Statistics of Raingage Data

G. DRUFUCA[1] AND I. I. ZAWADZKI[2]

Department of Physics, McGill University, Montreal, Canada

(Manuscript received 1 November 1974, in revised form 2 July 1975)

ABSTRACT

Ten years of raingage data are processed in order to evaluate duration, average and maximum rate, mean square, variance, autocorrelation function and total amount for each rain storm. A spatial interpretation of these quantities is also given. Further, various rainfall rate probabilities are evaluated.

1. Introduction

Estimates of statistical properties of precipitation patterns can be obtained from raingage records. In this work ten years (1961–71) of data from the McGill Observatory raingage have been analyzed. The data correspond to the period from May to September of each year, thus covering the rainiest season. The raingage used is the standard tipping bucket gage extensively used by the Canadian Meteorological Service (Canadian Department of Transport, 1952).

Precipitation patterns have a temporal and spatial structure, and although one raingage can detect only time variations at a fixed place, it is possible, with simple assumptions, to infer from raingage data the spatial properties of the precipitation. This problem is considered in Section 2.

The properties analyzed in this work are statistics of the precipitation such as duration and length of the events, total amount of rainfall, average rainfall rate, maximum rainfall rate, mean square rainfall rate, and standard deviation of rainfall rate. Autocorrelation functions of the pattern both in time and space are also calculated. Probability distributions are given for all these variables and also correlations between them. Some of these correlations are particularly high and invariant to the synoptic type of precipitation, thus suggesting a general model of precipitation (Section 3).

In Section 4 some rainfall rate probability distributions of interest are given. They are: conditional probability of rainfall rate given an average rainfall rate, rainfall rate probability distributions for various time resolutions, and joint probabilities of rainfall rate at two points versus separation distance.

[1] Present affiliation: Centro Telecommunicazioni Spaziali del CNR, Politecnico, Milan, Italy.

[2] Present affiliation: Université du Québec à Montréal, Canada.

2. Synthetic storms

As a storm or rainfall pattern moves over a raingage, the measured rainfall rate varies with time. The variations arise from two effects: the advection of the spatial pattern of rain and the changes that occur within the time required to pass over the raingage. These effects are both present in raingage records and are not separable without additional information about the structure and motion of the rain pattern.

A "synthetic storm" is a description of the rain pattern in terms of rainfall rate as a function of distance along a line in the direction of storm motion. It is obtained from a raingage record by converting the raw data of rainfall rate as a function of distance by employing the storm translation speed to transform time to distance.

Since the synthetic storm is based on the assumption that all time changes observed at a point at the ground arise from the advection of spatial variations, it will not give an exact description of the distribution of rain with distance. It may be, however, that the statistical properties of a large number of synthetic storms are approximately the same as the corresponding statistical properties of real storms. Some evidence of the validity of this hypothesis can be found in the special case of convective storms, as discussed by Zawadzki (1973a). He measured the Eulerian time autocorrelation function of a storm, which is related to the time variations as seen from a fixed position on ground, and the space autocorrelation function, with the space lag along the direction of motion converted to time using the velocity of the storm, and found that up to time lags of the order of 40 min the two functions are very close. Given that the average storm speed is of the order of 50 km h^{-1}, 40 min correspond to a space lag of about 30 km. On the other hand, the two-dimensional

