

A Statistical Description of Intense Rainfall

By MRS. A. E. FREENY and J. D. GABBE

(Manuscript received February 26, 1969)

This paper contains a statistical summary of the 14,000,000 measurements taken during 27 rainfalls in a six-month period in 1967 from a 96-station, rapid-response rain gauge network spread over a rectangular area 13 by 14 kilometers centered near Crawford Hill, New Jersey. The analysis emphasizes rain rates greater than 50 millimeters per hour, which interfere with radio transmission in the 10 to 30 GHz frequency range.

Heavy rain rates are relatively rare events, come in irregular bursts, and do not appear amenable to description by simple analytic distributions. This paper presents statistics concerning the behavior of rain rates at a point in space, the relationship of rain rates separated in space or time, and the relationship of average rain rates on pairs of paths in various configurations.

I. INTRODUCTION

This paper presents some statistics from the rainfall data collected on a rain gauge network during the period from June 1 to November 30, 1967. The network consists of 96 gauges spaced approximately 1.3 km apart on a rectangular grid centered near Crawford Hill, Holmdel, New Jersey. The design of the rain gauges and the equipment for recording data from the network are described elsewhere.^{1,2}

In communications, interest in rain-rate data arises from the relationship of attenuation of radio signals in the 10 to 30 GHz frequency range to the number and size of raindrops present in the transmission path. The quantity of water that the signal penetrates is directly related to the average rain rate on the path. Thus a major direction of our analysis was toward a statistical description of the behavior of rain rates at a point in space, the relationship of two rain rates separated in space or in time, and the relationship of average rain rates on pairs of paths in various configurations. Knowledge about these relationships, particularly for rain rates greater than 50 mm per hour,

where substantial attenuation occurs,³ is important for the design of microwave radio transmission systems.

A major characteristic of the rain-rate data taken in this experiment is its extreme variability, which increases as the rain rate increases. The separations between successive readings in time (10 s) and neighboring readings in space (~ 1.3 km) are too large to provide continuous representation of rain-rate behavior through time and space. Because the time series at the measuring stations cannot be considered even piecewise stationary during intense rainfalls, the usual time series techniques are not applicable. This leads to large numbers of descriptive statistics rather than a concise representation of the characteristics of rainfall.

A brief summary of the results and general conclusions given in Section VIII is: On the basis of 14,000,000 measurements obtained during 27 rainfalls that occurred in the Crawford Hill locale during the 1967 recording season, the empirical probabilities of observing point rain rates above 50, 100, 150, and 200 mm per hour are found to be 4.3×10^{-4} , 1.3×10^{-4} , 4.2×10^{-5} , and 1.0×10^{-5} respectively; the joint probability that the rain rate exceeds a given value at both of two stations simultaneously decreases rapidly at short distances as the separation between the stations increases, and goes through a minimum at a separation of about 12 km; the joint probability that the average rain rate on both of a pair of parallel paths exceeds a given value decreases as the path length increases, and shows a minimum for paths separated by 9 km; and the probability that the average rain rate will exceed 150 mm per hour on a single path 6.5 km long is 200 times greater than the probability that the average rain rate will simultaneously exceed 150 mm per hour on both of two parallel paths 6.5 km long and 6.5 km apart.

Detailed descriptions of the components of the analysis are:

Section II—treatment of the data, selection of a subset for analysis and procedures used for the detection and removal of spurious data;

Section III—general characteristics of observed rain, rainfall behavior at individual stations and in individual rainfalls, and the partitioning of the data made necessary by the variability of rain;

Section IV—statistics on point rain rates, both for the complete subset of the data and under the condition that the rain rate is greater than 50 mm per hour;

Section V—conditional and joint probabilities of various events for

pairs of stations at selected separations in space or time, and the relationship of the probabilities with distance and time;

Section VI—results of an analysis of average rain rates on pairs of paths in various configurations;

Section VII—engineering calculations and “benchmarks” for translating the relative probabilities calculated from the selected sample to “annual” probabilities based on the duration of the experiment, comparison of the results of various analyses using these data, and presentation of a result from another body of data.

II. TREATMENT OF THE DATA

Most bodies of data containing several million data points can be separated into segments of primary and secondary interest in the context of a particular analysis. Furthermore, raw data in such quantities inevitably contains some fraction of specious readings. In this section a brief description of the data collection system is followed by several presentations of rain-rate data and an outline of the data selection and screening procedures. The closing subsection contains remarks concerning the data retained for analysis.

2.1 *The Basic Data*

The data were acquired with a network of rapid-response rain-rate gauges most of which were mounted on telephone poles about 15 meters above the ground and well clear of all obstructions.^{1,2} The collecting surface of the gauge has an area of 478 cm²; the response characteristics of the gauge are such that the rain rate represents an average over less than one second. Although occasional anomalies in the data appear to be traceable to the gauges, they seem to give satisfactory readings for rain rates greater than about 10 mm per hour.

The network consisted of 96 stations, about 85 percent of which were operational at any given time. The area around Crawford Hill was divided into squares 1.29 km on a side by grid lines oriented north-south and east-west; one of the gauges was emplaced as close to the center of each square as practicable. The actual positions of the devices are shown on the map in Fig. 1. The data from the rain gauges, in the form of oscillator frequencies, is telemetered via telephone lines into the telephone central office serving the gauge location. At the central offices, which act as collecting points, the readings are commutated and forwarded to the Crawford Hill Laboratory, where the frequencies are detected. These frequencies and some auxiliary data are then automatically recorded on magnetic tape.

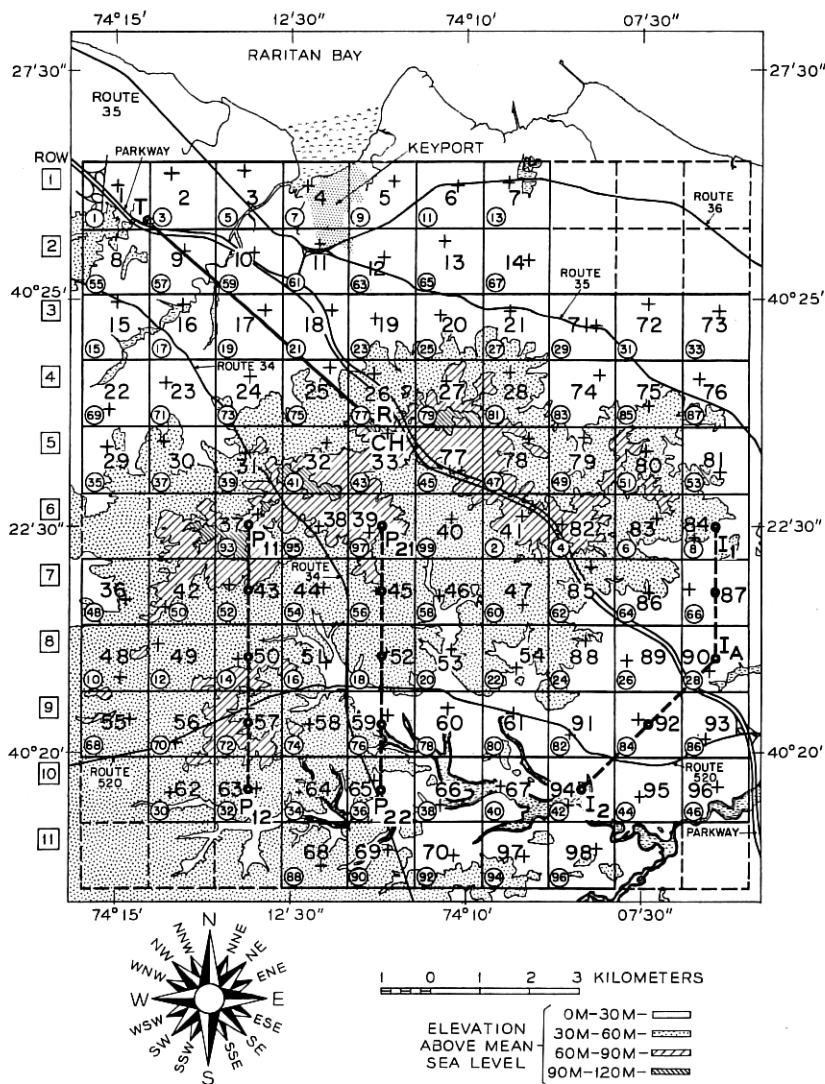


Fig. 1—The rain gauge network, showing the station numbers, scanning order (small circled numbers), and row numbers (in squares). Station 33 (CH) is at Crawford Hill, New Jersey.

During periods of rain, the network is scanned once every ten seconds to produce a "scan" of readings. The gauges are not read simultaneously, but are scanned at the rate of ten gauges per second in a sequence dictated by the telemetry arrangements and indicated by the numbers in the lower left corners of the grid squares in Fig. 1. The gauge at Crawford Hill (station 33) was read once per second separately from, and in addition to, the regular scan. This high frequency sample is recorded as part of the auxiliary data. The solid line between the transmitter (T) and receiver (R) shows an experimental microwave transmission path. A pair of the parallel paths (P_{11} — P_{12} , P_{21} — P_{22}) and a pair of the adjoining paths (I_1 — I_A — I_2) used in the path analysis are indicated by dashed lines. During the time covered by this report, the devices were not interchanged among grid locations, so that a station number always refers to the same device. The telemetry system is exposed to interfering signals which may produce spurious readings, creating some of the data-screening problems discussed in Section 2.2.

The data recorded at Crawford Hill are transferred from the original tapes onto tapes compatible with the computers used in the data analysis. For a variety of reasons this process could not be carried out successfully with about ten percent of the tapes; these data were lost. The next major processing programs transformed the data from oscillator frequencies to rain rates by applying calibration functions.⁴ As noted in Ref. 4 these calibrations have been optimized for the higher rain rates at the expense of some loss of relative accuracy below 25 mm per hour. The programs also (i) discarded time periods during which rain rates greater than 15 mm per hour were recorded at fewer than four stations during each scan for at least 20 consecutive minutes, (ii) indicated some kinds of questionable data, (iii) appended additional information, such as a list of stations known to be inoperative, (iv) organized the data into a format suitable for subsequent analysis, and (v) produced the isometric plots described in Section 2.2.

2.2 Data Selection and Screening

It is neither practical nor desirable to analyze all the data acquired from the rain gauge network. This is partly because more than 70 percent of the data were taken when the rain was too light to seriously affect microwave transmission in the frequency range of interest, partly because a tiny but important fraction of the data are spurious, and partly because the total number of data points is so large (over 14,000,000).

The first selection procedure operating on the data was "natural" selection, that is, owing to various equipment failures we have no data for a number of rainfalls. The second selection took place during the data processing where lengthy periods, during which virtually no rain was falling, were discarded. At this point about 430 hours of processable data remained. The procedures used to identify data of interest from the available 430 hours and remove spurious values are based on several graphical presentations of the data.

The isometric plots, which were produced for each rainfall in the final stage of the data processing, are the most basic graphical presentation. These plots give one a spatial and temporal appreciation of the rainfall as a whole, in addition to displaying particular features of the data in an illuminating manner.

Three hours of data from the rainfall of July 28, and one hour from July 11, 1967 are displayed in Figs. 2a and b. Each solid trace represents the rain-rate measurements from a single station, and is constructed by connecting six-reading averages in Fig. 2a and two-reading averages in Fig. 2b. The time in minutes is measured from the first data set. The traces are arranged in a horizontal and vertical grid which is isomorphic to the geographic grid in Fig. 1. Each trace is labeled with the station number. No trace is plotted if the station is known to be inoperative.

On the basis of the isometric plots, a sample of about 110 hours of rain was chosen for analysis in the following somewhat subjective but operationally convenient manner. An hour's worth of rain was included or rejected as a unit, with the exception of one isolated 20-minute shower which was included. A unit during which any six-reading average rain rate exceeded 50 mm per hour was included. A unit during which no six-reading average exceeded 30 mm per hour was excluded unless it fell between two units which were included on the basis of the 50-mm-per-hour criterion. Units with occasional six-reading averages greater than 30 mm per hour but less than 50 mm per hour were included if there were other units on the same day chosen on the basis of the 50-mm-per-hour criterion, and excluded otherwise. Subject to these constraints, the beginning and end of each time period was determined somewhat arbitrarily in terms of such operational conveniences as accepting an "entire rainfall" or eliminating a few bad magnetic tape records near the beginning or end of a rainfall. For example the entire three-hour period shown in Fig. 2a was accepted, and no effort was made to delete the first 20 minutes, during which there is little rain. On the whole, we were generous in the

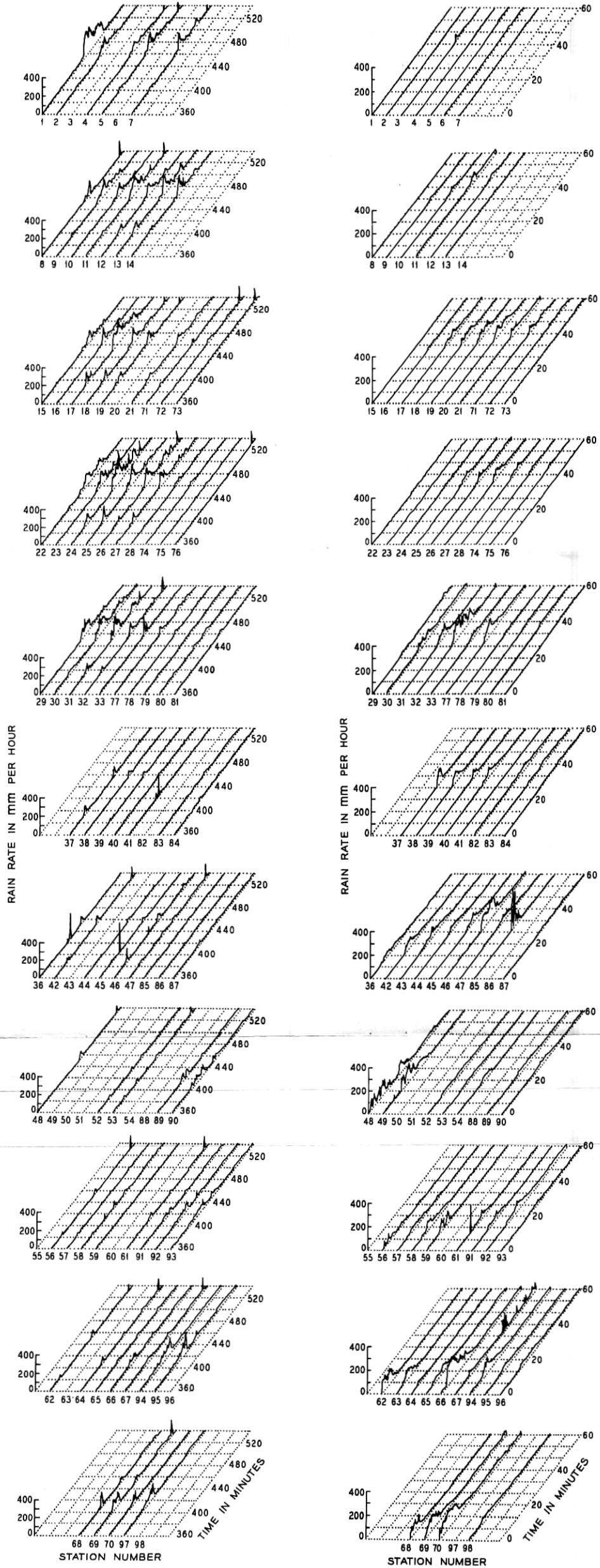


Fig. 2—Isometric plots of rainfalls for (a) July 28, 1967 and (b) July 11, 1967.

(a)

(b)

inclusion of data. This selection procedure reduced the data base to about 110 hours of "interesting" rain.

Attention was next directed to removing spurious values from the data. It is clearly impractical to attempt a point-by-point examination of so large a body of data (3.4×10^6 data points remained after the selection process), so various anomalies in the data were identified and studied with the aid of three additional presentations of information on the distribution of rain rate.

The most primitive displays of distributional information consisted of histograms giving the frequency of occurrence of rain rates for each station for each rainfall. Figure 3 contains four examples of the histograms from the rainfall of July 28, 1967, comprising 3840 readings per station, or about 11 hours of rain. The ordinate is the number of occurrences in 2 mm per hour intervals plotted on a logarithmic scale so that the frequencies of occurrence of the higher rain rates are visible in the diagram. Note the very substantial differences in the frequency of rain rates greater than 50 mm per hour at the four stations.

The frequency distributions were also displayed as probability plots of the rain-rate observations (ordinate) against quantiles of the standard normal distribution (abscissa).⁵ (In this presentation, a normal distribution would plot as a straight line with slope equal to the standard deviation.) Figure 4 contains four examples of the distributions of data pooled for an entire rainfall and two examples of the distribution of rain rates from individual stations. Adjacent curves are translated vertically by 100 mm per hour to increase the clarity of the plot. Curves *a* and *b* are from the rainfall of July 28, 1967 pooled over all operational stations. The data base for curve *a* contains 314,021 observations; that of curve *b* is smaller by 22 observations identified as outliers. Curves *c* and *d* are from the rainfall of July 11, 1967. Curve *c* is based on 285,098 observations pooled over all stations, curve *d* on the 281,869 observations that remained after deletion of the data from a "malfunctioning" station. Curves *e* and *f* are based on the 3840-observation sets collected during the rainfall of July 28, 1967 from stations 10 and 84, respectively. The distributions are far from normal, having either an overwhelming surfeit of low (less than 40 mm per hour) values or too long a high-rain-rate tail, depending on the point of view. The presentation has been retained because: (i) of the large number of analytic distributions with fixed parameters that we have tried, none provides a good fit to even a minority of the empirical distributions; (ii) the rain rates higher than 40 mm per hour, in

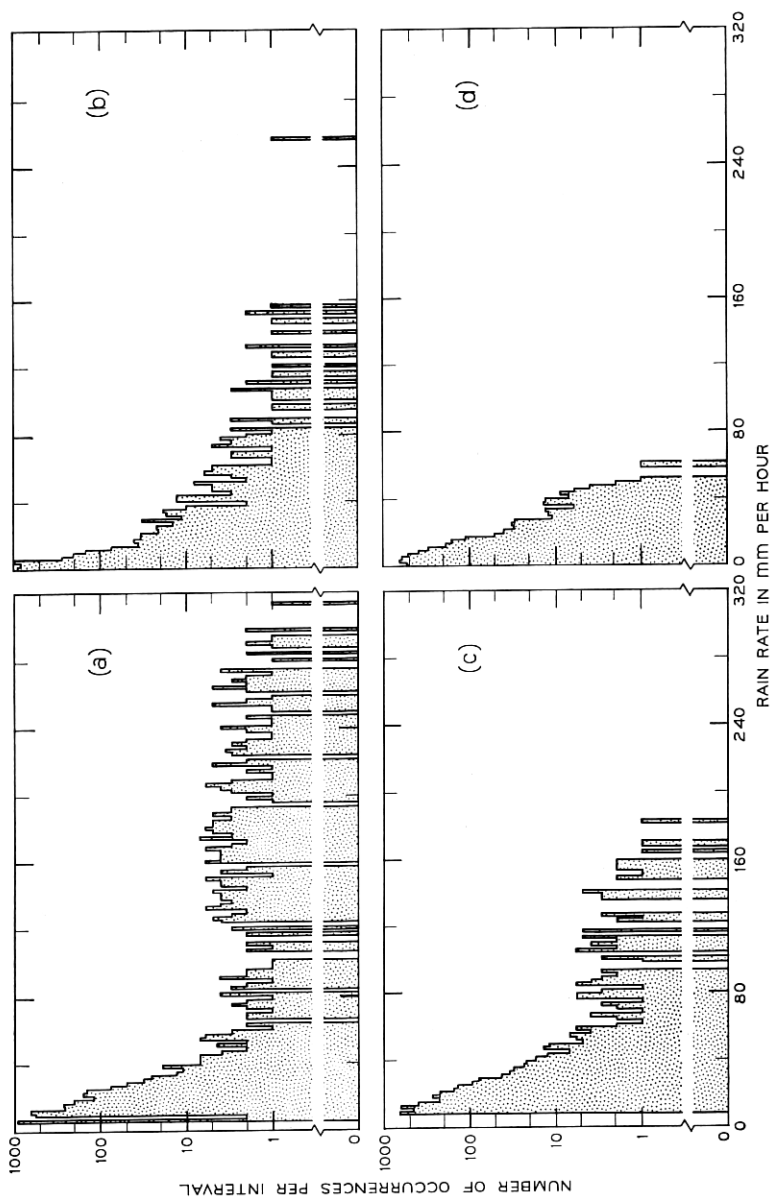


Fig. 3—Histograms of the frequency of occurrence of rain rates in 2 mm per hour intervals for (a) station 1, (b) station 10, (c) station 27, and (d) station 84.

