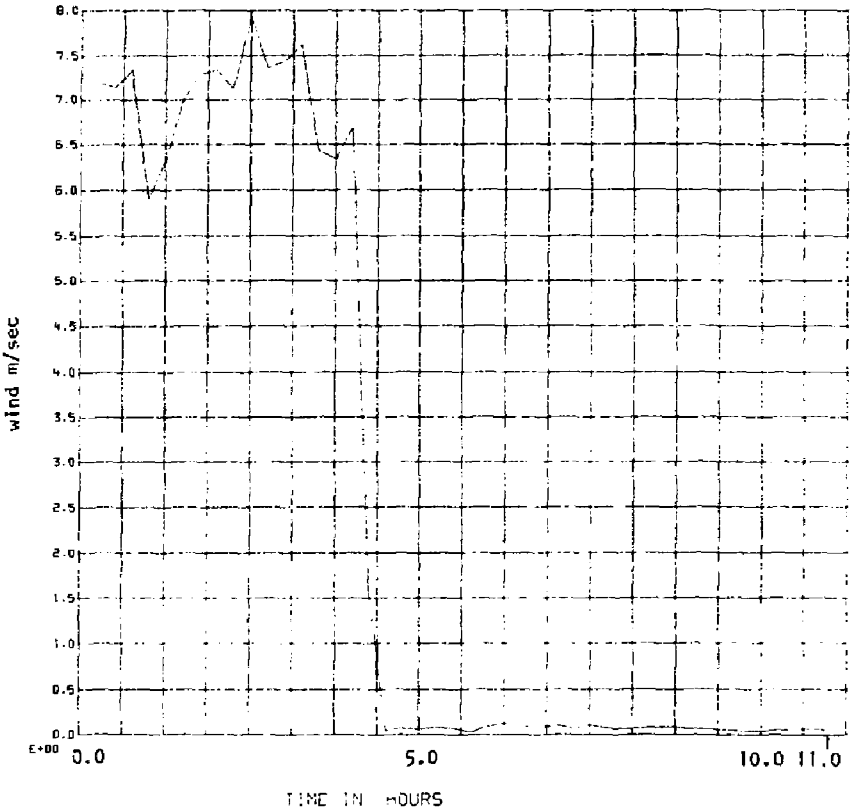


FIGURE 1c)

Time series plot of Savannah River Station TV304 on television tower at a height of 304 meters.

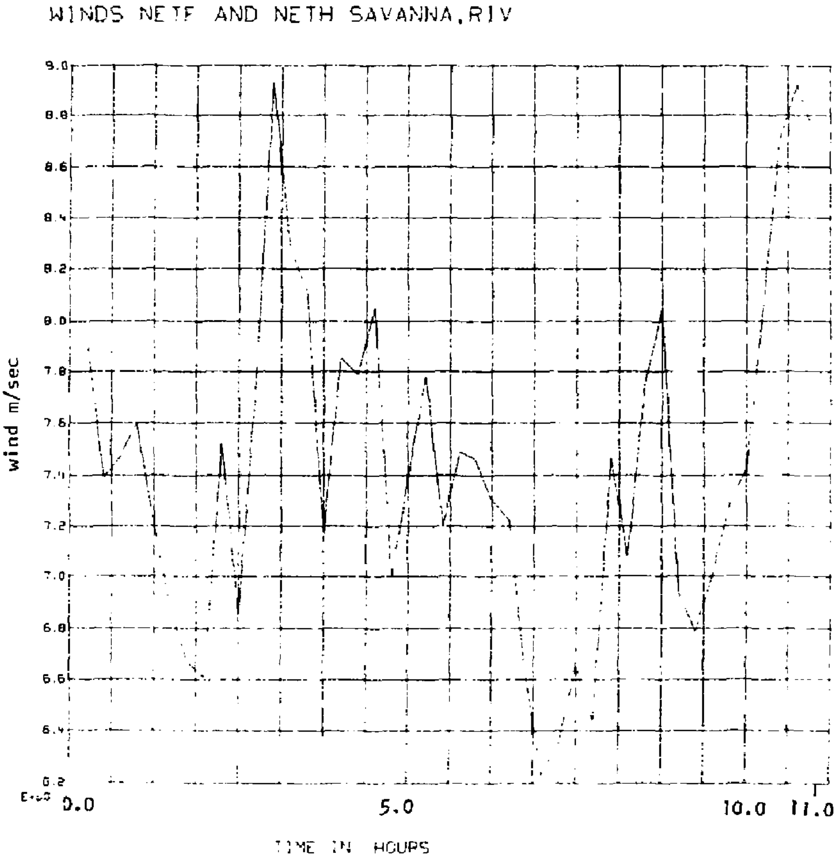


FIGURE 1d)

Time series plot of Savannah River Station NET F from anemometer network at a height of 30 meters.