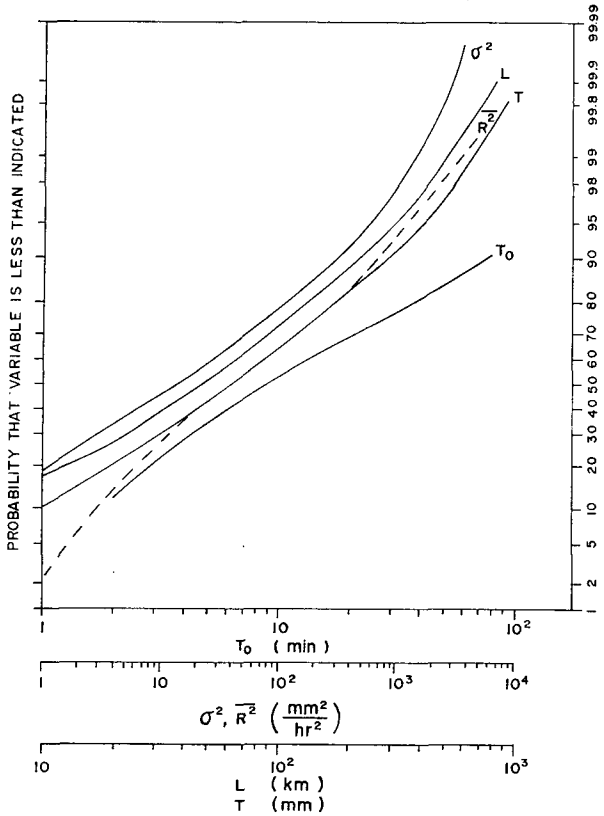

FIG. 1. Cumulative distributions of L, T, $\bar{R}^2$, σ^2, T_0.

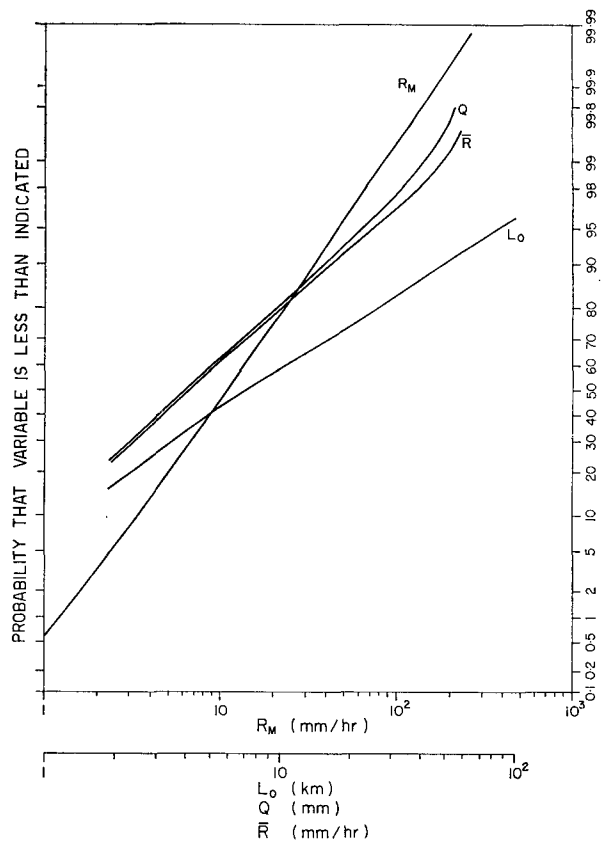

FIG. 2. Cumulative distributions of R_M, Q, $\bar{R}$, L_0.

space autocorrelation functions measured by Zawadzki also show a marked isotropy in the patterns indicating that statistics obtained along the direction of motion of the storm can be extended to the entire precipitation pattern.

Further evidence that the statistical properties of space variations in convective storms can be obtained from synthetic storms is given by a study of microwave attenuation by rain (Drufuca, 1974). Statistics of attenuation calculated from the synthetic storms and measured on an actual link agree well, down to probability levels where sample size becomes too small.

For widespread rain the hypothesis should hold even more because of the great uniformity, both in time and space, that these storms usually have.

As an estimate of the storm translation speed for each rain event the wind speed at 700 mb measured at Maniwaki has been used here as well as in Drufuca's work. This speed correlates well with the actual speed as measured by radar (Drufuca, 1974).

3. Statistics

For each raingage record a number of quantities have been calculated. They are listed as follows:

Duration T. The time between beginning and end of the precipitation. Two consecutive events are considered separately when they are at least one hour apart.

Length L. Obtained by multiplying the duration by the speed of translation of the precipitation pattern. Since the actual speed is not known, an estimate is used, the estimate being the 700 mb wind speed at the closest time, measured in Maniwaki (120 miles west of Montreal).