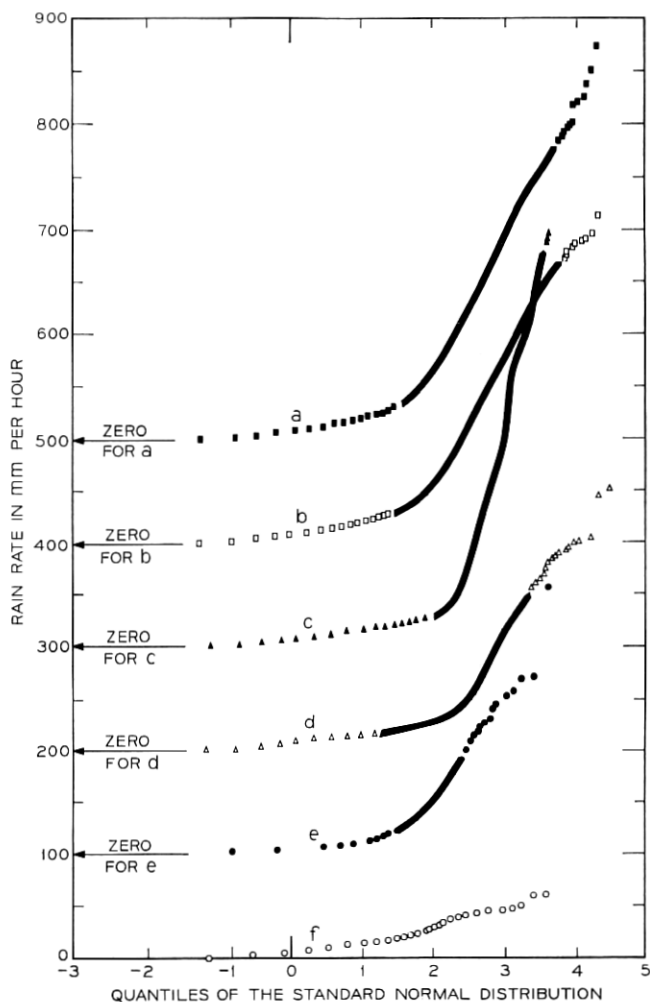


Fig. 4—Plots of empirical cumulative probabilities.

which we are particularly interested, are approximately linear on the normal plots (implying that these observations may be considered to be part of separate subpopulations); and (iii) the normal plots are readily produced and provide a convenient standard for comparisons among the empirical distributions.

Percentile information was also extracted from the frequency distributions and used to make pseudogeographic percentile plots. Figure 5 displays four of these. Each set of five dots connected by a

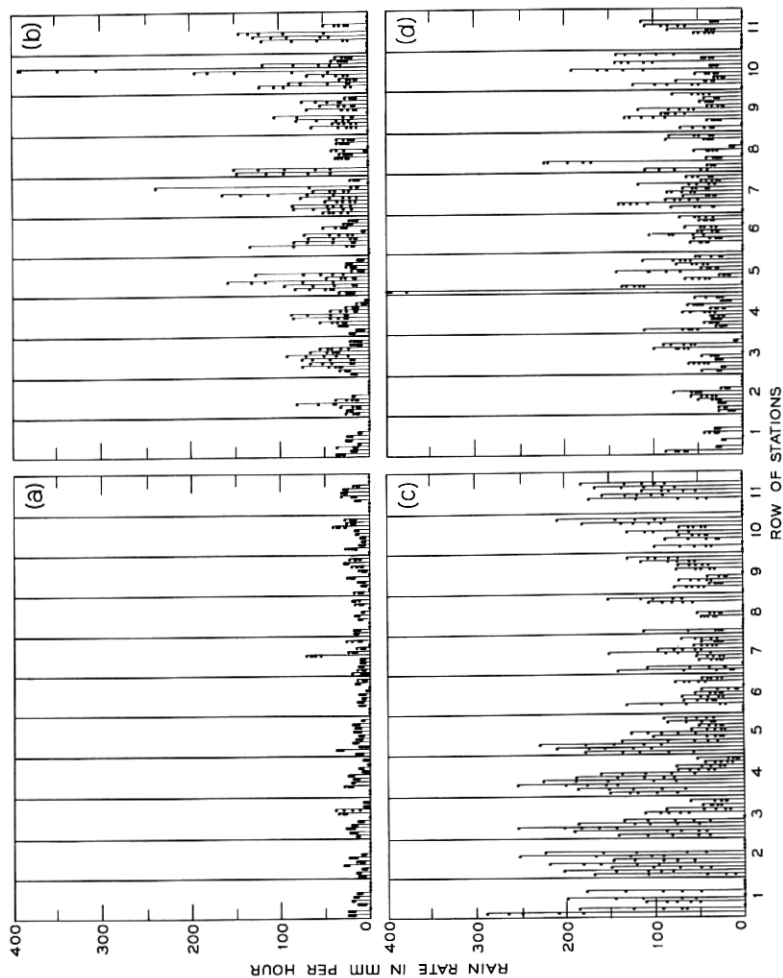


Fig. 5—Pseudogeographic percentile plots for rainfall of (a) July 10, (b) July 11, (c) July 28, and (d) November 17, all in 1967. Individual stations may be identified by the correspondence between sets of ten stations and the rows of Fig. 1. A single dot on the abscissa indicates an inoperative (or nonexistent) station.

vertical line represents the five points of the empirical cumulative rain-rate distribution corresponding to 99.9, 99.5, 99.0, 98.0, and 97.0 percent in descending vertical order. Each block along the abscissa corresponds to a row of stations on the grid of Fig. 1. Thus these plots display the geographic distribution of the high-rain-rate tail of the rain-rate distribution.

The first step in the data screening process was to remove data from malfunctioning stations. The term station covers all the equipment for measuring, telemetering, and recording data associated with a single rain gauge. Malfunctioning stations were identified by looking at the isometric plots for stations behaving differently from the surrounding stations, and at the pseudogeographic percentile plots for stations having distributional tails that appeared peculiar in the context of distributions of the surrounding stations. When a station was judged to have been malfunctioning, it was treated as though it

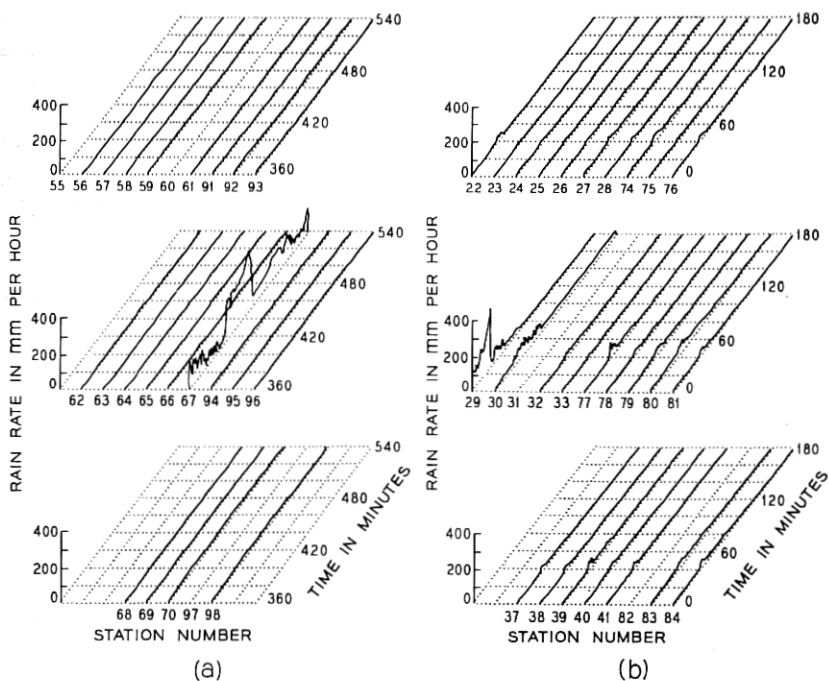


Fig. 6—Partial isometric plots of rainfalls for (a) July 11, 1967 and (b) November 17, 1967. The positions of the base lines of the traces correspond to the geographic positions of the stations, Fig. 1. Traces are six-reading averages.

had been inoperative throughout that rainfall; no more refined attempt was made to separate the spurious from the valid data.

Figure 5b shows that the tail of the distribution of data from station 67 during the rainfall of July 11, 1967, is much longer than the tails of the distributions recorded by the other stations. Part of the corresponding isometric plot, Fig. 6a, shows a rain rate in excess of 400 mm per hour at station 67 starting at minute 410 and lasting for nearly an hour, while the surrounding stations show no more than a light drizzle. Such a rainfall event is very unlikely in New Jersey and we discarded the data from station 67 for this storm. However the decision is not always so easy.

An earlier portion of the same storm is shown enlarged (two-reading-averages, rather than six-reading averages) in Fig. 2b. The behavior of station 67 after minute 25 is peculiar, but not outstandingly so, and without the subsequent evidence this data would have been retained. Another example of anomalous data is the behavior of station 29 during the storm of November 17, 1967. Figures 6b and 5d show this. The high-rain-rate tail of the distribution from station 29 is certainly very different from that of any other station. However, the isometric plot shows that the burst of intense rainfall (~ 400 mm per hour) lasted less than five minutes. As station 29 is on the edge of the network and there are only two stations which can be classed as close neighbors, there is little context in which to form a judgment of the validity of the data. In this case the data were retained.

These data-screening decisions may have substantial effects on the high-rain-rate tail of the distributions. Curves *c* and *d* in Fig. 4 are probability plots showing the distribution of the data from the rainfall of July 11, 1967 with and without the data from station 67, respectively. The effect on the shape of the rain-rate distribution for high rain rates is very evident. The probability of observing a rain rate greater than 100 mm per hour is changed by a factor of about 3; this factor increases rapidly at still higher rain rates (a factor of more than ten at 200 mm per hour). These results are typical for malfunctioning stations. Substantially all the retained data for rain rates greater than 350 mm per hour comes from station 29 during the November 17 rainfall, so the shape of the distribution in this region depends entirely upon this event.

After the malfunctioning stations had been disregarded, a number of individual bad data points were removed from the retained stations. Some of these data points are definitely not connected with the rainfall measurements. For example, the source of the numerous spikes

at time 530 on Fig. 2a is a faulty magnetic tape record. In other cases, such as the spikes at time 425 on station 36, 405 on station 45, and 370 and 435 on station 46, one cannot identify the extraneous sources. Outliers have been eliminated by examining the histograms and deleting any isolated points more than 50 mm per hour above what would otherwise be the last point in the tail of the frequency distribution. Such a point appears at 256 mm per hour in Fig. 3b. This point is also shown as the highest point on curve *e* in Fig. 4 (just under curve *d*) and is obviously inconsistent with the rest of the distribution.

Although the consequences of discarding outliers are less drastic than those of disregarding stations, the effect on the shape of the tail of the distribution is significant. Curves *a* and *b* in Fig. 4 show two probability plots of the data from the rainfall of July 28, 1968 pooled over all the operating stations. Curve *a* includes the outliers; curve *b* is without the 22 points identified as outliers by examining the station histograms individually. The five highest readings were among those judged outliers, and their deletion changes the shape of the distributions for rain rates greater than 250 mm per hour substantially. Extrapolations to rain rates above 300 mm per hour would give very different results in the two cases. The fate of spurious data in the more common rain rates is of less concern. They are overwhelmed by the valid readings and have little effect on the observed distribution.

In retrospect it is clear that the magnitude of the interference problem was not fully appreciated by the designers of the telemetry system. When one is interested in infrequent events (a probability of 10^{-6} corresponds to 30 seconds per year) cleanliness of the raw data is very important. We emphasize this point for future experimenters.

2.3 *The Retained Data*

The selection and screening processes leave a body of data (the "retained data") containing 3,418,623 observations at 96 grid locations on about 110 hours of rainfall. The observations are distributed roughly as in Table I.

These data were collected over 27 rainfalls of which nine may be classed as heavy (that is, at least 12 gauges showed rain-rate readings greater than 125 mm per hour). Parts a, c, and d of Fig. 5 show heavy rainfalls; part b shows a light rainfall.

The question of what population of rain rates the data represent is more fully discussed in Section 3.6. For the present, notice that the selection of the retained data is operationally defined and does not lead to a sample whose relationship to the sampled population

may be precisely defined in statistical terms. Moreover, the selection procedure severely decimates the low-rain-rate part of the rain-rate distribution. This is intentional because the experiment is specifically oriented toward measurements of rain rates high enough to interfere seriously with the transmission of microwave radio signals, and the rain gauges are not designed to provide reliable readings for rain rates below about 10 mm per hour (in fact, once wet the rain gauges record some low rain rate for long periods of time).⁶

The retained data may be used to calculate empirical probabilities regarding the occurrence of various events given the condition that the rain rate at some station is greater than 50 mm per hour. Viewed this way, the sample contains only 50,000 points, a number which must be considered small in the context of the enormous variations observed in rainfall. The top 1100 of these (rain rate greater than 200 mm per hour) are in addition particularly subject to the screening problems discussed in Section 2.2. Some evidence presented in Section VII suggests that the distribution of rain rates above 50 mm per hour derived from the retained data is not atypical of summer and fall rains in littoral northern New Jersey. Broader interpretations of the data should be viewed with appropriate caution.

III. GENERAL CHARACTERISTICS OF OBSERVED RAIN

The outstanding characteristics of the rainfalls observed are the variety of behavior, both within and among individual rainfalls, and the extremely rapid fluctuations in the measured rain rates. First we discuss a possible demographic model for rain rate and the poolings of the data necessitated by the variability. Then we discuss sampling rates at one station; typical spatial, temporal, and distributional behavior; and a systematic difference among the observations from different stations. Finally we consider the matter of the relationship between this sample and the general population of rain rates.

3.1 *A Demographic Model*

It is convenient to have a demographic model in terms of which the measurements can be discussed. The very elementary phenomenology offered in the following paragraph is only a convenient hypothesis which should not be regarded as being confirmed in any sense by the data, some of which is discussed later in this section. (Of course, the model would not be presented if it were contradicted.)

A large variety of populations of rain rates is produced by differ-

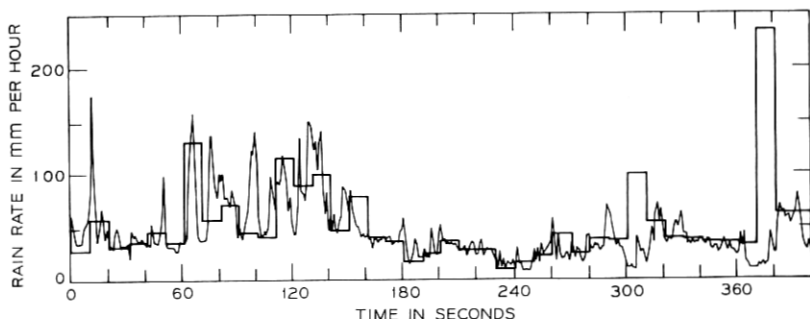


Fig. 7—Typical section of the time series from station 33 for the rainfall of July 25, 1967. The origin of the time axis has been arbitrarily set to zero.

ent, very local meteorological conditions. Each population has a characteristic distribution. A rain gauge records a sample of rain rates from the populations associated with those rain clouds which happen to pass over it. Thus a rain gauge observes a composite distribution which is the result of a two-stage sampling process, the first stage of which is the “natural” selection of the cloud, the second, the discrete scanning interval. Many rain-rate populations are inextricably intermingled as the data are recorded; indeed, as it may rain simultaneously from several strata of clouds, even a single rain-rate measurement may represent a mixture of populations.

Our extensive attempts to classify the heavy rain rates by distributional characteristics have been defeated by the large variety of distributions observed, and so subsets of the data have been pooled in various ways in attempts to synthesize some typical mixtures of populations. The unit of pooling is the data from one station during one rainfall. Some pools are: the data from each station pooled over all rainfalls for which the station was operational; the data from each rainfall pooled over all operational stations; the heavy-rainstorm pool, which contains all the data from the nine heavy rainfalls; and the grand pool which contains all 3.4×10^6 retained observations.

3.2 Behavior at Stations

Figure 7 shows two time series obtained from station 33 during the rainfall of July 25, 1967. The thin trace connects successive readings from the every-one-second (fast) sample. The thick trace represents the every-ten-second (slow) sample and is drafted as if the rain rate remained at the sampled value for ten seconds. (The equipment design prevents the every-ten-second and every-one-second readings from ever being coincident in time.)