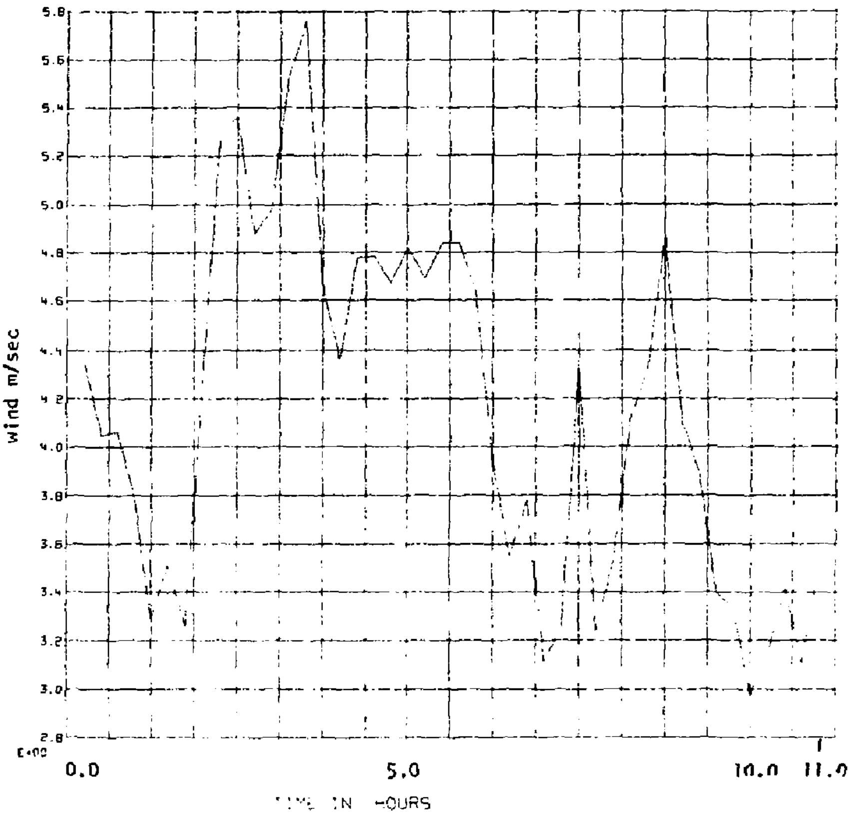


FIGURE 1e)

Time series plot of Savannah River Station NET H from anemometer network at a height of 30 meters.

TABLE 1

Elementary Statistics from Savannah River
Anemometer Network and 44 Fifteen Minute Averages

Station	Mean	Standard Deviation	Variance	3rd Moment About Mean	4th Moment About Mean	Kurtosis	Skewness
TV10	3.2 m/s	.94 m/s	.88	-.52	1.86	2.4	-0.63
TV36	4.7	.77	.59	.08	.66	1.9	.18
TV304	2.9	3.16	9.97	18.81	139.7	1.4	.60
NET F	7.43	.64	.41	.13	.44	2.61	.48
NET H	4.13	.72	.52	.07	.53	1.97	.20

Covariances

Anemometer	Pairs	Covariance
TV10	TV304	-.06
TV10	TV36	.57
TV10	NET H	.33
NET F	NET H	.15