Average rainfall rate $\bar{R}$ within the time T.

Total accumulation of water Q, over the time T.

Mean square rainfall rate $\overline{R^2}$. Calculated for each event from

$$\overline{R^2}=\frac{1}{T}\int_0^T R^2(t)dt,$$

where $R(t)$ is the rainfall rate as a function of time t and T is the duration of the event.

Variance σ^2. Calculated for each event from

$$\sigma^2=\overline{R^2}-\bar{R}^2.$$

Maximum rainfall rate R_M, within each event.

Eulerian time autocorrelation[3] *function* $E(\tau)$. De-

[3] Note that the definitions of the autocorrelation functions given here differ from the ones for stationary random processes given in literature. For a discussion of this problem see Zawadzki (1973a).

TABLE 1. Percentiles of the cumulative distributions of the indicated quantities.

	Percentiles									Maximum
	10	20	30	40	50	60	70	80	90	observed
Duration T (min)	10	20	32	47	67	90	120	164	270	930
Length L (km)	7.4	15	23	33	47	67	90	127	210	690
Decorrelation time T_0(min)	1.8	3	4.5	6	8.3	13.5	22	40	78	572
Decorrelation distances L_0 (km)	1.3	2.5	3.7	5	6.7	10	16	27	55	312
Average rainfall rate $\bar{R}$ (mm h^{-1})	1.1	1.8	2.5	3.2	4	5.5	7	10	17	76
Total accumulation Q (mm)	1	1.6	2.2	3	3.9	5	6.8	9.5	15.6	74
$\overline{R^2}$ (mm^2 h^{-2})	2.7	6.7	8.7	19.5	36	62	128	270	665	7662
σ^2 (mm^2 h^{-2})	0.4	1.3	3	6	16	30	55	110	300	5876
R_M (mm h^{-1})	3.3	6	9	12	15.6	18	21	26.6	36.6	456

fined as

$$E(\tau)=\frac{1}{T}\int_0^T R(t)R(t+\tau)dt.$$

Space autocorrelation function[3] $A(\delta)$. Defined as

$$A(\delta)=\frac{1}{L}\int_0^L R(x)R(x+\delta)dx,$$

where $x=\bar{v}t$ and $\bar{v}$ is the average speed of translation of the pattern as measured by the 700 mb wind at Maniwaki.

Decorrelation time. Defined as the time lag T_0 for which

$$E(T_0)=\frac{1}{e}E(0).$$

Decorrelation distance. Defined as the space lag L_0 for which

$$A(L_0)=\frac{1}{e}A(0).$$

In Figs. 1 and 2 the cumulative distributions of these quantities are plotted on probability paper. It is apparent that all these quantities have a distribution very close to log-normal whenever the sample size is adequate. Table 1 shows some percentiles of the cumulative distributions.

Some correlation exists between the various quantities, and the correlation coefficients are shown in Table 2. A strong correlation ($r=0.957$) exists between $\bar{R}$ and $(\overline{R^2})^{\frac{1}{2}}$. Fig. 3 shows the bivariate plot of these two quantities.

A regression equation between $\overline{R^2}$ and $\bar{R}^2$ can be written as

$$\bar{R}^2=0.4\overline{R^2}. \tag{1}$$

Another strong correlation ($r=0.953$) is found between σ and R_M. In this case a regression equation can be written as

$$\sigma^2=3\times10^{-2}R_M^2. \tag{2}$$

The corresponding bivariate plot is shown in Fig. 4.

Decorrelation time T_0 and duration T have also high correlation ($r=0.701$) with a regression equation

$$T_0=0.17\,T. \tag{3}$$

A better correlation can be obtained between T_0 and T if T is weighted by a quantity that measures the "uniformity" of the precipitation pattern. As the

TABLE 2. Coefficients of correlation between the indicated variables (above diagonal) and between the common logarithm of the variables (below diagonal).