The plot shows two notable features: the rain rate changes very rapidly from second to second, on occasion by a factor of three; and the magnitude of the fluctuations varies with the rain rate. These large, rapid fluctuations are characteristic of rain rates measured on a short (less than one-second) time scale with a small-area gauge.⁷ It is clear that the every-ten-second frequency of sampling the stations is not nearly fast enough to provide an accurate time-series representation of the rain rate at any gauge, but instead acts as a low-pass filter.

The rain-rate distributions of the slow and fast samples for an hour of the rainfall of July 25 are plotted against quantiles of the normal distribution in Fig. 8. Curve *a*, the slow sample, is based on 360 observations; curve *b*, the fast sample, is based on 3600 observations. The curves are nearly identical below the +2 quantile, indicating that the slow sample may be regarded as an unbiased sample of the fast sample and the distribution can be expected to be the same for large enough samples. The last three points in the upper tail of the slow sample affect the shape of the distribution above the second quantile significantly. One of these points is shown at time 370 seconds on Fig. 7. It is almost surely spurious, as are the other two outliers; the distributional effect of these three points should be discounted.

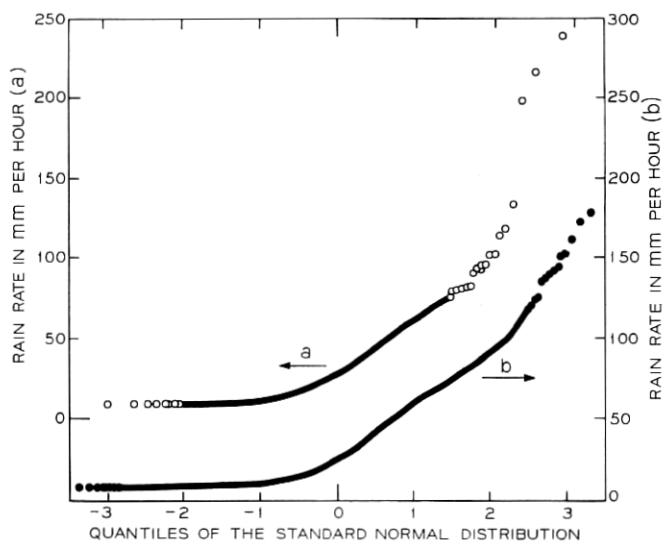


Fig. 8—Normal plots of the fast and slow samples from Station 33 for one hour of the rainfall of July 25, 1967.

Another characteristic of rainfall, which is apparent in Fig. 2, is that heavy rain comes in bursts. Although some of the showers last for almost 30 minutes, longer showers usually contain a number of short bursts. In general, these periods of rain cannot (at our sampling rate) be considered piecewise-stationary time series. Our many attempts to analyze periods during which the rain rate was greater than 35 mm per hour yielded results badly confounded by the nonstationarity produced by occurrences of these bursts. Thus one cannot naively apply time-series analysis techniques to the heavy rain rates; such common statistics as correlation and autocorrelation coefficients, and power spectra are not directly meaningful.

3.3 *An Overview of the Grid*

At about time 400 on July 28, 1967 it began to rain heavily on two separate portions of the grid, the southeastern portion and the north central portion (Fig. 2a). The southeastern rain built up slowly (in general) for about half an hour, decreased abruptly, and then resumed at lower intensities and intermittent intervals for another two hours. The north central rain was heavy for about ten minutes and was followed by 40 minutes of light rain. At about time 460, very heavy rain started in the northwest portion of the grid and continued for between 20 minutes and an hour before dying out. The general behavior of this rainfall is typical. It rains heavily first on one part of the grid then on another, the regions of heavy rainfall are fairly local (often only a few stations register heavy rain), and it seldom rains heavily for as long as an hour. The downpour is not continuous but has substantial variations in intensity throughout its lifetime.

In view of the rapid fluctuations of rain rate at the stations, it is not surprising that the correlation among the fine structures observed at different stations is poor. It would require an extraordinary mechanism to synchronize one-second rain-rate fluctuations over an area of several square kilometers. The temporal relationship among rain rates at the different stations may be seen on Fig. 2b. Even allowing for the fact that the stations are not sampled simultaneously (see Section 2.1) and that there may be a propagation lag, the structure of the rain-rate traces is not especially similar from station to station.

Of the 27 rainfalls, that of July 11 is one of the best examples of systematic motion of a rainstorm. The traces in Fig. 2b behave as though a rain cloud roughly 3 km wide and many kilometers long, with the long axis oriented in the WNW-ESE direction, moved NNE across the grid with a velocity of about 15 km per hour.

3.4 *Distributional Characteristics*

In discussing the distributional characteristics of rain rates, our attention is focused on rain rates greater than 50 mm per hour. Over our sample this is the upper $1\frac{1}{2}$ percent of the empirical distribution, although for particular stations and rainfalls the amount of data above 50 mm per hour varies from zero to five percent.

The purpose of this subsection is to demonstrate that there seem to be many different kinds of rainfall, so the data contain observations from many different populations with different rain-rate distributions and mixture ratios. Many common analytic distributions (the normal, log normal, gamma family, and so on) have been tried in attempts to find a simple description of the observed composite distributions which behaves reasonably in the tail region (rain rate ≥ 50 mm per hour); but none have been found that can serve even approximately.

We conclude that the large sampling variation and the complexity of the mixtures precludes obtaining reliable estimates of distributional parameters of rain rates above 50 mm per hour from any small sample. Some samples we consider small in this context are single stations for a season and the entire network for a single intense rainfall. For still higher rain rates the present overall sample may be inadequate (there are only 1100 observations for rain rates above 200 mm per hour). In the absence of a distribution on which to base precise estimates of formal statistics, such as confidence limits, discussion of these matters can best await the analysis of the 1968 data.

The remainder of this subsection indicates some of the distributional variety observed. The arrival of heavy rainfall in bursts, as indicated in Fig. 2, makes it obvious that the distribution of rain rates is very different for different segments of the same rainfall measured at the same station. The same holds true for the distributions observed at different stations for the same rainfall. This may be seen from the two probability plots for stations 10 and 84 (curves *e* and *f* in Fig. 4), from the differences in the tails of the distributions shown in Fig. 5, and the four histograms in Fig. 3. The rain-rate distributions differ greatly among stations even for the rainfall of July 11, whose active portion, shown in Fig. 2a, appears deceptively similar for many of the traces. Apparently many populations of rain rates coexist within rainstorms.

In an attempt to find a common mixture of populations, the stations have been pooled within rainfalls. Curves *b* and *d* in Fig. 4 are examples of such pooled distributions, and the difference between them is typical of what is observed. A glance at the percentile plots of Fig.

5 indicates that it is unlikely that pooling across these collections of stations can produce distributions with similar shapes above 50 mm per hour. The next pooling is within stations, across rainfalls, that is, at each station for the whole season. Naively this seems to be the procedure most likely to produce a "typical" sample and thus similar distributions. Cloudbursts, however, are limited in extent. Some of the stations were never hit by cloudbursts, whereas others appear to be deluged quite often. Figure 9a shows normal probability plots of the rain rates from each of four stations pooled for the season, and Fig. 9b shows the matching upper five percent of the empirical cumulative distributions. Adjacent curves have been translated vertically by 100 mm per hour. Curves *aa* and *ba* are for station 29, *ab* and *bb* for station 1, *ac* and *bc* for station 98, and *ad* and *bd* for station 76. The data base for each station contains about 39,000 observations. The distributions are seen to be quite different.

The overall pseudogeographic percentile plot, Fig. 10, shows the

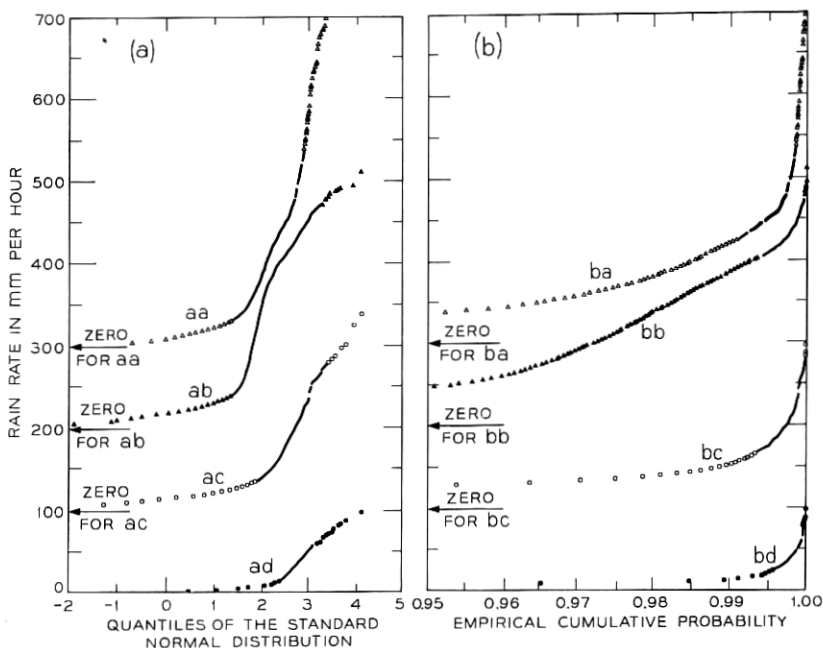


Fig. 9—Distributions of the rain rates at each of four stations pooled for the 1967 season: (a) plots against quantiles of the standard normal distribution; (b) the upper five percent of the empirical cumulative distribution.

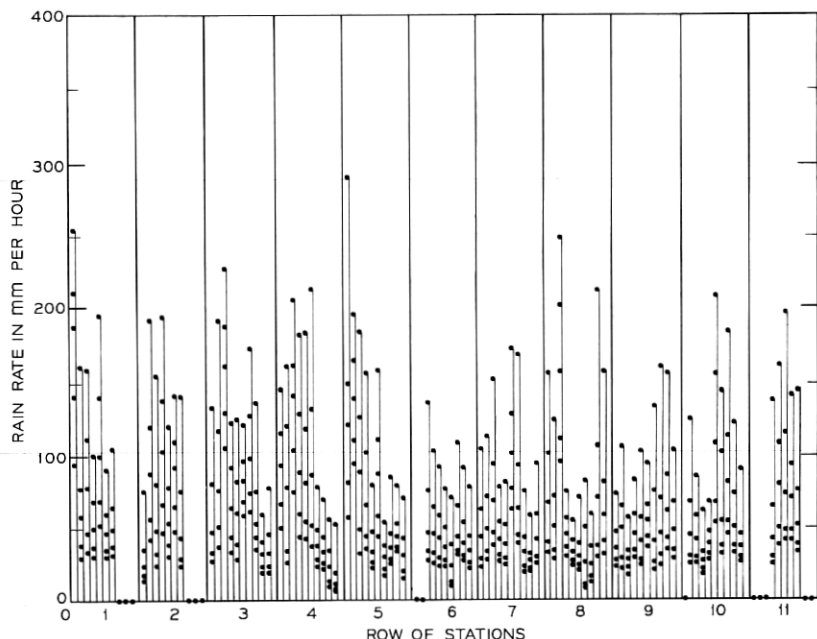


Fig. 10 — Pseudogeographic percentile plot. Individual stations may be identified by the correspondence between sets of ten stations and the rows of Fig. 1. The 99.9, 99.5, 99.0, 98.0 and 97.0 rain-rate percentiles of the empirical cumulative distributions are indicated for each station by the dots in decreasing vertical order. A single dot on the abscissa indicates a grid square that has no station.

upper tails of the distribution of rain rates from each station pooled across all the rainfalls (for which the station was operational) in the final sample. This plot covers the upper tail down to a rain rate of 50 mm per hour for most of the stations. The figure indicates the variety of distributions, and shows that the four samples selected for Fig. 9 are not atypical. It also confirms the notion that data from a single station over a season are not a sufficient sample on which to base rain-rate statistics for rain rates above 50 mm per hour.

3.5 Systematic Behavior

Figure 10 gives an idea of the areal distribution of the intense rain rates during the 1967 season. It rained most intensely on the northwestern part of the network and least intensely in the northeastern and southerly central portions. Intense rain was also observed by the stations in the southeastern section of the grid. A more detailed study

of the data reveal that station 1 always records more intense rain than most of the network and station 76 always records rain of lower intensity than most stations. Three possible explanations are confounded in the present data: first, there may be a systematic geographic effect; second, the effect may be inherent in the equipment of the stations (the calibrations of the gauges may have shifted); and third, one may be witnessing a sampling fluctuation. Action has been initiated to determine whether the effect is connected with the equipment. Analysis of data from 1968 may help to distinguish between the first and third possibilities.

3.6 *What Does the Sample Represent?*

The sample contains virtually all of the available data for rain rates greater than 50 mm per hour and about a quarter of the data below 30 mm per hour. As such it can be expected to be representative (with appropriate proportionality adjustments) of rainfall with rain rates greater than 10 mm per hour within a small area of northern littoral New Jersey over a relatively short period. Neither the year nor the locality are known to be exceptional in any respect.

The distributional stability of the observations above 50 mm per hour is poor, as has been indicated, which shows that the sample is small in the context of the demographic composition. In Section VII, the data are compared with the 1958 data taken at Island Beach, New Jersey by Mueller and Sims and found to be reasonably similar.⁸ Some informal indication of the dispersion of various statistics is given in the following sections by noting the quartiles and interquartile ranges of the distributions. (The quartiles are the 25 and 75 percent points of the distribution; the interquartile range is the difference between them.) While there is nothing to indicate that the results of this study are in any way atypical of heavy rainfall in this part of New Jersey, it is unlikely that they permit an accurate assessment of the extremes of rainfall that may not infrequently be encountered.

Extension of these results outside the immediate locality will, of course, be subject to even greater uncertainties.

IV. POINT-RAIN-RATE STATISTICS

4.1 *The Grand Pool*

Any interpretation of this section must take into account the fact that the retained data include all of the available data for rain rates

greater than 50 mm per hour but only about 25 percent of the available data for rain rates below 30 mm per hour.

Figure 11 is the grand histogram based on the 3,418,623 observations (about 110 hours of rain) retained from 1967. Log frequency of occurrence is plotted linearly on the ordinate in order to accommodate the range of frequencies. The interval of the histogram is 2 mm per hour. Although only about 25 percent of the data below 30 mm per hour were selected, the relative frequencies in this range are unlikely to have been substantially affected by this selection. However, the general tendency of the rain gauges to give biased readings which greatly overestimate the rain rates below about 15 mm per hour make this histogram unsuitable for hydrological use.

The empirical cumulative distribution is shown in Fig. 12a, and the upper four percent of the distribution is shown on an expanded scale in Fig. 12b. The shape of these curves may be adjusted to approximately represent the distribution of the entire 430 hours of rain recorded by scaling the probabilities below 0.96 by the factor $0.99/0.96$. Above 0.96 the new *probability*, P_N , would be given approximately by the expression $P_N = 0.75 + 0.25 P_B$, where P_B is the *probability* on the plot. The cumulative distribution points up the small fraction of the data which lies above about 50 mm per hour. The empirical probabilities, Fig. 12a, are transformed to quantiles of the standard normal distribution and the rain rates as ordinates are plotted against the quantiles in Fig. 12c. (Only the data above the mean, quantile 0, rain rate about 8 mm per hour, are shown.) If the data were a random sample from a normal distribution, the points would lie along a straight line. The data are not even approximately normally distributed, which must of course be the case as the frequency distribution (Fig. 11) decreases monotonically from the lowest interval.

As the distribution has such a long tail and is skew (no negative rain rates have been recorded although all that water must have gotten up there somehow), it seems reasonable to look at various long-tailed skew distributions such as the gamma, Wiebull, and extreme value. These distributions yield probability plots which do have an appreciably better appearance than the normal probability plot of Fig. 12c; however, they all show substantial deviations from linear behavior near rain rates of 100 mm per hour. Efforts to improve the linearity by adjusting the proportion of the data below rain rates of 30 mm per hour were not successful. Since a distributional description in such circumscribed conditions is of only very limited usefulness the details of these efforts are not reported here. Typical of the results is

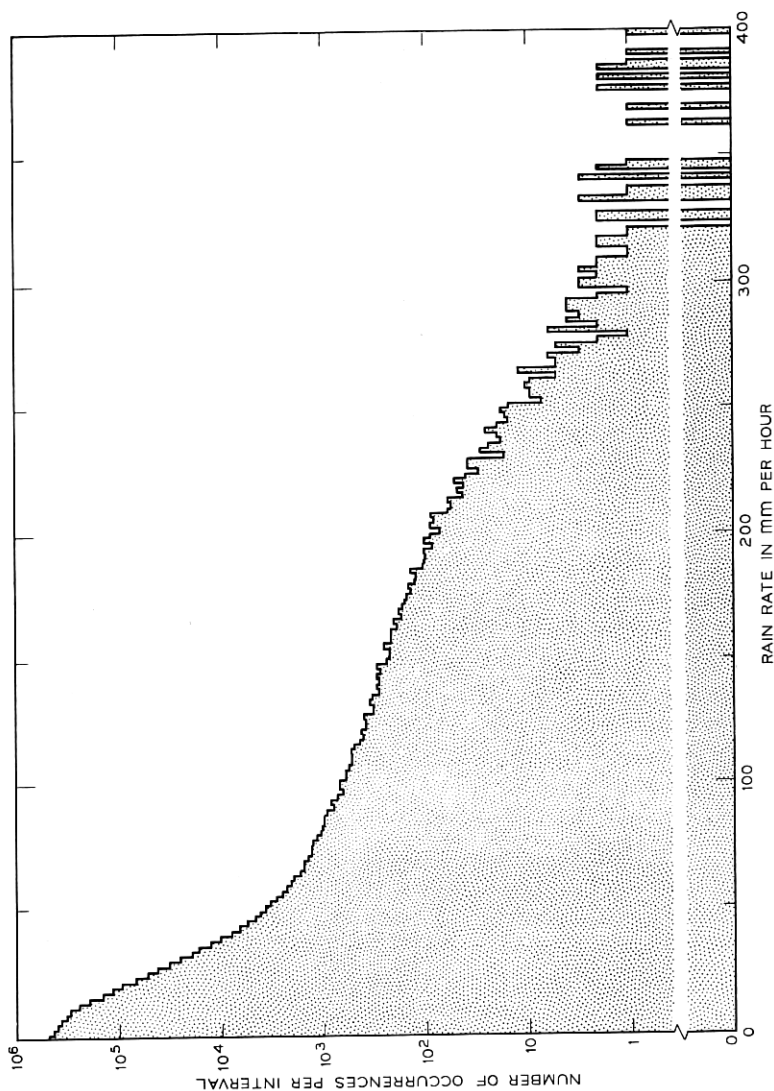


Fig. 11 — The grand histogram. There are 12 additional observations scattered above 400 mm per hour.