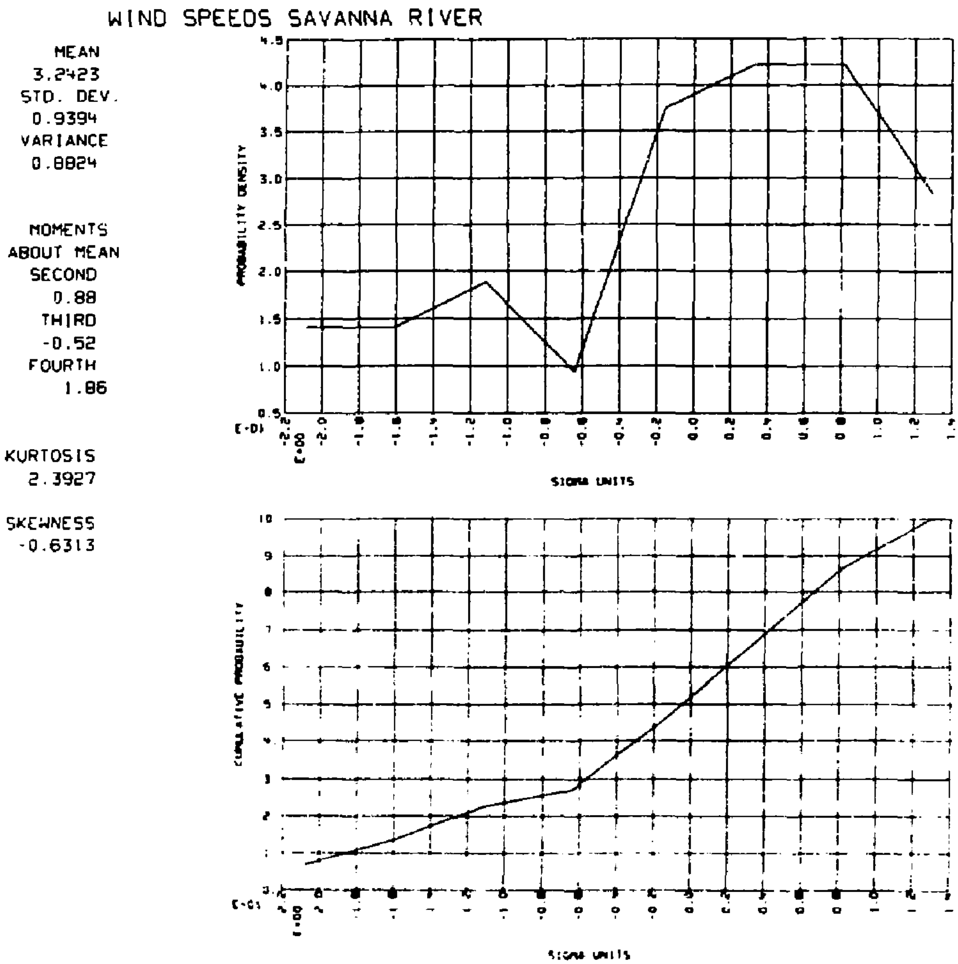


FIGURE 2a)

Wind station TV10 television tower at 10 meters.

WIND SPEEDS SAVANNA RIVER

MEAN
4.6919
STD. DEV.
0.7704
VARIANCE
0.5936

MOMENTS
ABOUT MEAN
SECOND
0.59
THIRD
0.08
FOURTH
0.66

KURTOSIS
1.8859

SKEWNESS
0.1842

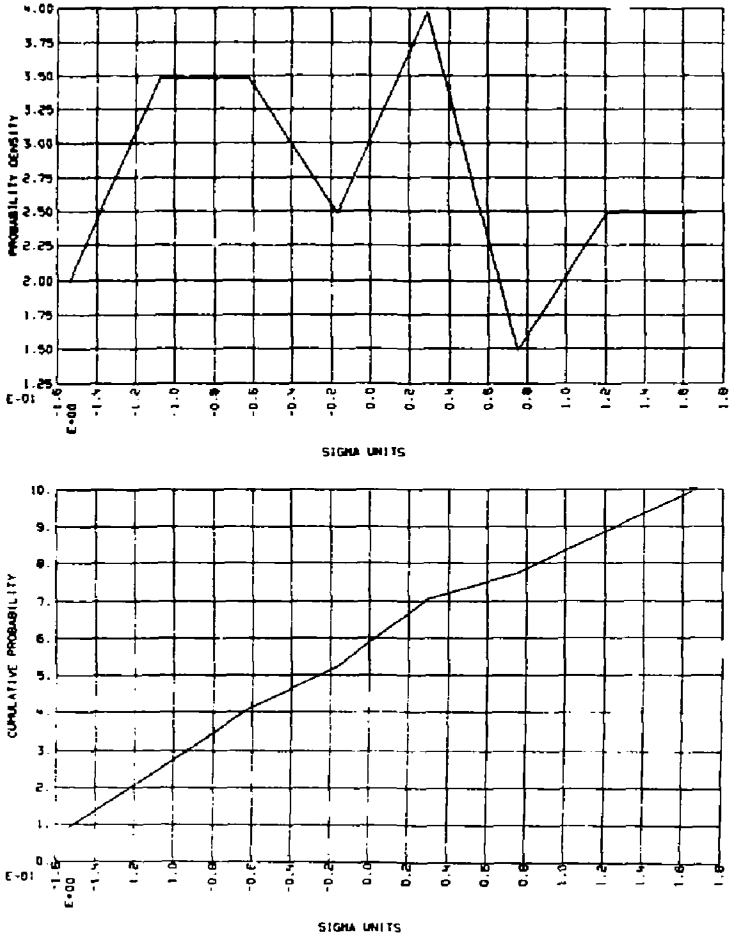


FIGURE 2b)

Wind station TV36 television tower at 36 meters.