	T	L	T_0	L_0	$\bar{R}$	R_M	$(\overline{R^2})^{\frac{1}{2}}$	σ	Q
T	—	0.940	0.701	0.657	−0.649	−0.175	−0.649	−0.330	0.671
L	0.866	—	0.626	0.696	−0.596	−0.120	−0.596	−0.267	0.644
T_0	0.742	0.582	—	0.946	−0.567	−0.587	−0.567	−0.645	0.360
L_0	0.665	0.738	0.863	—	0.528	−0.544	−0.528	−0.600	0.340
$\bar{R}$	−0.328	−0.306	−0.253	−0.245	—	0.727	0.957	0.846	0.129
R_M	−0.082	−0.042	−0.251	−0.240	0.488	—	0.727	0.953	0.481
$(\overline{R^2})^{\frac{1}{2}}$	−0.402	−0.379	−0.330	−0.315	0.957	0.5	—	0.846	0.129
σ	−0.240	−0.202	−0.289	−0.278	0.801	0.881	0.802	—	0.402
Q	0.567	0.535	0.283	0.277	0.114	0.5	0.153	0.374	—

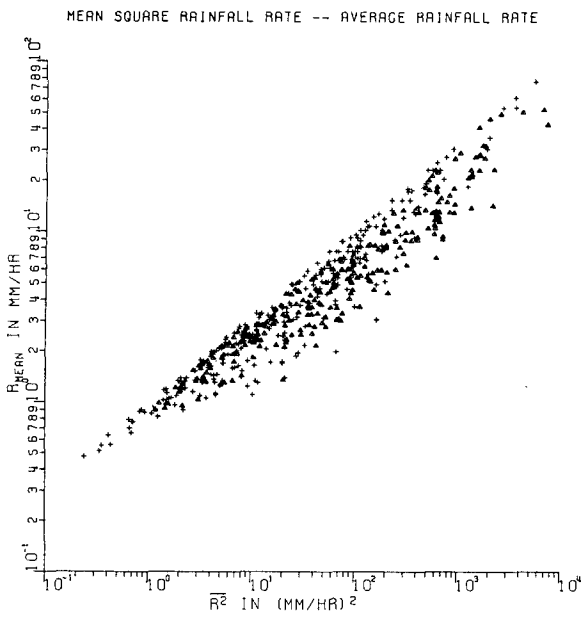

FIG. 3. Correlogram of $\bar{R}$ and $\overline{R^2}$.

weight, the ratio $\bar{R}^2/\overline{R^2}$ was used. This ratio has value 1 for uniform precipitation and value less than 1 for non-uniform precipitation. When this weight is applied to T smaller values will be obtained for patterns with higher variability. The present correlation is very high ($r=0.941$) and a regression equation can be written as

$$T_0=0.4T\frac{\bar{R}^2}{\overline{R^2}}.$$

For practical purposes it would be useful to know the statistical relations between $\bar{R}$, Q and R_M. For this purpose joint conditional probability densities have been evaluated. Table 3 shows this function for $\bar{R}$ given Q, Table 4 for R_M given Q, and Table 5 for R_M given $\bar{R}$. From these tables it is possible to calculate the joint probability densities. It is enough to multiply each number by the number of cases of its row and divide by the total number of cases (527). These functions can be of interest in practice to infer the value of one variable from measurements of the others. For example, if only Q is given, Table 3 allows the evaluation of $\bar{R}$ and Table 4 allows the evaluation of R_M.

TABLE 3. Conditional probability density* of $\bar{R}$ given Q.

Q (mm)	R (mm h^{-1}) 0–2	2–4	4–8	8–16	16–32	>32	Number of cases
0–2	29.3	19.7	19.7	22	5	4.3	
	14.6	9.85	4.9	2.75	0.31	0.098	136
2–4	33	29	20	8	8	2	
	16.5	14.5	5	1	0.5	0.045	136
4–8	22.7	34.5	21	14.3	5	2.5	
	11.4	1.72	5.25	1.8	0.31	0.057	119
8–16	15	27.5	25	22.5	7.5	2.5	
	7.5	13.8	6.25	2.8	0.47	0.057	80
16–32	8.9	20	28.9	22.2	15.5	4.5	
	4.45	10	7.2	2.8	0.97	0.10	45
>32	—	18.2	27.3	18.2	27.3	9	
	—	9.1	6.8	2.3	1.7	0.2	11

* The upper figure of each entry is the conditional probability density in percent in the indicated class, the lower figure the mean conditional probability density in an interval of 1 mm h^{-1} of $\bar{R}$ in the indicated class.