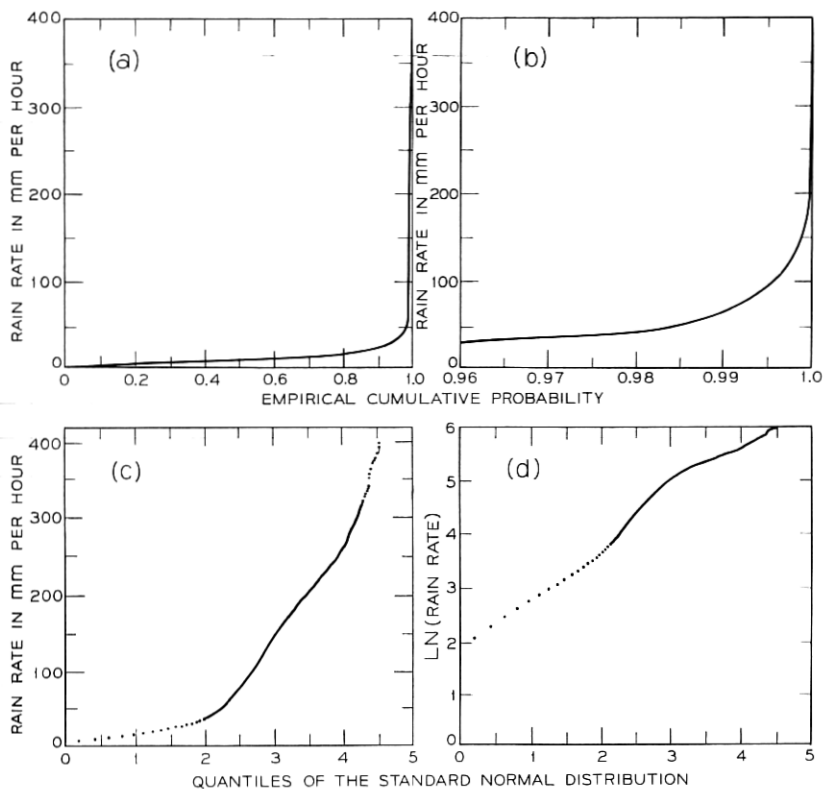


Fig. 12—Distribution of the grand pool: (a) cumulative distribution, (b) enlargement of the upper tail of cumulative distribution, (c) normal probability plot, and (d) log-normal probability plot. Only the part of the distribution above the mean is shown in parts (c) and (d).

Fig. 12d, which is a plot of log rain rate vs. quantiles of the standard normal distribution (that is, a probability plot of the “log normal” distribution). As in Fig. 12c, only data above the mean are shown. Those familiar with probability plots will recognize that the inflection at the 2.5 quantile cannot be removed by adjusting the fraction of data assumed to lie above 30 mm per hour.⁵

In the present data-analytic situation, where there is no simple analytic distributional description of the data, and indeed the data may represent a drawing in unknown proportions from many populations, it is not possible to compute formal confidence limits for the various estimates of probabilities. An informal indication of the

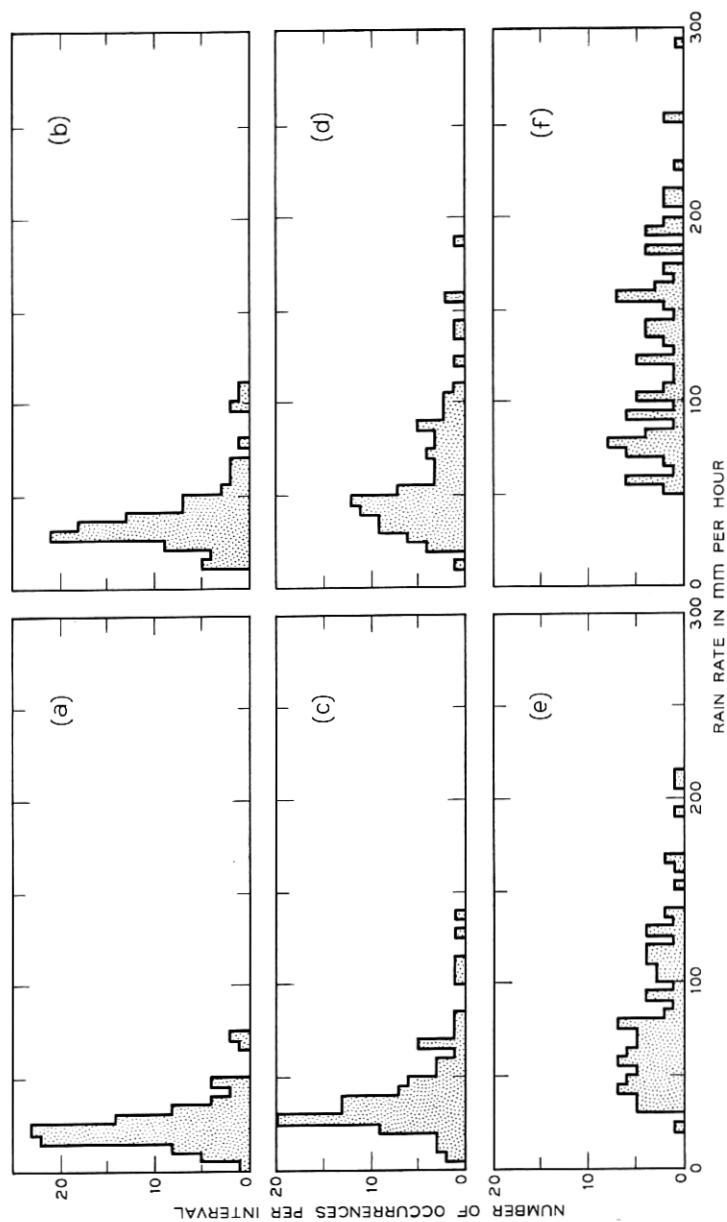


Fig. 13—Histograms showing the distribution of the (a) 95.0th, (b) 97.0th, (c) 98.0th, (d) 99.0th, (e) 99.5th, and (f) 99.9th percentiles from the rain-rate distributions of the 96 stations where the data from each station has been pooled over the 1967 season. Intervals are 5 mm per hour.

dispersion in the data may be obtained by examining the behavior of the distribution from individual stations, some of which are displayed in Fig. 9. The behavior of all the stations is summarized in histograms showing the frequency distribution of rain rates corresponding to 95, 97, 98, 99, 99.5 and 99.9 percent for the data from each station pooled over all the rainfalls. These histograms are presented in Fig. 13. All 96 stations were used. While 11 stations were operational only about half of the time and four were operational less than one-third of the time, inspection of the distributions from some of these stations indicates that an assumption of random loss of observations is not unreasonable. Summary statistics comprising the value of rain rate from the grand pool, together with mean, upper, and lower 25 percent points (quartiles), and the interquartile range (spread of the middle 50 percent) of the histograms, are given in Table II.

The overall rain-rate percentiles are larger than the means of the corresponding histograms. The amount by which they are larger increases as the overall percent (and thus the rain rate) increases until the largest overall percentile corresponds to the upper quartile of the histograms. This indicates that most of the high-rain-rate data came from only a few stations. The interquartile ranges are between 50 and 75 percent of the means of the histograms and tend to increase with rain rate. The histograms of the 99th and higher percentiles suggest a bimodal distribution, as if the data contained a component of very high rain rates which (because it is infrequent) is never observed at a large number of the stations. Thus Fig. 13 indicates a substantial sampling uncertainty above the 99th percentile; this uncertainty would be much greater for data collected at only a few stations.

While it is possible to make histograms similar to those of Fig. 13

TABLE II—STATISTICS ON THE DISTRIBUTION OF VARIOUS PERCENTILES FROM THE RAIN-RATE DISTRIBUTION

Percent of rain rate distribution	Rain rates in mm per hour				
	Value from the grand pool	The 96 stations, each station pooled for the season			
		Lower quartile	Upper quartile	Interquartile range	Mean
95.0	28	17	29	12	25
97.0	34	22	38	16	32
98.0	41	27	47	20	41
99.0	62	37	75	38	59
99.5	91	50	109	59	81
99.9	162	79	160	81	127

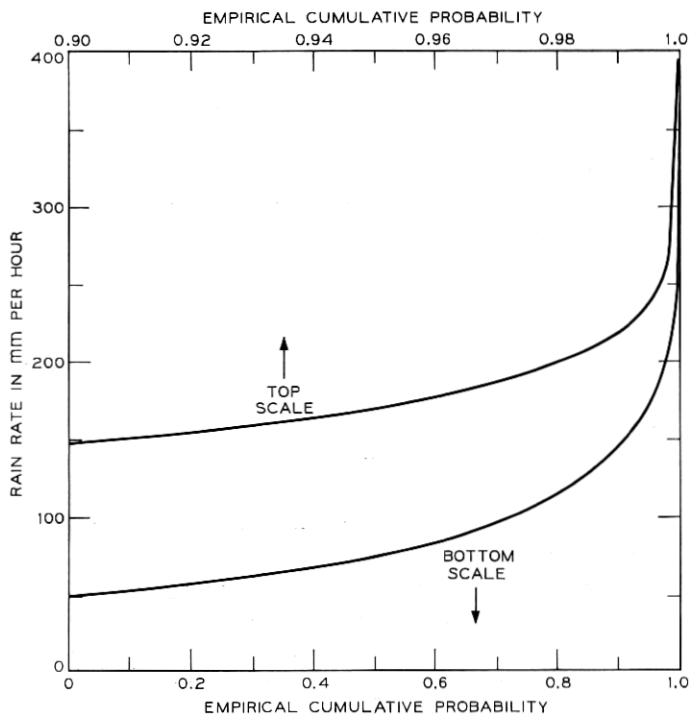


Fig. 14—Cumulative distribution of the grand pool subject to the condition that the rain rate is greater than 50 mm per hour. The data base contains 48,481 observations.

for other poolings of the data (say across stations for each rainfall) this has not been done because the rain-rate percentiles are affected by the dilution of the high-rain-rate data when periods of light rain are included. As a result of the selection procedure (Section II) this dilution varies greatly (by a factor of five) from rainfall to rainfall. Thus percentiles for different rainfalls are not really comparable, but percentiles from poolings across rainfalls are. This remark does not apply to the data when the condition in Section 4.2 is applied.

4.2. *Conditioned Probabilities*

All the probabilities in this section are based on the condition that the rain rate is greater than 50 mm per hour. This simple condition may be applied straight-forwardly to the point-rain-rate data because in arriving at the point-rain-rate statistics each data point is

treated as independent, and no relationships in time or space are taken into account. The 50-mm-per-hour condition is chosen because our selection procedure accepts virtually all data meeting this criterion. Thus at the cost of severely limiting the range of rain rates examined, one may work with probabilities that are well-defined and independent of the subjective criteria used to determine the beginning and end of periods of "interesting" rain.

The frequency distribution of rain rates greater than 50 mm per hour (in 2 mm per hour intervals) is given by the part of Fig. 11 which lies above this value. The corresponding empirical cumulative distribution appears in Fig. 14, which shows the full range from 0 to 1.0, as well as the upper tail on an expanded scale from 0.9 to 1.0. Once more there is no obvious overall correspondence between the data and the more common distributions. We have however fitted the function

$$N = \frac{1}{2}\alpha \exp(-\beta R), \quad R \geq 65 \text{ mm/hr}, \quad (1)$$

where N is the number of observations in a 1 mm per hour interval, R is the rain rate, and α and β are fitted coefficients, to the tail of the frequency distribution of Fig. 11. The fitting procedure was weighted least squares, with an *ad hoc* weighting that emphasizes the goodness

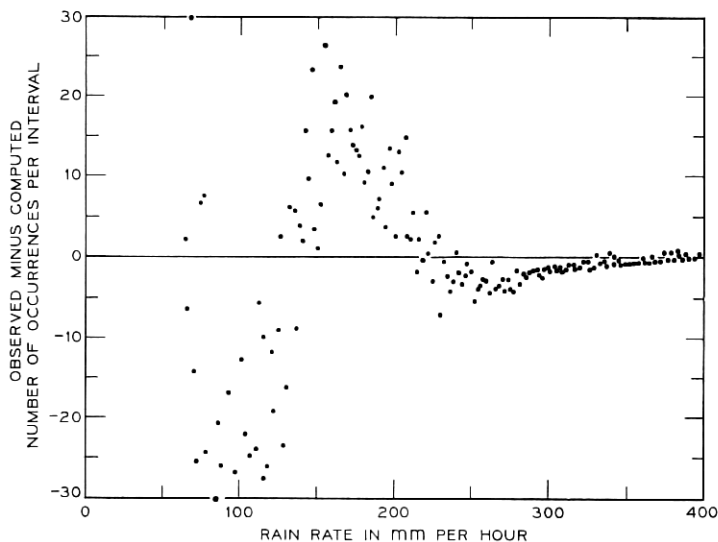


Fig. 15 — Residuals from the fit to the tail of the frequency distribution. The interval is 1 mm per hour.

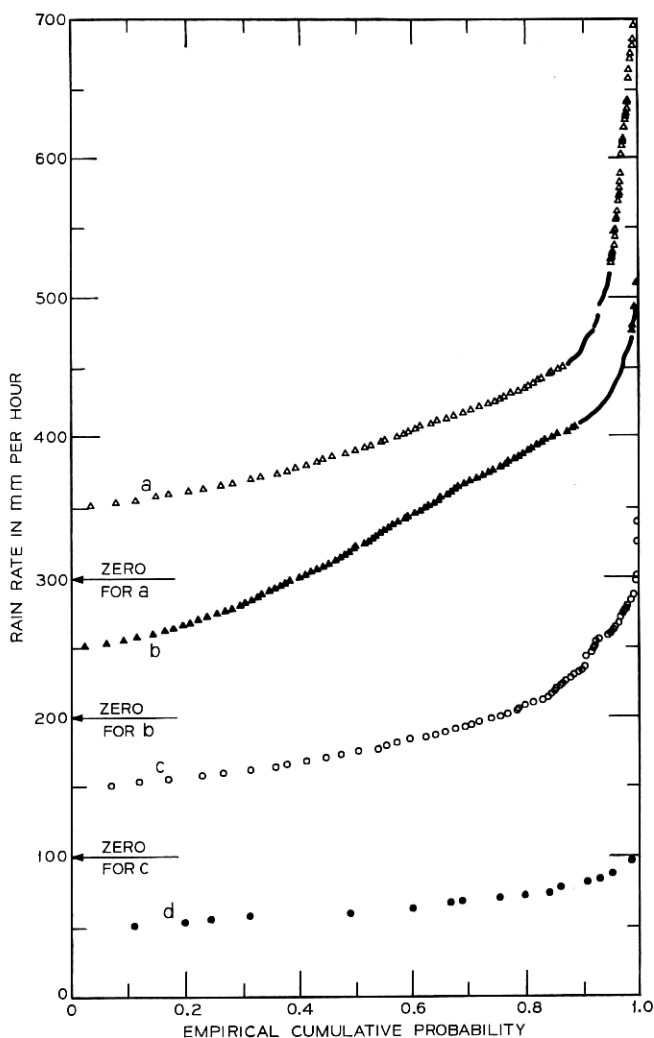


Fig. 16—Distributions of rain rates at each of four stations pooled for the 1967 season subject to the condition that the rain rate is greater than 50 mm per hour. Adjacent curves have been translated vertically by 100 mm per hour to increase the clarity of the plot. Curves *a*, *b*, *c*, and *d* are for stations 29, 1, 98, and 76, respectively and the respective data bases contain 1384, 1817, 415, and 45 observations.

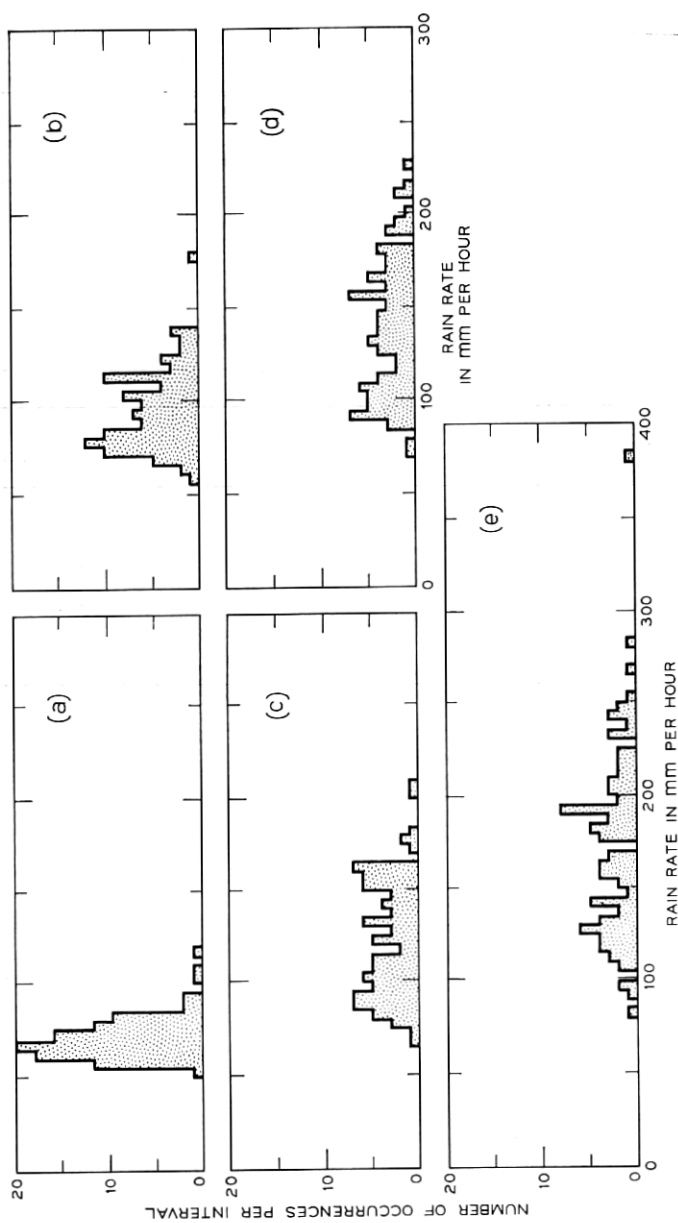


Fig. 17—Histograms showing the distribution for (a) 50.0th, (b) 75.0th, (c) 90.0th, (d) 95.0th, and (e) 99.0th percentiles from the rate-rate distributions of the 96 stations, for rain rates greater than 50 mm per hour. Intervals are 5 mm per hour.