WIND SPEEDS SAVANNA RIVER

MEAN
2.9088
STD. DEV.
3.1578
VARIANCE
9.9717

MOMENTS
ABOUT MEAN
SECOND
9.97
THIRD
18.81
FOURTH
139.69

KURTOSIS
1.4048

SKEWNESS
0.5973

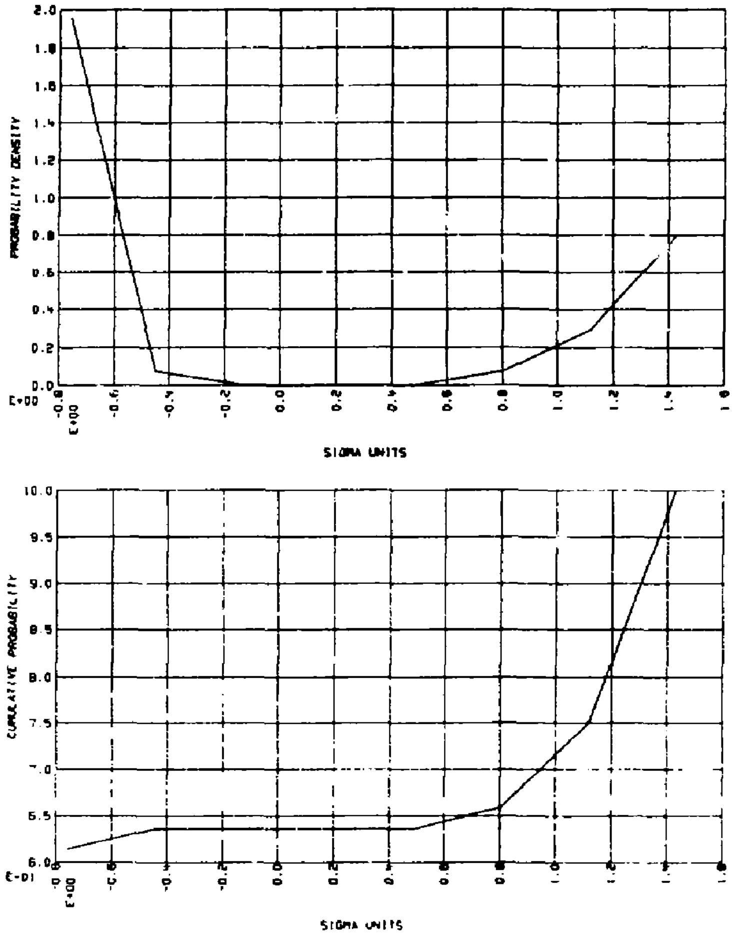


FIGURE 2c)

Wind station RV304 television tower at 304 meters.