4. More statistics

A statistic of interest that can be obtained from raingage records is the distribution of rainfall rate. This distribution depends on the time interval over which rainfall rate is averaged. Fig. 5 shows the cumulative distributions of rainfall rate, for various averaging times. The upper line is for "instantaneous" rainfall rate in the sense that it is obtained directly from the gages without further integration. In the figure the probability of exceeding a given rainfall rate is expressed as minutes in 10 years. Comparing the slopes of the various curves one can say that the raingage employed in this study has an equivalent averaging time of less than 2 min. From these curves the probability distributions of time-accumulated rainfall can be readily obtained for the various time intervals by multiplying the rainfall rate of the lower scale by the time interval. For example, if the abscissa is multiplied by $\frac{1}{2}$ the curve labeled 30′ will indicate

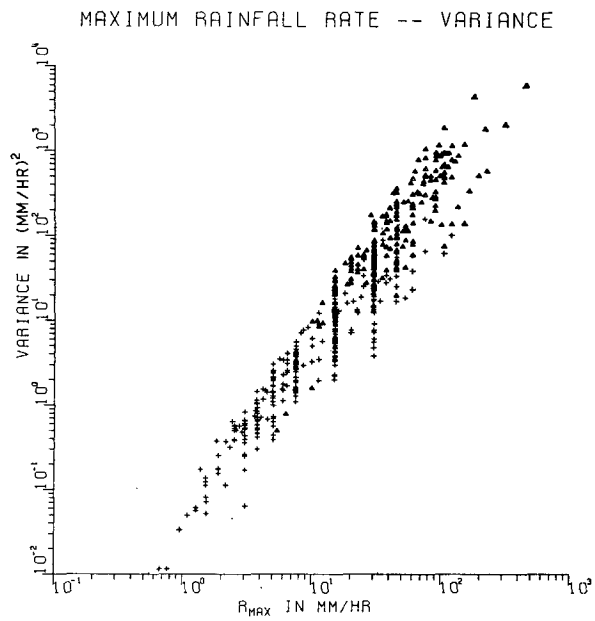

FIG. 4. Correlogram of R_M and σ^2. The concentration of data on fixed values of R_M is caused by the method employed in reading tipping bucket records.