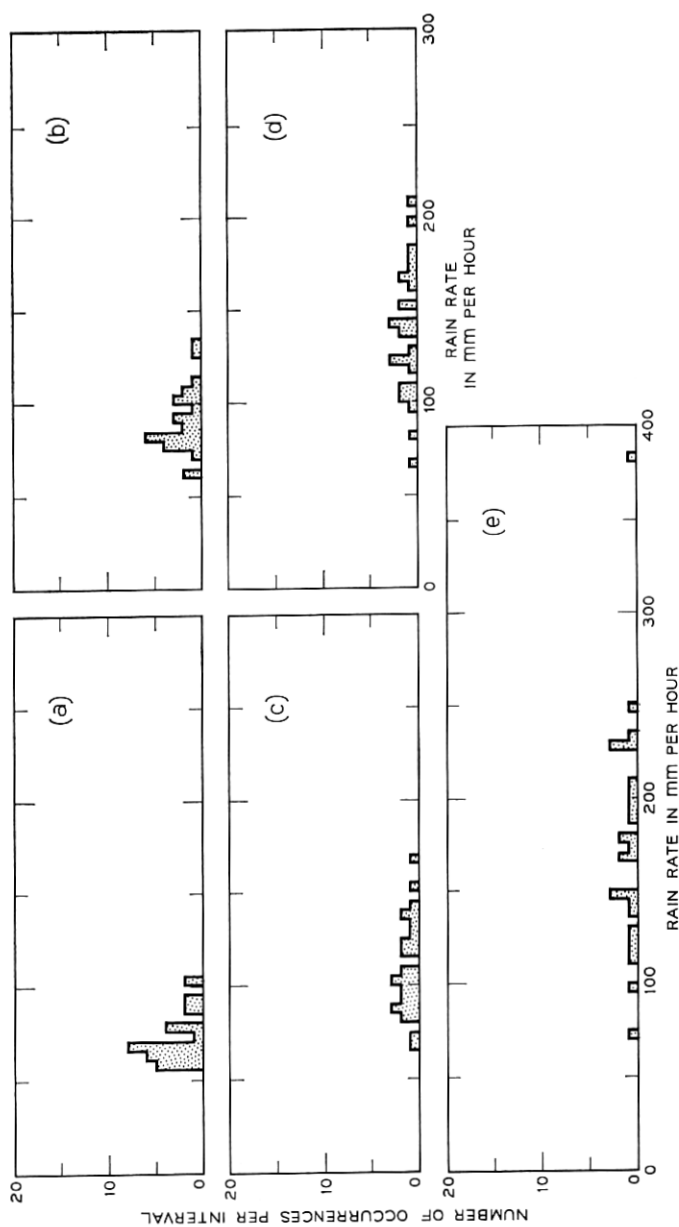


Fig. 18—Histograms showing the distribution for (a) 50.0th, (b) 75.0th, (c) 90.0th, (d) 95.0th, and (e) 99.0th percentiles from the rain rate distributions of the 27 rainfalls, for rain rates greater than 50 mm per hour. Intervals are 5 mm per hour.

TABLE III—STATISTICS ON THE DISTRIBUTION OF VARIOUS PERCENTILES FROM RAIN-RATE DISTRIBUTIONS*

Percent of rain rate above 50 mm per hour	Rain rates in mm per hour									
	Value from the grand pool	The 96 stations, each pooled for the season				The 27 rainfalls, each pooled over all stations				Mean
		Lower quartile	Upper quartile	Inter- quartile range	Mean	Lower quartile	Upper quartile	Inter- quartile range		
50	74	65	78	13	73	64	77	13	71	
75	106	80	112	32	95	81	105	24	93	
90	148	97	152	55	125	100	139	39	121	
95	174	110	167	57	142	108	167	59	139	
99	222	132	204	72	174	141	208	67	179	

* For rain rates greater than 50 mm per hour.

of fit for the higher rain rates. The estimates a and b of α and β are 7754 and 0.02408, respectively. The properties of the model are best displayed by the residuals, which are plotted against rain rate in Fig. 15. The fit is seen to be moderately good for rain rates greater than 200 mm per hour, and fair for rain rates between 65 and 200 mm per hour. Below 65 mm per hour the function and the observations diverge rapidly.

The model provides a useful smoothing function between rain rates of 65 and 400 mm per hour, and also represents the best currently available extrapolation of these data above 400 mm per hour. However, the systematic behavior of the residuals shows that an exponential function is not really a suitable distributional description of the data.

Figure 16 shows empirical cumulative distributions for rain rates above 50 mm per hour for the same four stations whose unconditioned cumulative distributions appear in Fig. 9. This once more illustrates the large differences among stations. Histograms of the frequency with which certain percentiles fall in various rain-rate intervals are once more used to quantify the dispersion in the data. Figure 17 contains the results obtained when data are pooled across the season's rainfalls for each station; Figure 18 contains analogous information for data pooled across stations for individual rainfalls. The outliers on the two 99 percentile histograms are both caused by the same event, a very high rain rate at one station during one rainfall (see Section 2.2). Smaller poolings generally contain too few data points above 50 mm per hour to be informative. The results are summarized in Table III.

Most notable in Table III is the great similarity between the distributions of upper-tail percentiles for the two poolings. There is no obvious reason why this should be so, and the result may be peculiar to this set of data. As in Table II, the interquartile range increases with the mean (ranging in value from about 20 to about 40 percent of the mean) indicating the increasing uncertainty associated with the small sample of high rain rates. Above 90 percent, the value from the grand pool lies above all other values in the same line of the table, showing once again the influence of a few local cloudbursts.

V. SPATIAL AND TEMPORAL RELATIONSHIPS

Figure 2 shows the general space-time relationship of rain rates: When it is raining heavily at a station it is likely to be raining heavily

at nearby stations, and it is also likely to be raining heavily a short time later. This section quantifies these general observations.

If the time series representing the rain-rate observations were stationary the relationships would be given by the cross correlation as a function of distance and the autocorrelation as a function of lag (time difference). However as indicated in Section 3.1 the time series are not even piecewise stationary in the regions containing the high rain rates. As the situation is more complex than that represented by stationary time series, the statistical summary is somewhat less succinct. The spatial relationships are described in terms of the distribution of the rain rate at a point *there* given certain conditions of rain at point *here*. The distance between *here* and *there* and the rain rate at point *here* are varied as parameters. The temporal relationships are described analogously in terms of distributions of the rain at time *hence* given certain conditions of the rain at time *now*. The lag, that is, the interval between *now* and *hence*, and the rain rate at time *now* are varied as parameters.

5.1 *Spatial Relationships*

The frequency distributions underlying this presentation were collected as follows. The readings from each scan of the network were assumed to be simultaneous. The very rapid variations in rain rate (Fig. 7) and lack of detailed correlation over distances of the magnitude of the station separations (Section III) provide some statistical justification for this treatment. Each station (*here*) was then paired with every other station (*there*), giving $(96)^2$ pairs. Each pair was classed as belonging to one of 238 possible histograms according to the value of the rain rate *here*, which is sorted into 17 intervals, and the nominal distance (stations are assumed to be at the center of the grid square) between the stations, which is sorted into 14 separation intervals. Results are labeled with the mean of the rain-rate interval, and the mean of the nominal distances in the separation interval. For the rain rate *here* the interval is increased as the rain rate increases; for the separation, the interval is about 20 percent of the mean. Once the histogram was selected, the frequency of the rain rate *there* was accumulated in intervals of 5 mm per hour.

The 110 hours of rain data contain about 4×10^4 scans of the network and each scan yields about 10^4 pairs giving a total data base of 4×10^8 pairs. The data are accumulated separately for each of the 27 rainfalls and then pooled for many presentations.

In spite of the large number of pairs, there are relatively few pairs at the high rain rates. Thus the high-rain-rate results show noticeable sampling fluctuations and remain sensitive to the data-screening procedures. Probabilities involving rain rates of less than 50 mm per hour are distorted by the data-selection processes as already discussed.

Three typical histograms (from the set of 238) are shown in Fig. 19. Figures 19a and b show the smooth behavior produced by the large numbers of pairs at low and intermediate rain rates; Figure 19c shows the fluctuations that result from the relatively small numbers of pairs at high rain rates. These fluctuations are reflected in occasional irregular behavior seen in other displays in this section.

Three sets of empirical probabilities for the rainfall of July 25, 1967 are shown in Fig. 20. The curves give the probability of the rain rate *there* being less than R mm per hour when the rain rate *here* falls in the interval indicated on the curve. Figure 20a shows that when it is raining lightly at station *here*, it is unlikely to be raining heavily at stations 1.3 km away. As the rain rate at station *here* increases it becomes increasingly more likely to be raining heavily at nearby stations.

The strength of this relationship decreases with increasing separation until at about 8 km separation the probability of observing a rain rate less than R mm per hour is largely independent of the rain rate observed at the distant station; this is demonstrated by the contraction of curves in Fig. 20b into a narrow band. As the separation increases still further the curves reverse (that is, the probabilities associated with curve 5 ± 5 mm per hour are generally lower than those associated with curve 200 ± 20 mm per hour for the same R . This may be seen on Fig. 20c which shows the results for stations separated by 12 km. The reversal means that if it is raining heavily at station *here* it is less likely to be raining heavily at a station 12 km away than if it were raining lightly at station *here*. This effect is noticeable at separations up to 16 km, but few pairs of stations are so widely separated (the diagonal measurement of the network is 18 km). This behavior suggests that regions of heavy rainfall have a "range of influence" of about 16 km (twice the 8 km distance corresponding to Fig. 22b) and that the "centers" of these regions are separated by more than the dimension of the network, that is, more than about 16 km.

The probabilities may also be plotted against distance, with the rain-rate-*here* intervals as parameters and different values of R appearing on different plots. Figures 21a shows the behavior of the prob-

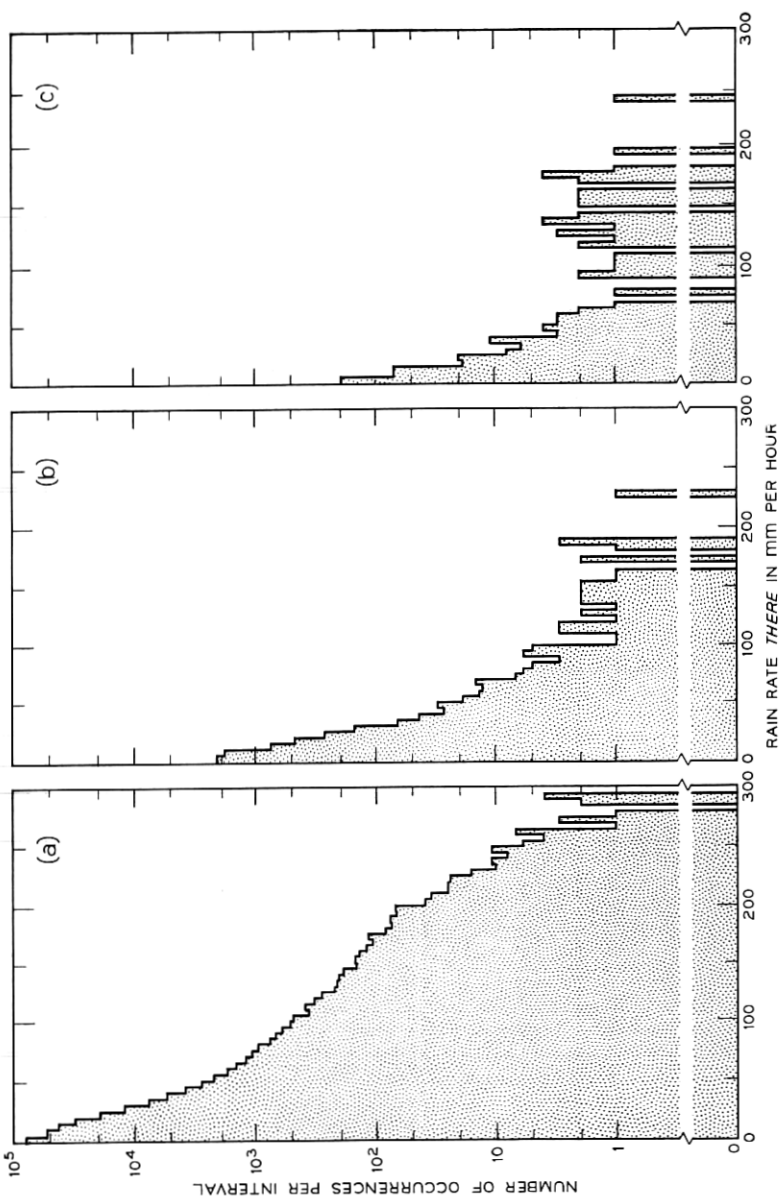


Fig. 19 — Histograms showing the distribution of the rain rate *there* for three separations, rain-rate-*here* combinations pooled over all rainfalls. (a) 6.3 ± 0.6 km, 45 ± 5 mm per hour; (b) 14.1 ± 1.4 km, 155 ± 15 mm per hour; (c) 6.3 ± 0.6 km, 300 ± 20 mm per hour. The interval of the histogram is 5 mm per hour.

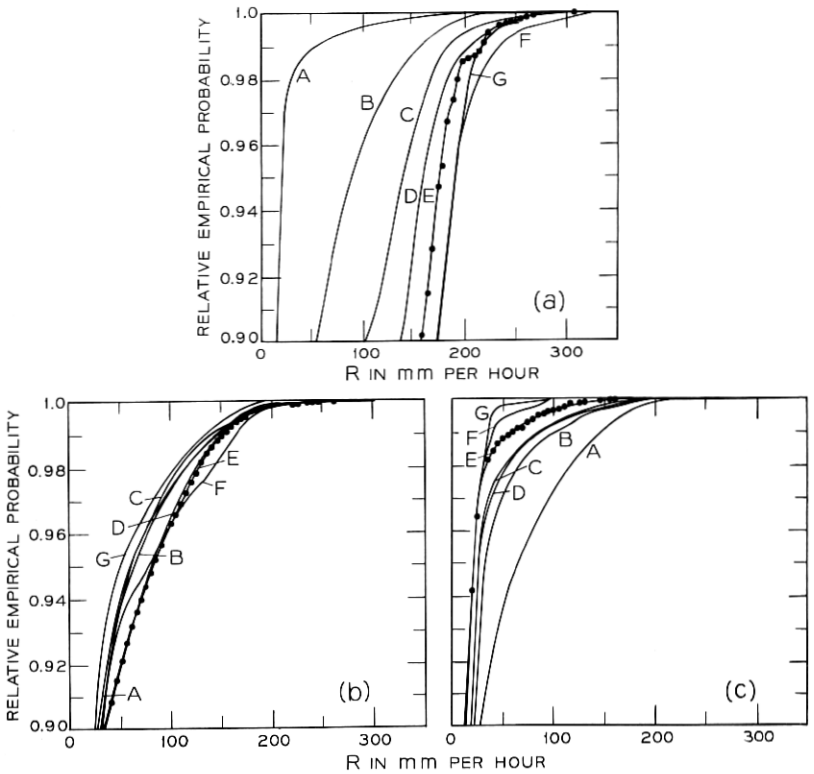


Fig. 20 — Relative empirical probability that the rain rate *there* is less than R mm per hour for various intervals of the rain rate *here* during the rainfall of July 25, 1967, for stations separated by (a) 1.3 ± 0.2 , (b) 7.5 ± 0.8 , and (c) 12.3 ± 1.2 km, respectively. Points shown on this and following plots indicate actual calculated values.

Rain Rate *Here* (mm per hour)

	A	B	C	D	E	F	G
Min.	0	20	40	70	110	170	200
Max.	10	30	50	90	140	200	240

ability that the rain rate *there* is less than 35 mm per hour (the threshold) as a function of distance for various intervals of the rain rate *here*. The probability that the rain rate *there* is less than 35 mm per hour increases monotonically with distance when the rain rate *here* is greater than 30 mm per hour. For rain rates *here* that are less

than 30 mm per hour, the probability for rain rate *there* being less than 35 mm per hour goes through a minimum at a separation that depends on the rain rate *here*. Figure 21a also displays the crossover of the probability curves at a separation of about 8 km. For smaller separation it is less probable for the rain rate *there* to be below 35

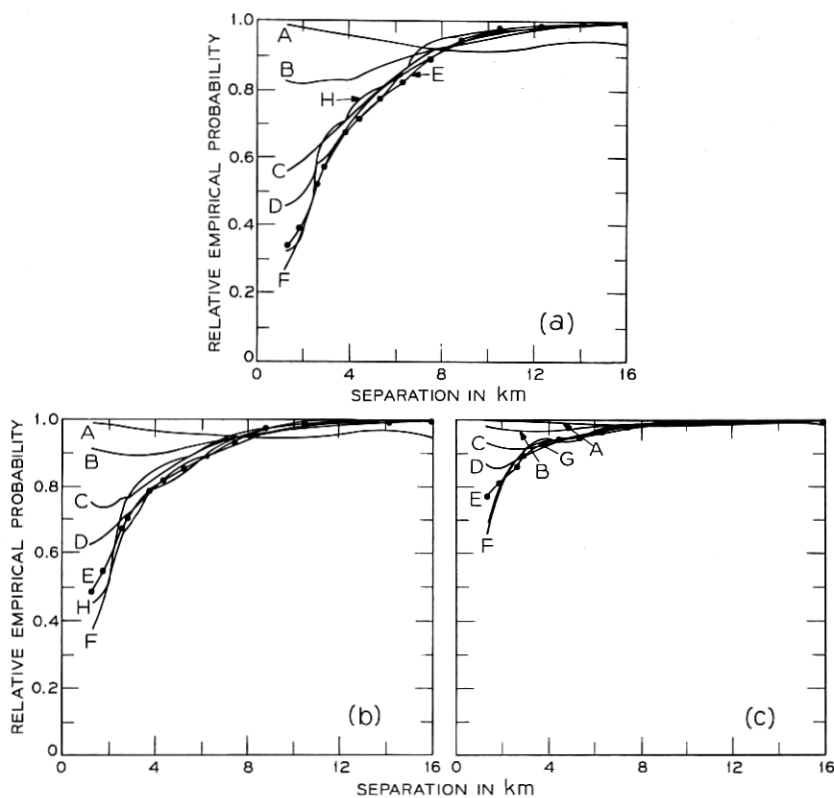


Fig. 21 — Relative empirical probability that the rain rate *there* is less than R as a function of separation with various intervals of the rain rate *here* as a parameter, for the rainfall of July 25, 1967. R is (a) 35, (b) 60, and (c) 125 mm per hour.