WIND SPEEDS SAVANNA RIVER

MEAN
7.4343
STD. DEV.
0.6412
VARIANCE
0.4111

MOMENTS
ABOUT MEAN
SECOND
0.41
THIRD
0.13
FOURTH
0.44

KURTOSIS
2.6089

SKEWNESS
0.4790

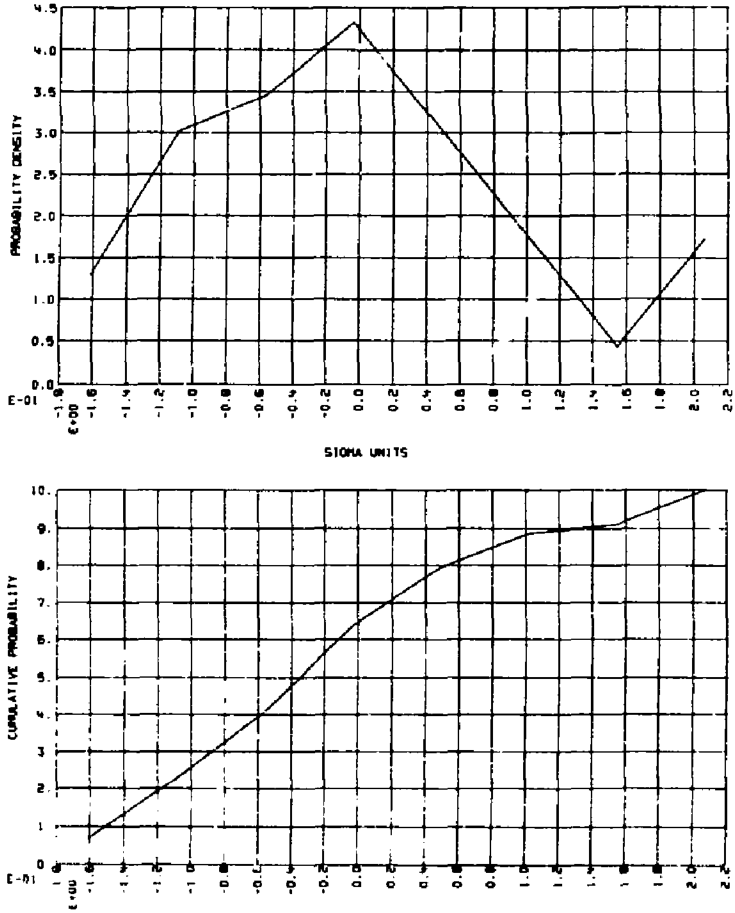


FIGURE 2d)

Wind station NET F anemometer network at 20 meters.

WIND SPEEDS SAVANNA RIVER

MEAN
4.1331
STD. DEV.
0.7193
VARIANCE
0.5175

MOMENTS
ABOUT MEAN
SECOND
0.52
THIRD
0.07
FOURTH
0.53

KURTOSIS
1.9663

SKEWNESS
0.1976

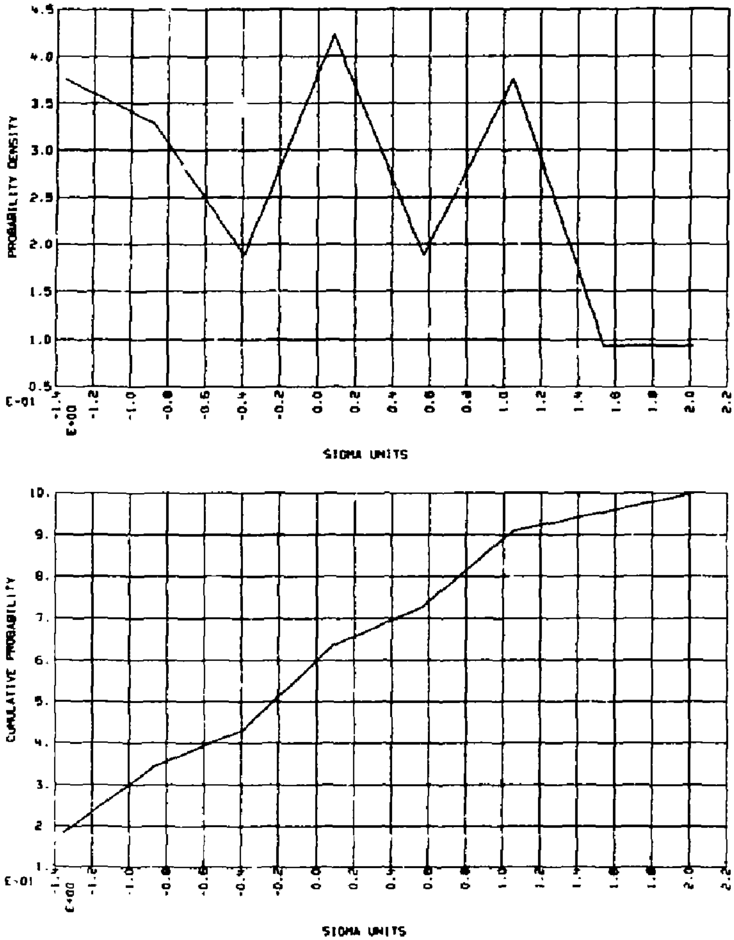


FIGURE 2e)

Wind station NET H anemometer network at 20 meters.