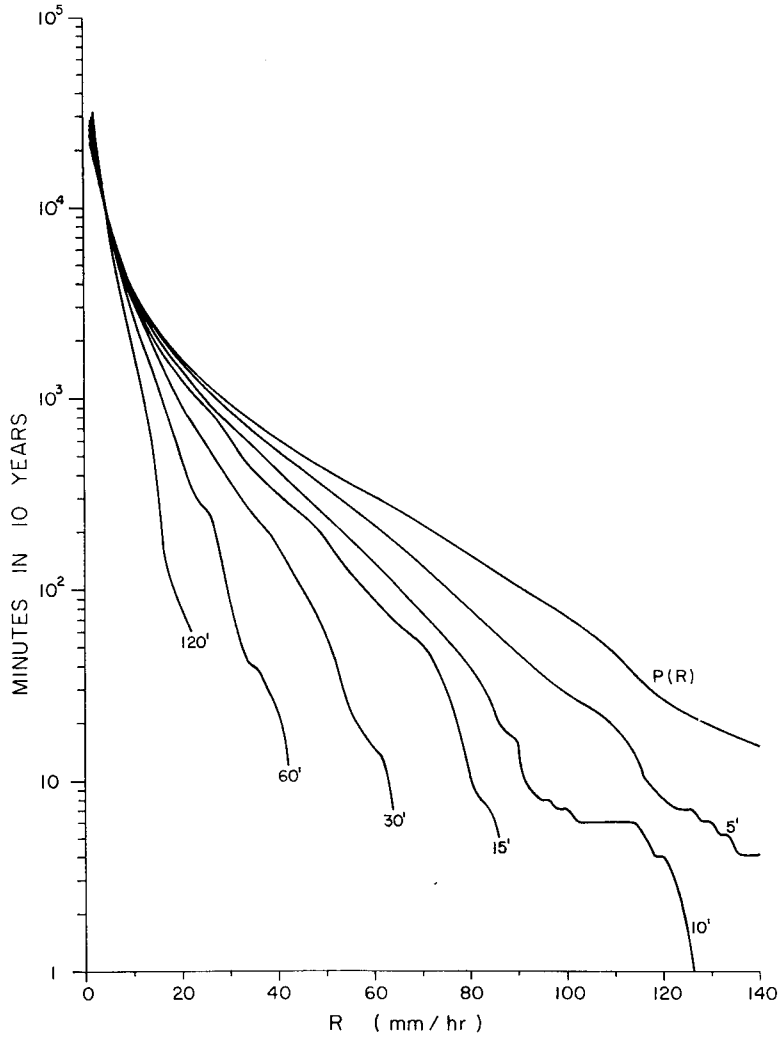

FIG. 5. Cumulative distributions of rainfall rate for various averaging times.

the minutes in 10 years that a 30 min accumulated rainfall exceedes the abscissa value.

In Fig. 6 the cumulative conditional probability distribution of rainfall rate for a given mean rainfall rate is shown. For comparison the overall probability is plotted with the same normalization (dashed). Although these curves were obtained from measurements of percentages of time that a given rainfall

TABLE 4. Conditional probability density of R_M given Q in percent (upper figure) and referred to an interval of 1 mm h^{-1} (lower figure).

Q (mm)	R_M (mm h^{-1}) 0–1	1–2	2–4	4–8	8–16	16–32	32–64	64–128	128–256	256–512	Number of cases
0–2	2.2	7.3	15.4	16.2	31.6	18.5	6.6	2.2	—	—	
	2.2	7.3	7.7	4.05	3.95	1.15	0.2	0.024			136
2–4	—	3.7	16.2	16.8	19.9	29.4	11.8	2.2	—	—	
	—	3.7	8.1	4.2	2.5	1.8	0.37	0.034			136
4–8	—	—	8.4	17.6	17.6	21	27.8	5.9	1.7	—	
	—	—	4.2	4.4	2.2	1.3	0.87	0.092	0.013		119
8–16	—	—	—	3.8	11.1	13.3	11.1	44.4	11.1	2.2	
				1.7	1.39	0.83	0.35	0.69	0.087	0.0086	45
>32	—	—	—	—	18.2	9.1	18.2	27.2	18.2	9.1	
					2.28	0.57	0.57	0.42	0.14	0.036	11