Rain Rate *Here* (mm per hour)

	A	B	C	D	E	F	G	H
Min.	0	20	40	70	110	170	200	240
Max.	10	30	50	90	140	200	240	280

mm per hour if it is raining heavily *here* than if it is raining lightly *here*. At separations exceeding 8 km, the inverse is true. Figures 21b and 21c show the analogous plots for rain rates less than thresholds of 60 mm per hour and 125 mm per hour, respectively. The pattern is the same, and if 70 and 125 mm per hour respectively are substituted for 30 mm per hour in the previous discussion it remains applicable. In particular the crossover still occurs at 8 km separation. As might be expected, the probability of finding the rain rate *there* lower than the threshold increases with increasing threshold.

For completeness we include four plots in Fig. 22, which show the behavior with separation of the probability that the rain rate *there*

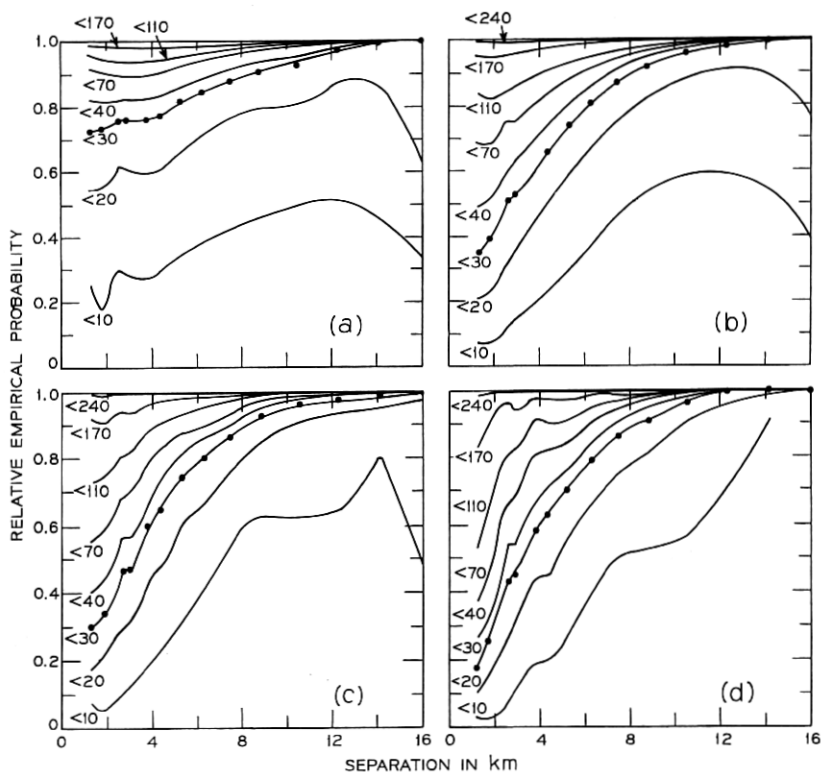


Fig. 22 — Relative empirical probability that the rain rate *there* is as indicated on the curve as a function of separation for various intervals of the rain rate *here* for the rainfall of July 25, 1967. The intervals of rain rate *here* are (a) 20 to 30, (b) 50 to 70, (c) 90 to 110, and (d) 200 to 240 mm per hour. Curves are labeled in units of millimeters per hour.

is less than certain values for the rainfall of July 25, 1967. Each plot is for a given interval of the rain rate *here*. The behavior of the curves is generally smooth without any abrupt changes of slope. The direction of curvature and position of the maxima and minima depend on the particular interval of rain rate *here* displayed. We are unable at this time to evaluate the significance of the fine structure that appears occasionally.

The patterns of the other heavy rainfalls have been examined and are similar to that of the July 25, 1967 rainfall. In particular the "range of influence" of the rainfalls is approximately 16 to 20 km. When the data from all 27 rainfalls (or even for the nine heavy rainfalls) are aggregated, the differences among the rainfalls conceal the individual patterns, and the reversal of the order of the curves that take place between Figs. 20a and 20c disappears. However, the concept of a "range of influence," that is, a separation at which the rain rate *there* is relatively independent of the rain rate *here*, is still valid for the aggregate when the rain rate *there* is below about 125 mm per hour. The "range of influence" is about 20 km.

Of particular engineering interest is the probability that the rain rate is *greater* than a particular value at two stations simultaneously. Results in this form are presented in Fig. 23, which is a plot of the joint probability that the rain rate is greater than R mm per hour at both of two stations against the separation in km. These probabilities change smoothly with distance, and have a minimum at a separation of about 12 km for rain rates with sufficient data to establish the shape of the curves. The joint probabilities were also computed for each of the 27 rainfalls separately; the upper quartiles of the 27-rainfall distributions (a distribution for each separation) are shown for minimum rain rates of 50 and 110 mm per hour by the crosses in Figs. 23a and b, respectively. The lower quartiles are indistinguishable from the abscissa on the scale of the plots. The greater-than-50-mm-per-hour curve lies close to its upper quartiles; the greater-than-110-mm-per-hour curve lies far above the corresponding upper quartiles. These results again demonstrate the sensitivity of the upper tail of the aggregated distribution to a small number of "cloudbursts." Although the plots for individual rainfalls differ substantially in behavior and statistical stability, about one third of the rainfalls display a minimum in the joint probabilities at a separation of about 12 km.

Because of the bias introduced by the method of selecting the data, the numerical values of joint probabilities for rain rates below 50 mm per hour are distorted relative to those for rain rates above that value.

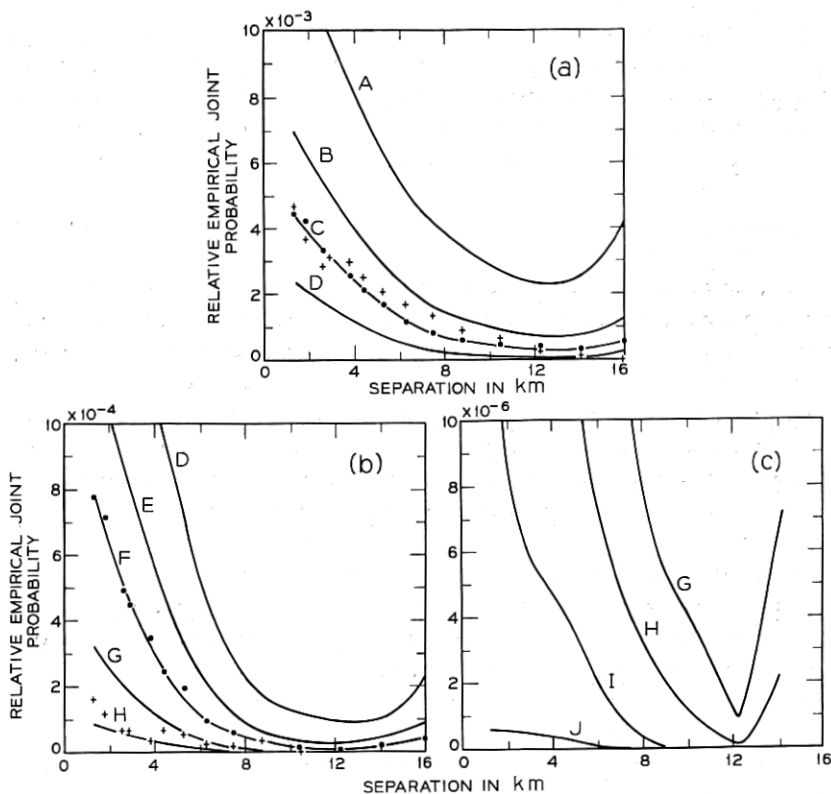


Fig. 23—Relative empirical joint probability that the rain rate at both stations exceeds the value with which the curve is labeled versus separation in kilometers. Crosses in (a) and (b) indicate the upper quartiles of the distributions of joint probabilities that the rain rate is greater than 50 and 110 mm per hour, respectively, when the 27 rainfalls are considered separately.

Minimum Rain Rate (mm per hour)

A	B	C	D	E	F	G	H	I	J
30	40	50	70	90	110	140	170	200	240

5.2 Temporal Behavior

The frequency distributions on which descriptions of the temporal behavior are based are accumulated in a manner completely analogous to that used to collect the spatial distributions. Instead of distance intervals, time intervals (lags) are used, and each station is "paired"

only with itself. Thirteen lags were used and there are about 100 stations and 40,000 scans. This leads to 5×10^7 pairs which (as 17 intervals of rain rate *now* are used) fall within $13 \times 17 = 221$ histograms. The value of the rain rate *hence* is sorted into intervals of 5 mm per hour to form the histograms. All the data presented in this section are pooled across the 27 rainfalls.

The probability that the rain rate *hence* is less than R mm per hour for various intervals of the rain rate *now* is given by the curves in Fig. 24. While the curves shift toward lower rain rates at the longer lags, the pattern remains essentially the same with time. If it is raining heavily *now* it is more likely to be raining heavily a short time *hence* than if the rain *now* were light. Curves G and H, which show rain rate *now* greater than 240 mm per hour are much more variable than the others because of the small number of observations in this range.

Figure 25 shows curves giving the empirical probability that the rain rate *hence* will be less than the value indicated on the curve versus lag for rain rates *now* in the ranges 20 to 30, 50 to 70, 90 to 110, and 200 to 240 mm per hour, respectively. These plots show that the probability that the rain rate will fall at or below its present value is about 0.7 (increasing to 0.9 at the highest rain rate) and relatively independent of the lag (the increase is about 0.1 over the 360 seconds).

The tendency of the curves to move in the direction of the line corresponding to a probability of 0.8 is noticeable at the small lags. This means that the probability, that the rain rates a short time *hence* are very different from the current rain rate, is small. The relationship has both short and long term components. The short term component dies down within about five minutes, as evidenced by the decrease in the slopes of the curves with increasing lag, and is probably associated with the fine structure of the precipitation. The long term component is present after five minutes as evidenced by the different values for the same curves on the four plots. This component is probably associated with some average property of the local precipitation. Except for the highest rain rates (Fig. 25d), the spacing between the curves after a 360-second lag is larger for the lower rain rate *hence* limits (in spite of the fact that the rain rate interval is also smaller for the lower rain rate *hence* limits), indicating that the probability density for the rain rate *hence* peaks at the low rain rates after a lag of 360 seconds for values of the rain rate *now* below about 200 mm per hour.

An interesting feature appears in Fig. 25d. All the curves with rain

rate *hence* limits of less than 240 mm per hour show a dip at a lag of 20 seconds, and the curve labeled 240 mm per hour shows a double dip. This indicates a tendency for the heavy rain rates to have a fine structure oscillation with a period of about 20 seconds. The conjecture is confirmed by examination of plots (not presented here) for rain rate *now* intervals with lower bounds greater than 200 mm per hour.

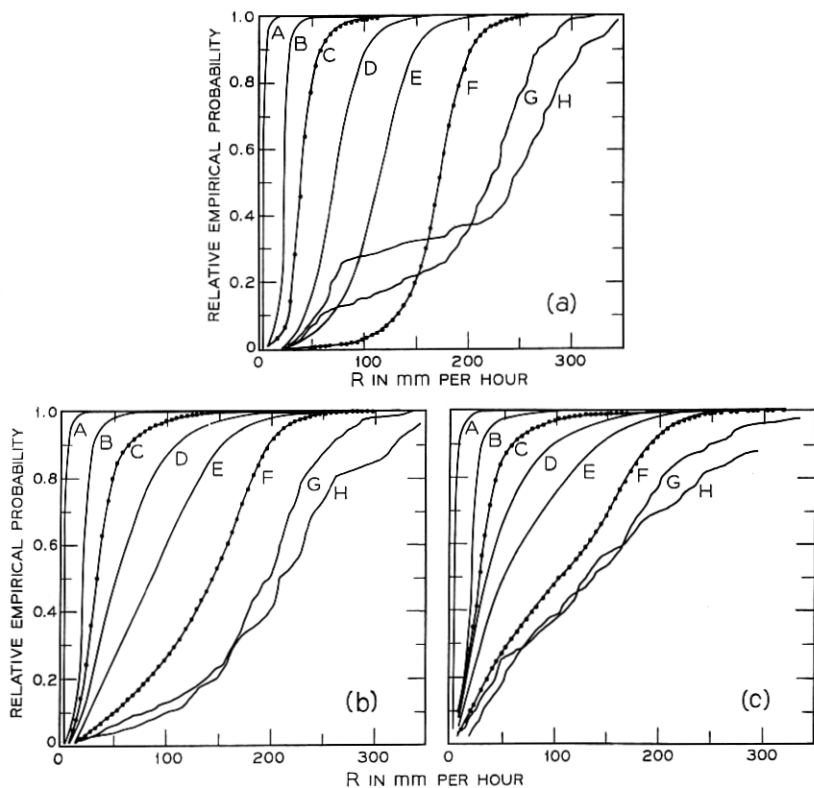


Fig. 24—Relative empirical probability that the rain rate *hence* is less than R mm per hour for various intervals of the rain rate *now* for lags of (a) 10, (b) 120, and (c) 360 seconds.

Rain rate (mm per hour)

	A	B	C	D	E	F	G	H
Min.	0	20	40	70	110	170	240	290
Max.	10	30	50	90	140	200	280	320

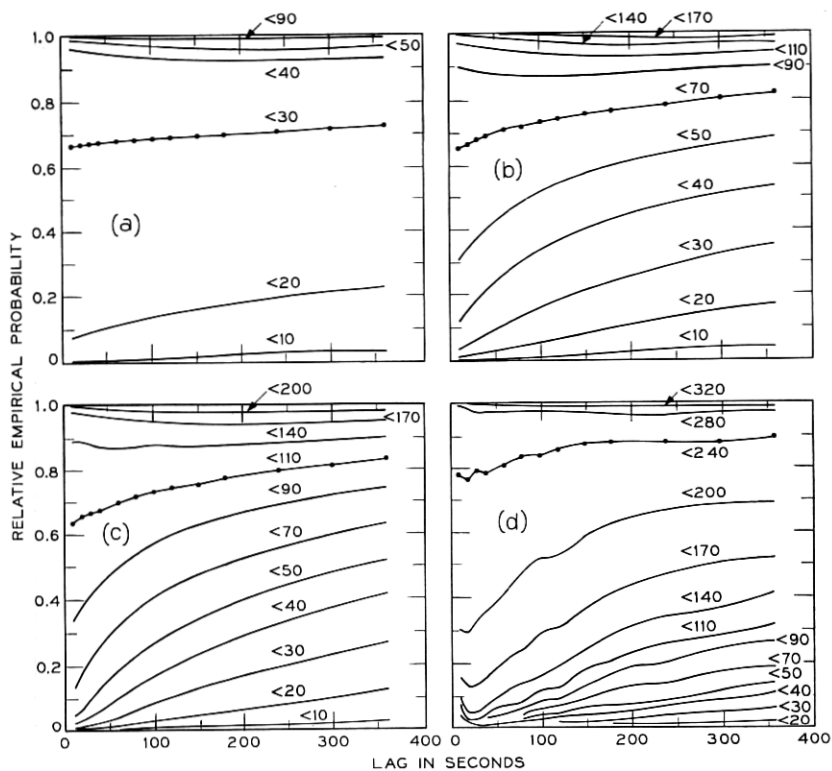


Fig. 25 — Relative empirical probability that the rain rate *hence* is as indicated on the curve as a function of lag for intervals of the rain rate *now* of (a) 20 to 30, (b) 50 to 70, (c) 90 to 110, and (d) 200 to 240 mm per hour. Curves are labeled in units of mm per hour.

Another presentation, Fig. 26, plots the empirical probability that the rain rate *hence* is less than R mm per hour (the threshold) for various intervals of the rain rate *now* as a parameter. In agreement with the previous plots, one notes that the probability that the rain rate exceeds its present value decreases as both lag and rain rate *now* increase and is smaller than about 0.2. There is a tendency for the rain-rate intervals below the threshold (for example, the 40 to 50 mm per hour curve in Fig. 26b) to be slightly concave upward, with a minimum at about 180 seconds. This might be associated with a characteristic growth time for showers; but the averaged effect is very weak.

Lastly, the empirical probability that the rain rate exceeds a given

value both now and after various lags is given in Fig. 27 for a number of rain rates. The probabilities are seen to decrease slowly and monotonically with time up to lags of 360 seconds. The crosses in Fig. 27 are analogous to those in Fig. 23 and indicate the upper quartiles of the distributions obtained (a distribution for each lag) when the joint probabilities of minimum rain rates of 50 and 110 mm per hour are calculated separately for each of the 27 rainfalls. (The lower quartiles are too close to zero to show on the plots.) That the joint probabilities in time are also very sensitive to small numbers of heavy-rain-rate events is demonstrated by the fact that the mean value is higher than the crosses for a large fraction of the data. However the

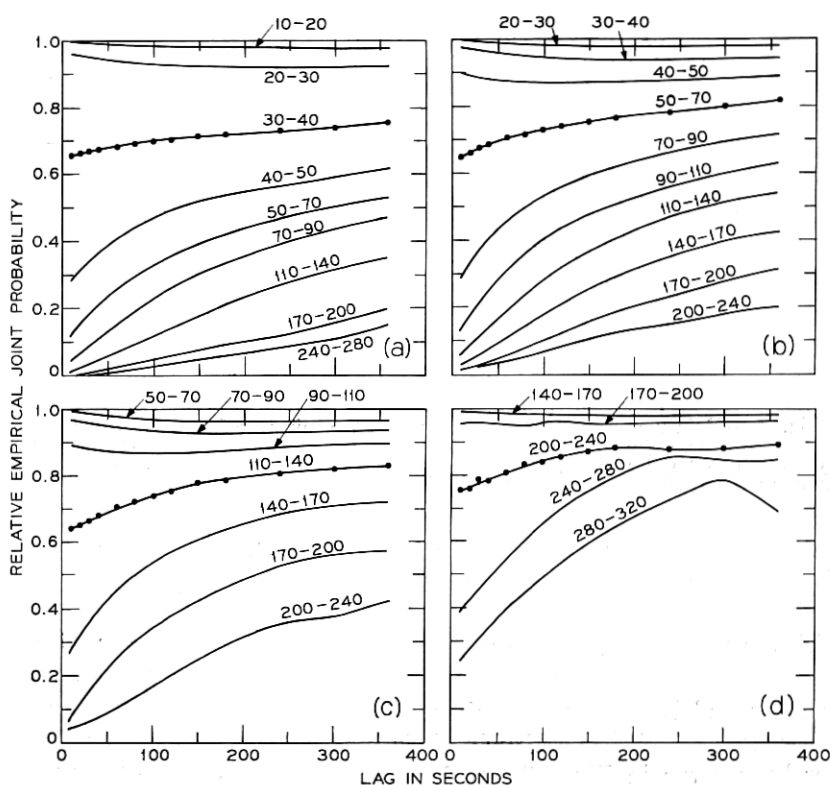


Fig. 26—Relative empirical probability that the rain rate *hence* is less than R mm per hour as a function of lag with intervals of the rain rate *now* as parameter. R is (a) 35, (b) 60, (c) 125, and (d) 220 mm per hour. Curves are labeled in units of mm per hour.