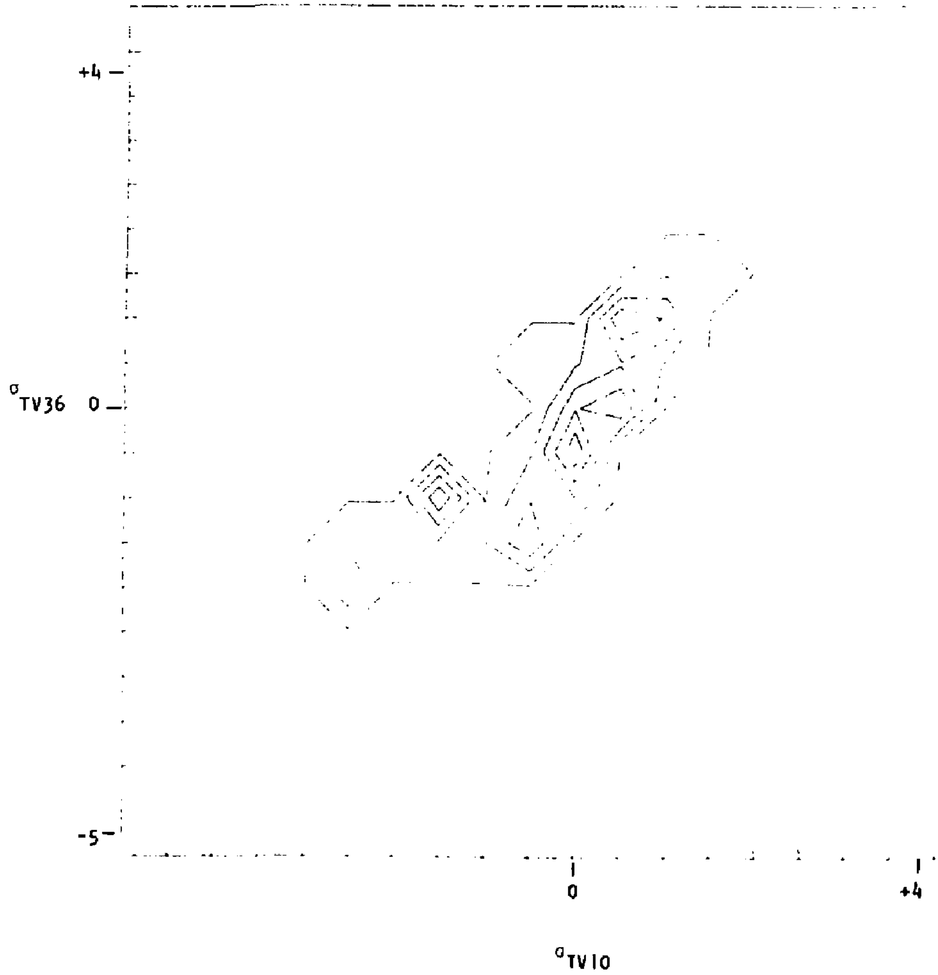


FIGURE 3a)

Contour plot of joint probability density between station TV36 and TV10.

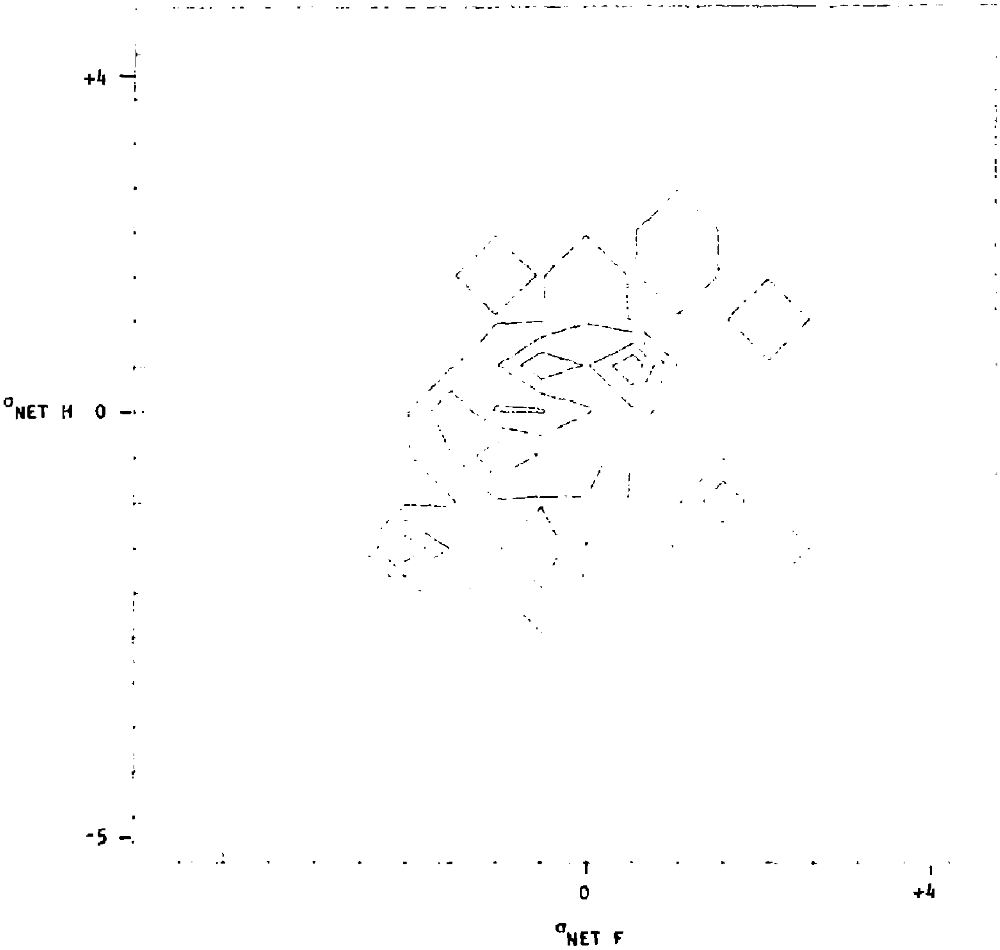


FIGURE 3b)

Contour plot of joint probability density between station NET F and NET H.