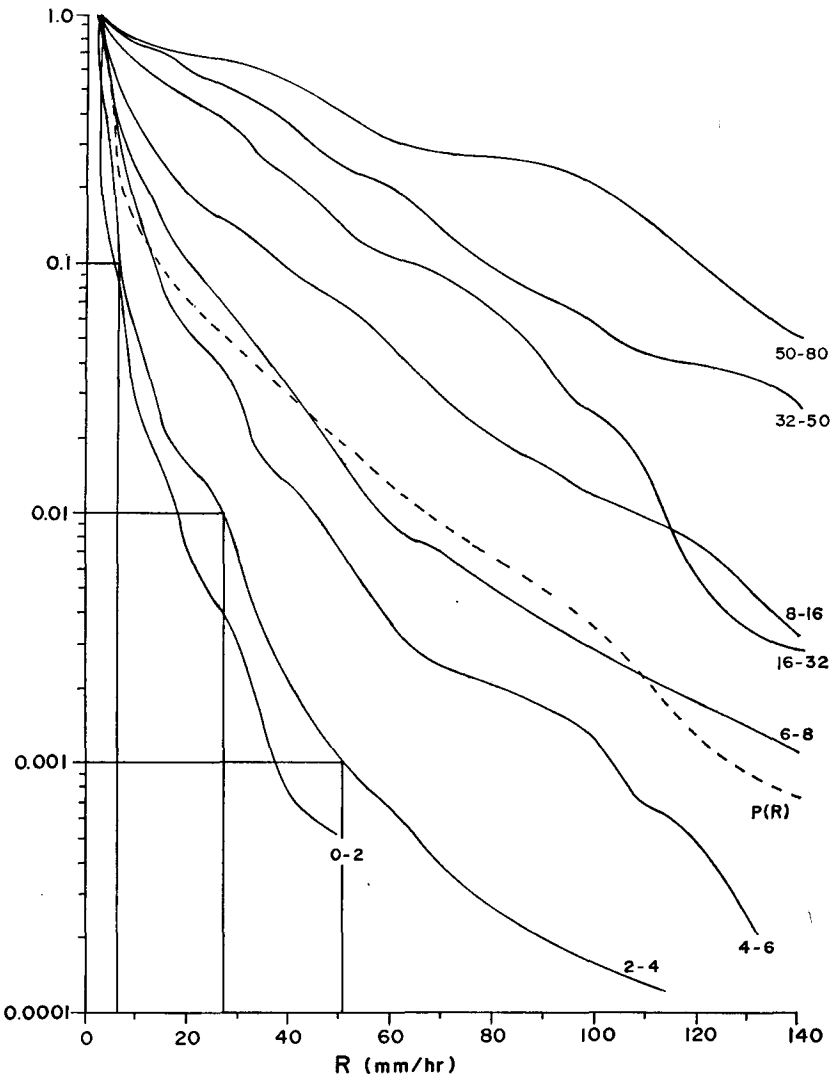

FIG. 6. Cumulative conditional probability distribution of rainfall for the indicated mean rainfall rate intervals. The overall probability is plotted for comparison (dashed).

TABLE 5. Conditional probability density of R_M given $\bar{R}$. The two figures of each entry have the same meaning as in Table 4.

$\bar{R}$ (mm h^{-1})	R_M (mm h^{-1}) 0–1	1–2	2–4	4–8	8–16	16–32	32–64	64–128	128–256	256–512	Number of cases
0–2	2.3 2.3	11.7 11.7	32 16	29 7.25	14.8 1.85	9.6 0.54	1.6 0.05	—	—	—	128
2–4	—	—	10.7 5.4	20.7 5.18	33.6 4.2	20.7 1.3	11.4 0.36	2.9 0.045	—	—	140
4–8	—	—	—	9.6 2.4	21 2.6	33.3 2.1	28.2 0.88	6.1 0.095	1.8 0.014	—	114
8–16	—	—	—	—	21.6 2.7	19.3 1.2	28.4 0.89	23.9 0.37	5.7 0.044	1.1 0.004	88
16–32	—	—	—	—	—	35.7 2.2	28.6 0.89	31 0.48	4.7 0.037	—	42
>32	—	—	—	—	—	—	33 1.03	53 0.83	6.5 0.051	6.5 0.025	15