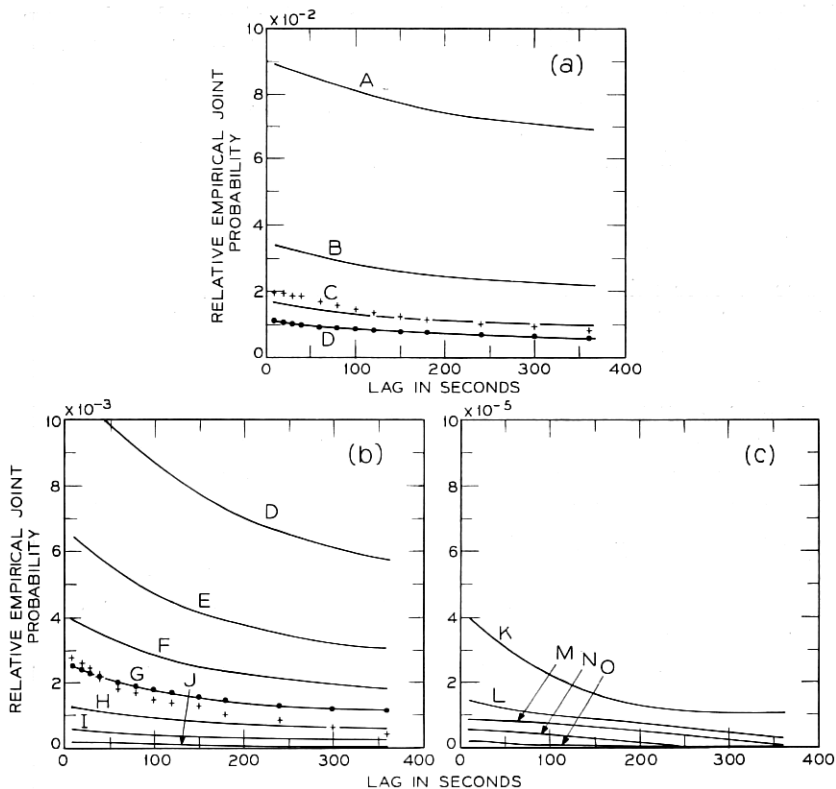


Fig. 27—Relative empirical joint probability that the rain rate exceeds the value with which the curve is labeled both *now* and *hence* (but not necessarily between the two times).

Minimum Rain Rate (mm per hour)

A	B	C	D	E	F	G	H
20	30	40	50	70	90	110	140
I	J	K	L	M	N	O	
170	200	240	280	320	360	400	

situation is less sensitive than in the case of joint probabilities in space (Fig. 23).

This presentation does not provide direct information on the probability that the rain rate will *remain* above a certain value for the time given by the abscissa. We have not attempted to provide sta-

tistics on the length of time the rain rate exceeds given values because (as shown in Fig. 7) the rain rate changes rapidly within the ten second sampling interval and the form of distribution of which the rain rates within the interval are a sample is not known.

The remark on sampling bias at the end of Section 5.1 applies to the temporal as well as the spatial probabilities.

VI. PATH STATISTICS

The average rain rate along the path between a microwave transmitter and receiver affects the amplitude of the received signal. This section presents the probability that the average rain rate exceeds a given value on a single path or on two paths simultaneously for a variety of conditions. We calculate the probabilities directly from the data without recourse to any of the statistics and intermediate results previously presented. Because of the very large amount of computation involved, it is essential to reduce the data base and regularize the geometry as much as possible. We chose to examine only the nine heavy rainfalls (see Section 2.3), and from these we selected only those 6725 scans during which at least one station showed a rain rate in excess of 50 mm per hour.

The geometry has been regularized by rectangularizing the network to 11 east-west rows of ten stations each. The averages of the values from the surrounding stations have been substituted for values missing because stations were inoperative (or fictional, such as the northeast corner) or where the values were judged to be outliers. Substituting for missing values has only a small effect (statistically) on the distribution of average rain rates but a large effect on the population of paths in various categories. This matter has been examined in some detail and, because of the configuration of heavy-rain-rate occurrences, it has been concluded that omission of paths containing missing observations introduces artifactual sampling variations that distort the results of the path analysis very seriously in unpredictable ways, and that this approach is unacceptable. The distortion introduced by using averages of surrounding values to estimate missing observations is that of generally lowering the probability of occurrence of high rain rates. As noted in Section 7.3, the effect is most severe on "single station" paths.

Omission of data from rows 1, 2, and 11 of the grid, in an attempt to make the grid rectangular without filling in corners, seriously biases the results by throwing out too much of the heavy-rain-rate data and

many of the long paths. As justified in Section V, stations were treated as if they were located at the centers of their grid squares and scanned simultaneously. The paths were taken as starting and ending at stations; in computing the average rain rate the values at the stations at the ends of the path were given the weight $\frac{1}{2}$ and the values at the intermediate stations, the weight 1.

The diameter of a rain gauge is only 30 cm and the distance between grid square centers is 1.3 km, so that much less than $1/4000$ of the path is sampled. As the rain rate has a great deal of fine structure and the number of samples per path is small (less than 11), these averages are themselves subject to very substantial fluctuations about the "true" path average.

We considered many possible paths and pairs of paths through the network. Because a single station is generally included in several paths of the same length and orientation, the separate path averages are not statistically independent. This factor has not been taken into account explicitly in our analysis.

6.1 *Parallel Paths*

Parallel paths are characterized by three parameters: length, separation and orientation. (To allow consistency of notation, length zero indicates paths which include only a single station and separation zero indicates a single path.) We consider four orientations: two square, north-south, and east-west; and two diagonal, northeast-southwest and northwest-southeast. A pair of parallel north-south paths of length 5.2 km and separation 2.6 km is indicated by the two dashed lines, $P_{11}-P_{12}$ and $P_{21}-P_{22}$ toward the western side of the grid in Fig. 1.

The relative probability that the rain rate on both paths exceeds the value on the abscissa is given by the ordinate for nine separations as parameters and six lengths of path as parts of Fig. 28 for nine heavy rainfalls in 1967. (The abscissa is then the minimum average rain rate for the pair.) The results for the north-south and east-west paths have been pooled. The largest number of pairs of paths entering into the data base for an individual curve is 1,338,275 (Fig. 28a curve *B*) the smallest is 80,700 (12 per scan, Fig. 28f curve *I*).

Except for a small amount of straggling in the lowest decade, the curves are smooth and well behaved. In order to display the small probabilities and high rain rates effectively, the logarithm of the probability has been used as the ordinate, and rain rate has been

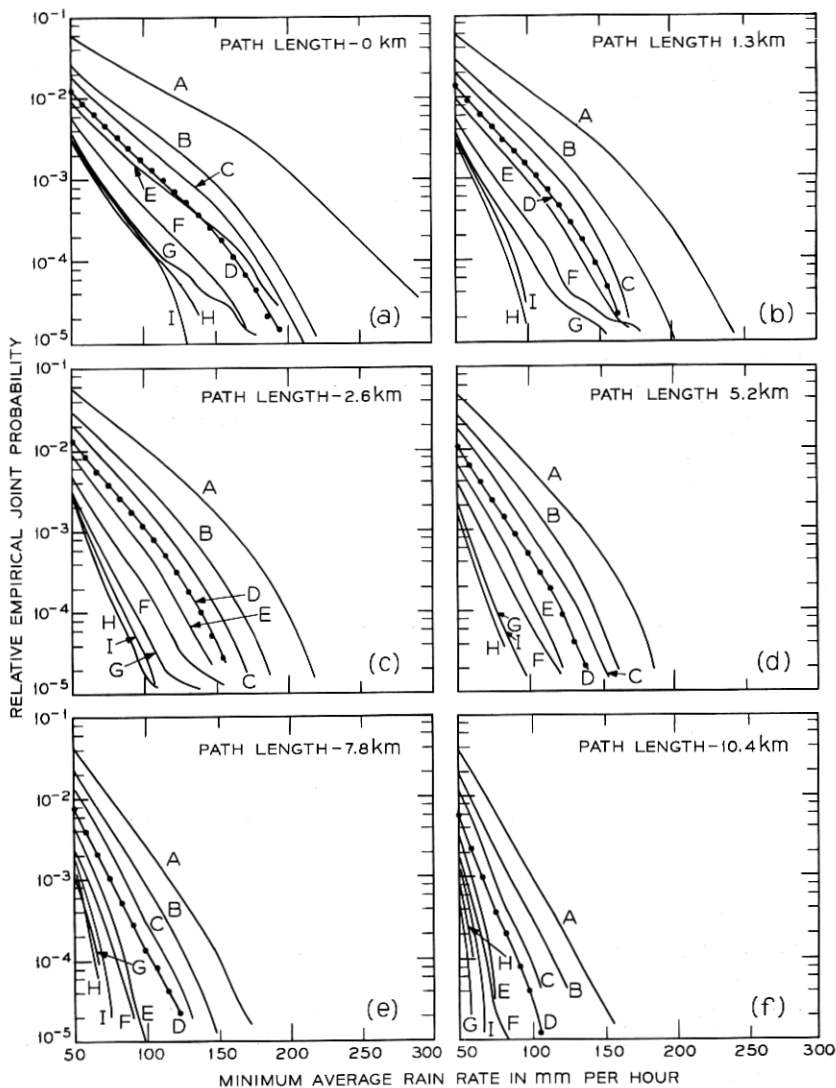


Fig. 28 — Relative empirical probability that the average rain rate on both of two parallel paths will exceed the value on the abscissa. Path lengths are (a) 0 km, (b) 1.3 km, (c) 2.6 km, (d) 5.2 km, (e) 7.8 km, and (f) 10.4 km. Points indicate actual calculated values.

Separation (km)

A	B	C	D	E	F	G	H	I
0	1.3	2.6	3.9	5.2	6.5	7.8	9.1	10.4

plotted linearly as the abscissa. The curves are generally quite straight on this semilog plot, but tend to droop somewhat at the high rain rates. This indicates that for purposes of extrapolation at the high rain rates the customary log-probability versus log-rain-rate plot is unsuitable and that improvement should be sought in the opposite direction of scaling (for example, by raising the rain rate to a power slightly greater than 1). For paths up to 10 km long, the joint probabilities decrease with increasing separation until a minimum is reached at a separation of about 9 km.

For larger separations the joint probability increases again. This result was observed for single stations in Section V. Comparison among the parts of Fig. 28 shows that for fixed probabilities the minimum average rain rate decreases as the path length increases at all separations. Thus the product of the minimum average rain rate and the path length increases more slowly than the path length itself. (This product is approximately proportional to the quantity of water along the path.)

Results from the diagonal paths generally confirm the results from the square paths. Pooling has not been carried out, however, because of the differences in path lengths and separations, and the fact that the absence of long diagonal paths at large separations tends to weight the stations in the center of the network very heavily.

The curve for zero length and zero separation corresponds to single stations, that is, point-rain-rate statistics. Similarly the curves for zero length and various separations are subsets of the space pairs of Section V. However, the data bases are not the same; missing observations have been estimated for the subset of the data used for the path analysis, making direct comparison of these results very difficult. Section VII presents a rough comparison and a discussion of the differences.

In order to minimize edge effects (the stations on the edges are included in fewer paths, that is, less heavily weighted than the central section even for the square paths) the maximum length of path considered should be short compared with the dimensions of the network. This is certainly not the case for the paths more than 5 km long. Furthermore, the results appear quite sensitive to the particular patterns of rainfall observed. The relative values, while subject to some distortion introduced by the estimation of missing points, are probably fairly good; however, we do not believe that the numerical values of the results in Fig. 28 should be considered better than an order of magnitude for the high rain rates at this stage of the analysis.

This conservative view is vindicated by the large dispersion among the results for the four orientations considered individually. For example, Fig. 29 shows the results for all parallel paths about 7.5 km long and 3.7 km apart pooled over the nine heavy rainfalls categorized by orientation. The "orientation" effect exceeds a factor of 2 at a minimum average rain rate of 60 mm per hour and a factor of 10 above about 100 mm per hour. Part of the dispersion may be real, that is, a result of systematic factors (such as the direction of the prevailing winds or the location of local updrafts) and part may be a result of differences in the population of stations which go to make up the different sets of paths. The dispersion also indicates the additional caution with which results from single lines of stations should be viewed.

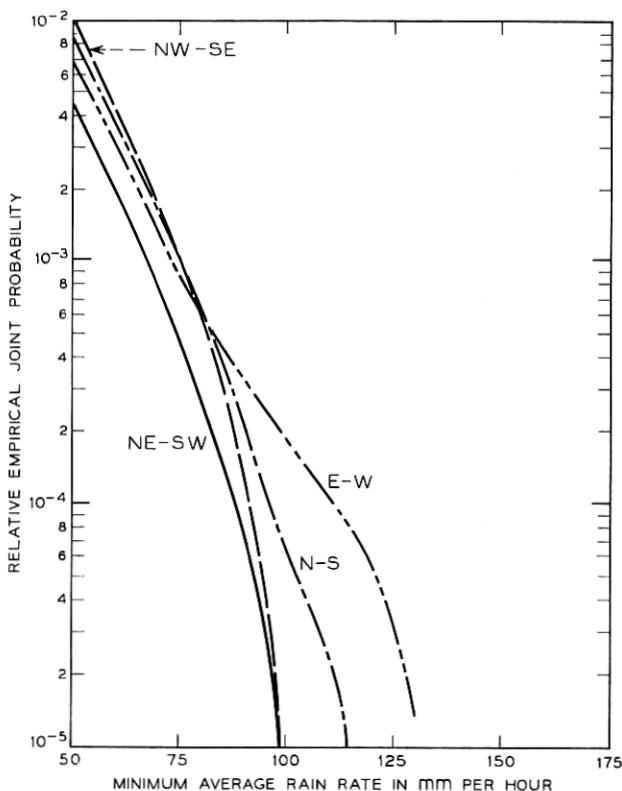


Fig. 29—"Orientation" effect. The four curves represent the two square and two diagonal orientations in the network.

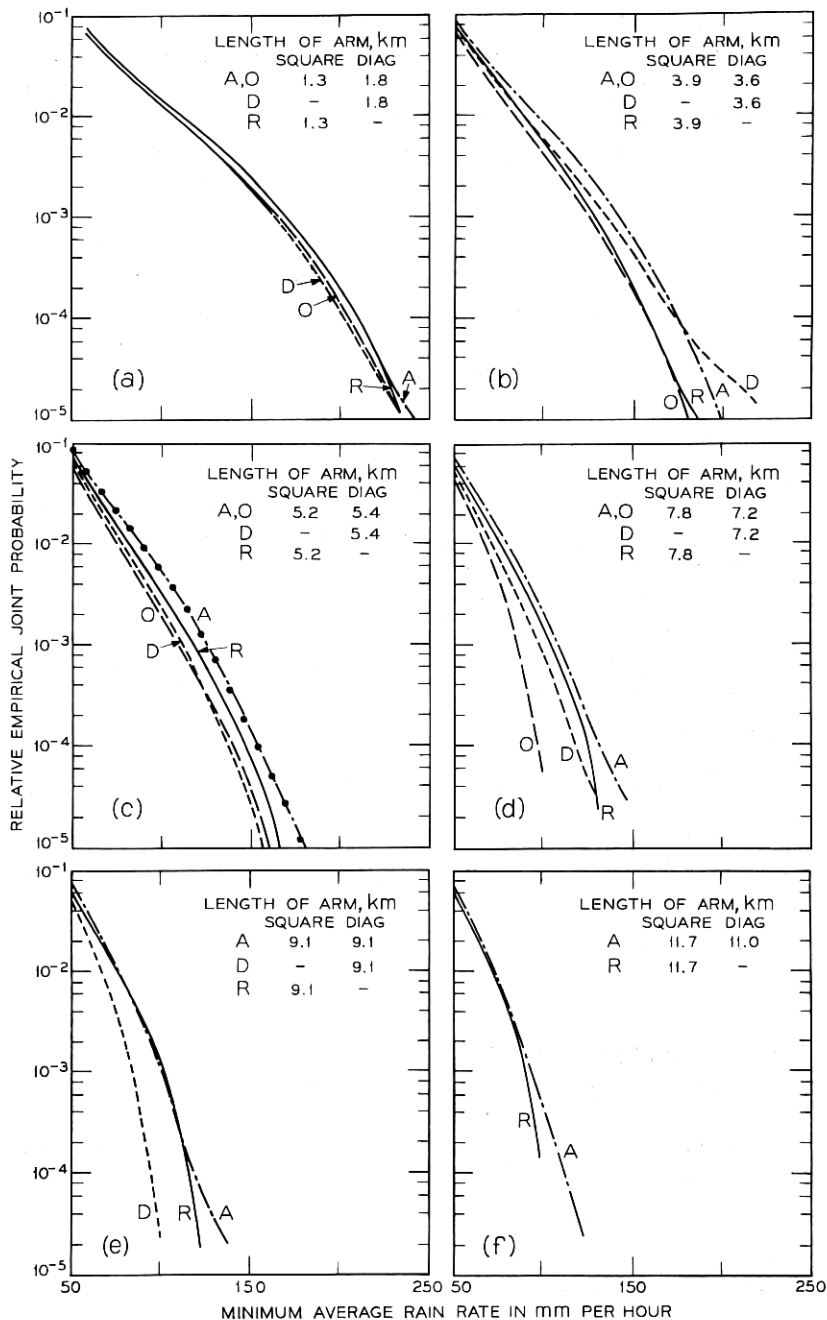
6.2 Adjoining Paths

The three parameters characterizing adjoining paths (two paths having an end point in common) are the length of the arms, the angle between the arms, and the orientation of the apex. We consider three angles, $\pi/4$ (acute), $\pi/2$ (right), and $3\pi/4$ (obtuse) and the eight orientations corresponding to rotations through an angle of $\pi/4$. Right angles containing diagonal paths (**D**) are differentiated from right angles containing square paths (**R**) to avoid confounding the different lengths of arm, thus there is a total of four categories: acute (**A**), obtuse (**O**), **D**, and **R**. Two paths with arms about 3 km long joined in an obtuse angle are indicated by the two dashed lines I_1 — I_2 toward the eastern side of the grid in Fig. 1.

The relative probability that the average rain rate on both paths exceeds the value on the abscissa is given by the ordinate for the four categories as parameters and six lengths of arm in Figs. 30a through f. The data, which are from the nine heavy rainfalls in 1967, have been pooled across orientations. The smallest and largest data bases for individual curves contain 53,800 and 4,842,000 points, respectively. For the same length of path, the results for adjoining paths lie between those for single paths (zero separation) and those for two paths separated by 1.3 km; and show the same decrease in probability with increasing length and rain rate.