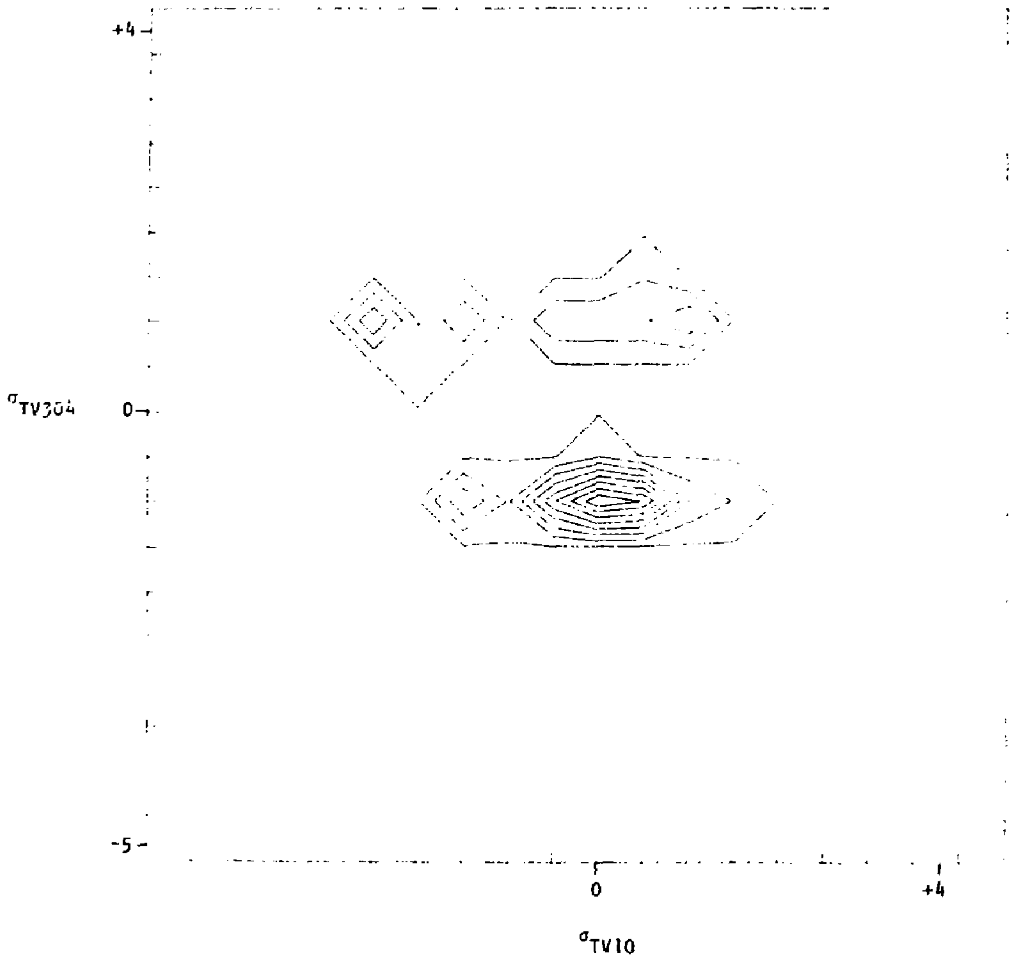


FIGURE 3c)

Contour plot of joint probability density between station TV10 and TV304.

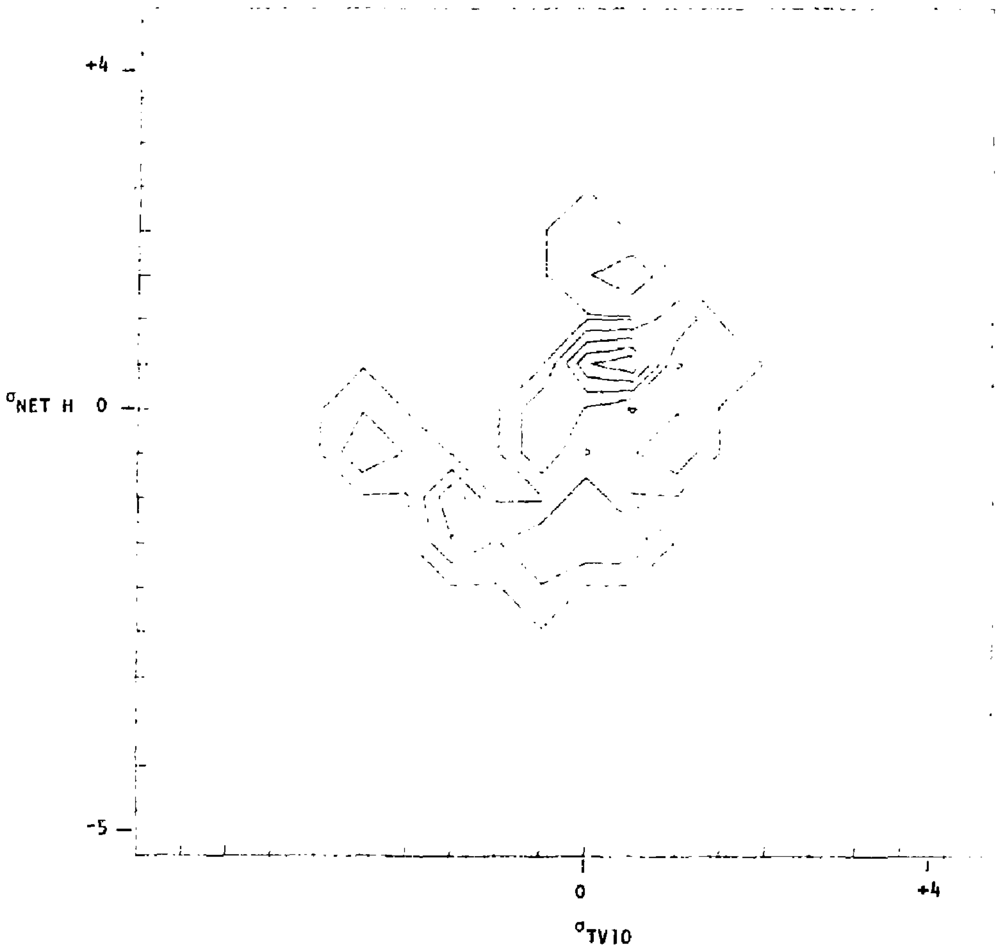


FIGURE 3d)

Contour plot of joint probability density between station TV10 and NET H.

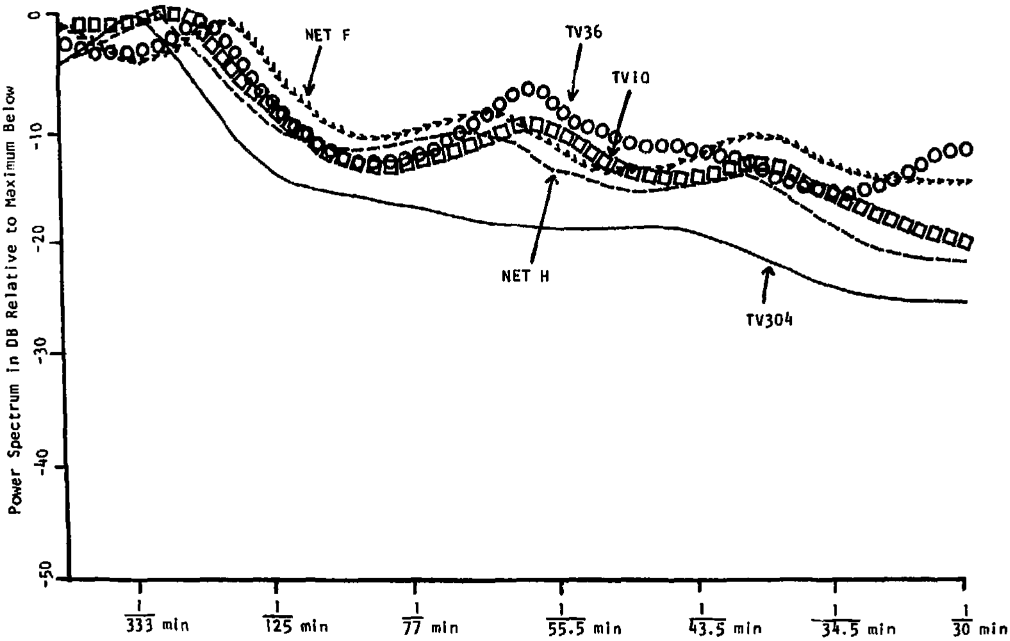


FIGURE 4

Maximum entropy spectral densities for 44 hourly (four 15 minute pts/hr) Savannah River for the Savannah River anemometer stations.