A given rain rate is exceeded most frequently for the acute angles and least frequently for obtuse angles. The factor between the two increases with both path length and rain rate, and ranges from less than two (Fig. 30a) to about ten (Fig. 30d). Except for the acute angle, long paths cannot transit the center of the network so the results for various path lengths are based on different populations of stations, an effect which may introduce some bias. Very long path pairs with obtuse and diagonal angles cannot be formed in the network, and are missing from Figs. 30e and f. Once again the concatenation of uncertainties suggests that these results should not be considered better than an order of magnitude in absolute values.

Figure 31 contains results for the four pairs of adjoining paths of the square family (**R**) whose arms are 5.2 km long. The paths whose apex angles point to the northwest yield probabilities substantially higher than the other three orientations. This result is not peculiar to this category of angle or length of path. Symmetry considerations would lead one to expect the results for diagonally opposed pairs (that is, NW, SE and NE, SW) to be the same. The symmetry expectation,



← Fig. 30 — Relative empirical probability that the average rain rate on both of two adjoining paths will exceed the value on the abscissa. Curves are labeled with the apex angle; see Section 6.2.

however, is quite generally violated. The explanation is that for a small grid the population of path pairs is very different for the two sets of pairs. In this case the SE, NE, and SW path pairs can never have their apexes in the northwest corner of the grid where heavy rain rates are most frequent. This sensitivity to sampling bias further confirms our reluctance to regard the numerical probabilities derived from the results of this section as better than an order of magnitude.

Some further work on statistical dependence, edge and other geometric effects, and dispersion in the data is being considered.

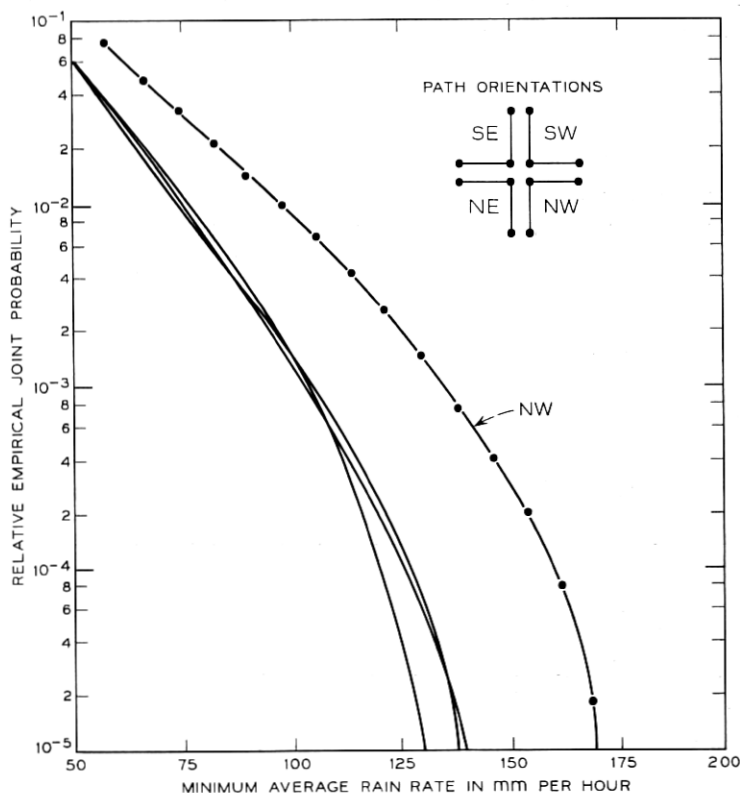


Fig. 31 — Asymmetry effect. Relative empirical probabilities that the average rain rate on both of two adjoining paths which constitute one of the four sets of pairs that make up the square family (R) with arms 5.2 km long will exceed the value on the abscissa. Inset: the four sets of pairs represented by the four curves.

VII. BENCHMARKS FOR ENGINEERING APPLICATIONS

All the probabilities presented so far have been calculated with respect to their own particular data bases. Thus the results do not contain assumptions about the representative properties of the data or subsets thereof, correction factors for missing data, or other adjustments which may require future revision. In this section we supply, for those who must make engineering calculations, factors which allow conversion of the relative probabilities to annual probabilities. The annual probabilities should be used only in the light of the various caveats and discussions regarding representativity, precision, and accuracy contained in this paper.

All benchmarks are computed at rain rates of 50 mm per hour. They may be applied to higher rain rates directly, but can be applied to lower rain rates only if they are scaled (see Section 4.1) to compensate for the bias introduced by excluding low-rain-rate data in the selection process.

7.1 *A Point-Rain-Rate Benchmark*

During the time covered by these data the average probability of observing a rain rate greater than 50 mm per hour was 4.3×10^{-4} . This number may be used to transform the scale of Fig. 14 as follows:

$$P(R) = 4.3 \times 10^{-4}[1 - p(R)],$$

where $P(R)$ is the annual probability that the rain rate exceeds R and $p(R)$ (the empirical probability that the rain rate is less than R) is given by Fig. 14. This benchmark is based on a nominal observation period of 158×10^4 scans (June 1, 1967 through November 30, 1967), an allowance of 15 percent for data known to have been lost, and a correction of all the stations to 40,069 scans (110 hours of rain) on the assumption that the loss of data is random over the data base.

For convenience, some of the point-rain-rate data have been replotted in Fig. 32a on the log-log scales often used in engineering. The figure shows the Crawford Hill 1967 data and the 1958 Island Beach, New Jersey, data of Mueller and Sims.⁸ (The curve in Fig. 32a was derived from Muellers and Sims Figs. 25 and 26 by D. C. Hogg.) Rain rates greater than 100 mm per hour are six times more probable in the Crawford Hill than the Island Beach measurements. However, rain rates above 250 mm per hour are equally frequent indicating

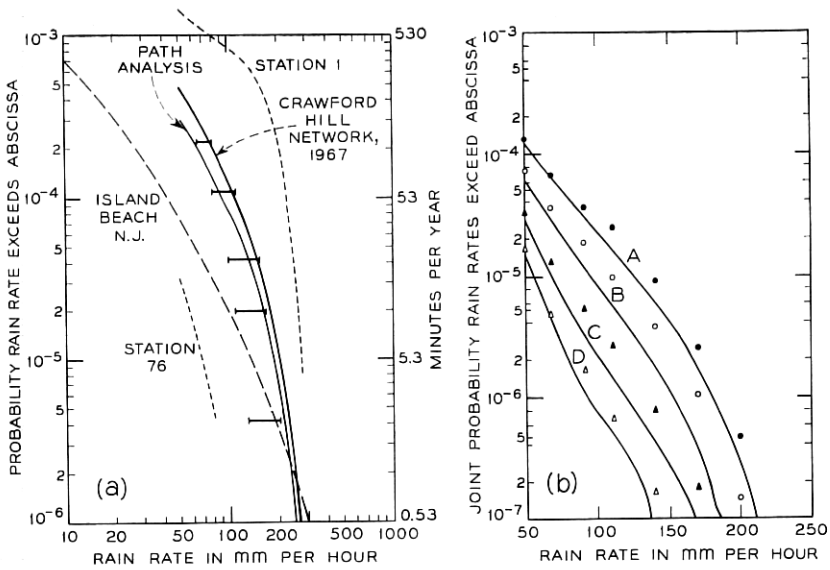


Fig. 32—(a) Annual probabilities for point rainfall rates. (b) Annual joint probabilities for point rainfall rates at two stations separated in space.

Label		Separation (km)
A	(●)	1.3
B	(○)	3.9
C	(▲)	6.5
D	(△)	9.0

somewhat different distributions of the rain rates measured. Fig. 32a also shows the results from stations 1 and 76. The spread between them is more than two orders of magnitude and encompasses most of the stations in the network. In view of this spread there is no reason to regard either the Crawford Hill aggregate or the Island Beach results as atypical. The bars in the figure are the interquartile ranges from the distributions of the results from the 96 stations (taken from Table III). The aggregate curve lies above the upper end of the bar when the rain rate is higher than 150 mm per hour, emphasizing the importance of rare events at these rain rates. If essential, extrapolation of the Crawford Hill curve to higher rain rates may best be accomplished by using the formula in Section 4.3 to construct the necessary ratios.

7.2 *A Spatial-Temporal Benchmark*

For rain rates above 50 mm per hour the relative probability scales of Figs. 23 and 27 may be converted to annual probabilities by multiplying by the factor 3.0×10^{-2} . This factor contains a correction of about 15 percent for rainfalls known to have been missed.

7.3 *Path Probabilities Benchmark*

The path probabilities benchmark is calculated on the premise that the 6725 scans analyzed contain all the pertinent data. The consequences of this premise are further discussed in this section. The probability scales of Figs. 28 through 31 should be multiplied by the factor 5.0×10^{-3} to convert to annual probabilities. This factor also contains the correction (about 15 percent) for rainfalls known to have been missed.

Curve A of Fig. 28a, which represents single stations, is replotted using this benchmark on Fig. 32 as the thin solid line labeled "path analysis." For single stations the corrected probabilities from the path analysis are about 25 percent below those obtained from the point-rain-rate analysis for rain rates over 100 mm per hour and 35 percent lower for rain rates near 50 mm per hour. This suggests that about 30 percent of the data above 50 mm per hour have been excluded from the data base for the path analysis. However, this has been confounded by the inclusion of estimates for missing observations in the path analysis. We feel that most of the data above 50 mm per hour were included, and the lower result for the path analysis is mainly artifactual.

Relative joint probabilities of pairs of stations for four different separations are taken from Fig. 28a, corrected to annual probabilities, and replotted as the solid curves in Fig. 32b, together with the corrected relative probabilities for the same nominal separations (the points) taken from Fig. 23. The geometry of the pooling in the two sets of data differs for all except for curve A of Fig. 32b. The path analysis included only separations along the square grid, while the spatial analysis included all stations (both square and diagonal) within appropriately separated concentric circles. As previously noted, the data base for the path analysis is a subset of that used for the joint probabilities, and the probabilities in the path analysis have been lowered by the inclusion of estimates for missing observations. Figure 32b demonstrates the results; the probabilities from both

analyses are in good agreement for low rain rates and diverge as the rain rate increases. The values from the spatial analysis (the points in Fig. 32b) are clearly more reliable.

Results for high rain rates on paths consisting of single gauges are most affected by the estimation of missing observations, so that Fig. 32b presents the worst cases for the path analysis. As agreement is good near 50 mm per hour and the missing-data effect is more serious at the higher rain rates and less severe for the longer paths, no single compensatory adjustment of the benchmark can be made.

7.4 Seasonal Correction

Data concerning fading on the microwave transmission path (T-R in Fig. 1) indicate that during 1967 most of the heaviest rain rates, which are usually generated by summer thunderstorms, occurred between June 1 and December 1.⁹ This seasonal effect decreases the probability of observing the highest rain rates by a factor of two, without materially affecting the probability of observing light rain. Insufficient data are available to correct our distributions for the seasonal variation. However, an approximate adjustment may be made by multiplying the annual probabilities associated with rain rates greater than about 100 mm per hour by the factor $\frac{2}{3}$, if the probabilities are to be applied on an annual basis. Correction for the seasonal effect would improve some aspects of the comparison between the Island Beach and Crawford Hill data (Fig. 32) and worsen others.

VIII. SUMMARY AND CONCLUSIONS

This article presents some statistical summaries of rainfall data acquired on the Crawford Hill, New Jersey, rain gauge network between June 1 and November 30, 1967. The emphasis of the analysis is on rain rates greater than 50 mm per hour, which cause substantial attenuation of electromagnetic waves in the 10 to 30 GHz frequency range. The objects of analysis are: the behavior of rain rates at a point in space, the relationship of two rain rates separated in space or in time, and the relationships of average rain rates on pairs of paths in various configurations.

The network consisted of 96 stations approximately uniformly distributed over a rectangular area 13 by 14 km.² The gauges are small in area (478 cm²), measure the rain rate averaged over less than one second, and are read out in sequence every ten seconds.¹ About 430

hours of data were recorded; the 110 hours of greatest interest have been examined in detail. The data are generally satisfactory for rain rates greater than about 10 mm per hour, although some substantial data-screening problems were encountered. There is no indication that these data are atypical of rain rates in northern littoral New Jersey, but they constitute a very small sample at the high rain rates; there is only informal indication of the large range of variation to be expected.

Heavy rainfall is found to come in irregular bursts. The time series containing these bursts are nonstationary (not even piecewise stationary at the sampling rate used), and no simple description of the wide variety of observed distributions of the rain rates has been found. Several poolings (for example, one station for the season, all stations for a rainfall) of the data were examined in an unsuccessful effort to find subsets of the data that behaved consistently enough to be described by a single distribution.

Figures 11 and 12 summarize the distribution of the point-rain-rate data. The high-rain-rate tail (*i*) is longer than would be expected on the basis of the normal distribution, (*ii*) can be approximated by an exponential function [equation (1)], (*iii*) depends heavily on a small number of cloudburst-like events (and thus shows a lot of scatter, see Fig. 13 and Table II), and (*iv*) is sensitive to the data-screening procedures used in the data processing. Restricting the data to rain rates greater than 50 mm per hour gives a better-defined subset of the data than does the original data-selection process and allows a further examination of the dispersion in the data (Table III), but fails to reduce the distributional variety observed.

Empirical probabilities for point rain rates above 50 mm per hour are given in Table IV. These measurements are somewhat higher (up to a factor of 6 at 50 mm per hour) for rain rates below 250 mm per

TABLE IV—EMPIRICAL PROBABILITIES FOR POINT RAIN RATES*

<i>R</i> mm per hour	Probability (rain rate > <i>R</i>)	Total time (rain rate > <i>R</i>) minutes per year
50	4.3×10^{-4}	230
100	1.3×10^{-4}	70
150	4.2×10^{-5}	22
200	1.0×10^{-5}	5

* Data are for Crawford Hill network during 1967.

hour than the Island Beach, New Jersey measurements of Mueller and Sims.⁸ However, better agreement would be obtained if a seasonal correction were applied.

When it is raining heavily at a station, nearby stations are also likely to show heavy rain. More detailed examination of the spatial characteristic of the data in terms of the joint probability of the rain rate being greater than R mm per hour at each of two stations simultaneously shows a rapid decrease of the joint probability with increasing separation for short separations and a minimum in the joint probability when the stations are separated by about 12 km (Fig. 23). Other spatial relationships also change qualitatively when the separations exceeds about 12 km. The temporal behavior may be characterized by the behavior of the joint probability of the rain rate at a station exceeding R mm per hour at both the beginning and end of a given time period (Fig. 27). This probability declines smoothly and relatively slowly as the interval increases from 10 to 360 seconds.

Of particular engineering interest are the joint probabilities that the average rain rate will simultaneously exceed a particular value on all of a number of paths. Figure 28 summarizes results of analysis of various sets of paths within the network for parallel paths; Figure 30 summarizes the same results for adjoining paths. Figures 29 and 31 give some indication of the dispersion among subsets of the paths.

For parallel paths up to 10 km long, the joint probability is minimum for paths separated by about 9 km. The joint probabilities for parallel paths 6.5 km long (and ordinary probabilities for single paths) are tabulated in Table V. As seen in Table V, for paths of this length the joint probability may be more than 200 times lower than the ordinary probability. This indicates the improved reliability of microwave transmission that may be obtained with appropriate choice of redundant paths. While the relative values of the probabilities obtained in this analysis are likely to be reasonably good, for reasons given in Section VI the numerical values should be regarded as correct only within an order of magnitude. The joint probabilities associated with average rain rates on adjoining paths are somewhat less than the probabilities associated with single paths through the network.

Section VII supplies benchmarks for converting the relative probabilities presented in the body of the paper to the annual probabilities needed for engineering calculations and discusses an additional correction for seasonal effects. Annual probabilities for point rain rates are plotted in Fig. 32a, and results of various analyses are compared and found to be consistent.

TABLE V—JOINT PROBABILITY THAT THE AVERAGE RAIN RATE ON BOTH OF TWO PARALLEL PATHS EXCEED R^*

R mm per hr	Separation km	0 (Single path)	1.3	3.9	6.5	9.1
50		2.5×10^{-4} (2 hr per yr)	1.5×10^{-4} (75 min per yr)	5×10^{-5} (25 min per yr)	1.5×10^{-5} (7.5 min per yr)	7.5×10^{-6} (4 min per yr)
100		2.0×10^{-5} (10 min per yr)	7.5×10^{-6} (4 min per yr)	1.3×10^{-6} (30 s per yr)	1.0×10^{-7} (3 s per yr)	$< \sim 10^{-8}$
150		1.5×10^{-6} (30 s per yr)	1.0×10^{-7} (3 s per yr)	$< \sim 10^{-8}$	$< \sim 10^{-8}$	$< \sim 10^{-8}$

* The paths are 6.5 km long. Probability is tabulated by the separation between paths. (Ordinary probability for the single paths.)

IX. ACKNOWLEDGMENT

The authors thank D. C. Hogg and R. A. Semplak for the data and for lively discussions thereof; Mrs. C. L. Clark and Mrs. R. J. Hedin for their assistance with data reduction and analysis.

REFERENCES

1. Semplak, R. A., "Gauge for Continuously Measuring Rate of Rainfall," *Rev. Sci. Instruments*, *37*, No. 11 (November 1966), pp. 1554-1558.
2. Semplak, R. A., and Keller, H. E., "A Dense Network for Rapid Measurement of Rainfall rate," *B.S.T.J.*, this issue, pp. 1745-1756.
3. Semplak, R. A., and Turrin, R. H., "Some Measurements of Attenuation by Rainfall at 18.5 GHz," *B.S.T.J.*, this issue, pp. 1767-1787.
4. Freeny, A. E., "Statistical Treatment of Rain Gauge Calibration Data," *B.S.T.J.*, this issue, pp. 1757-1766.
5. Wilk, M. B., and Gnanadesikan, R., "Probability Plotting Methods for the Analysis of Data," *Biometrika*, *55*, No. 1 (March 1968), pp. 1-17.
6. Semplak, R. A., unpublished work.
7. Hogg, D. C., unpublished work.
8. Mueller, E. A., and Sims, A. L., "Investigation of the Quantitative Determination of Point and Areal Precipitation by Radar Echo Measurements," Technical Report ECOM-00032-F, U. S. Army Electronic Command, Fort Monmouth, New Jersey, December 1966. (Contract DA-28-043 AMC-00032(E) with the Illinois State Water Survey, University of Illinois, Urbana, Illinois.)