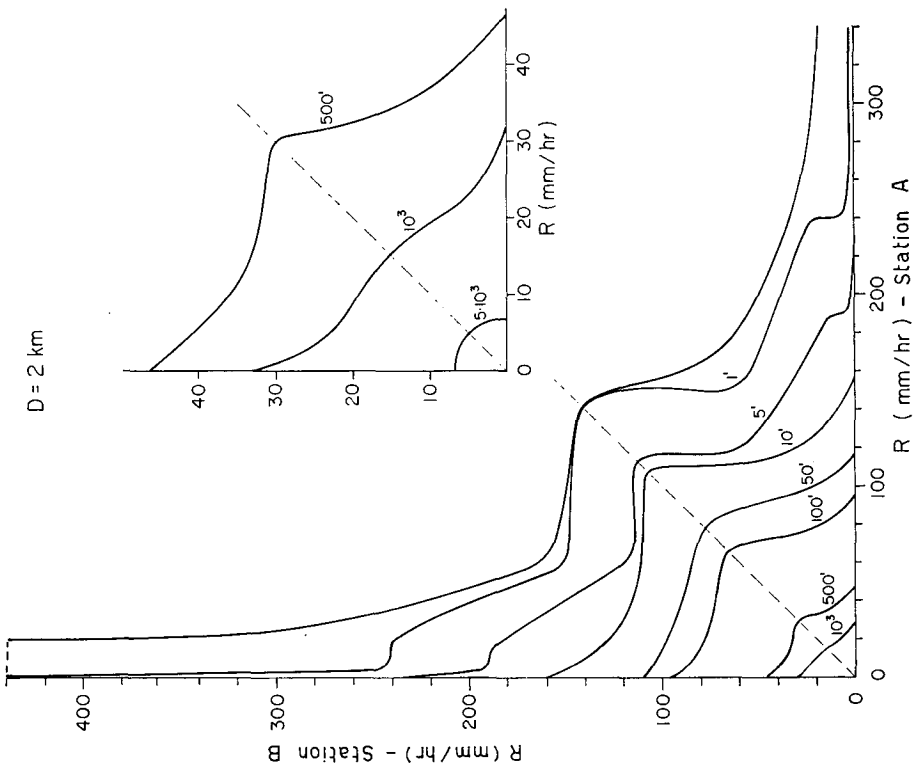

FIG. 8. As in Fig. 7, for $D=2$ km.

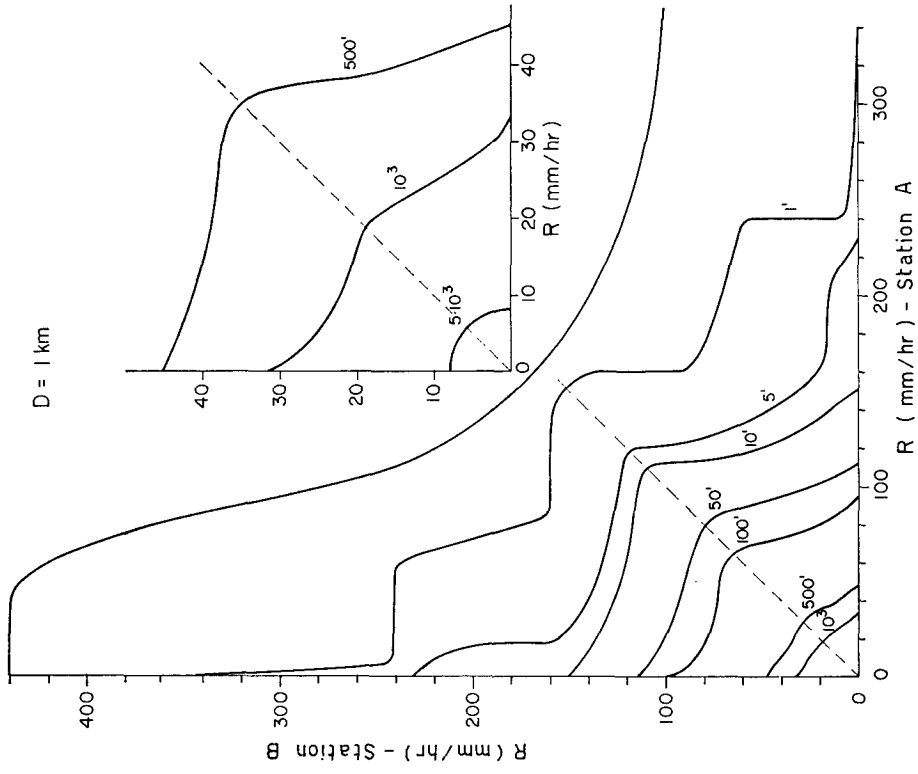

FIG. 7. Cumulative joint probability that given rainfall rates are exceeded simultaneously at two points separated by a distance $D=1$ km. The probabilities are expressed as minutes in 10 years.

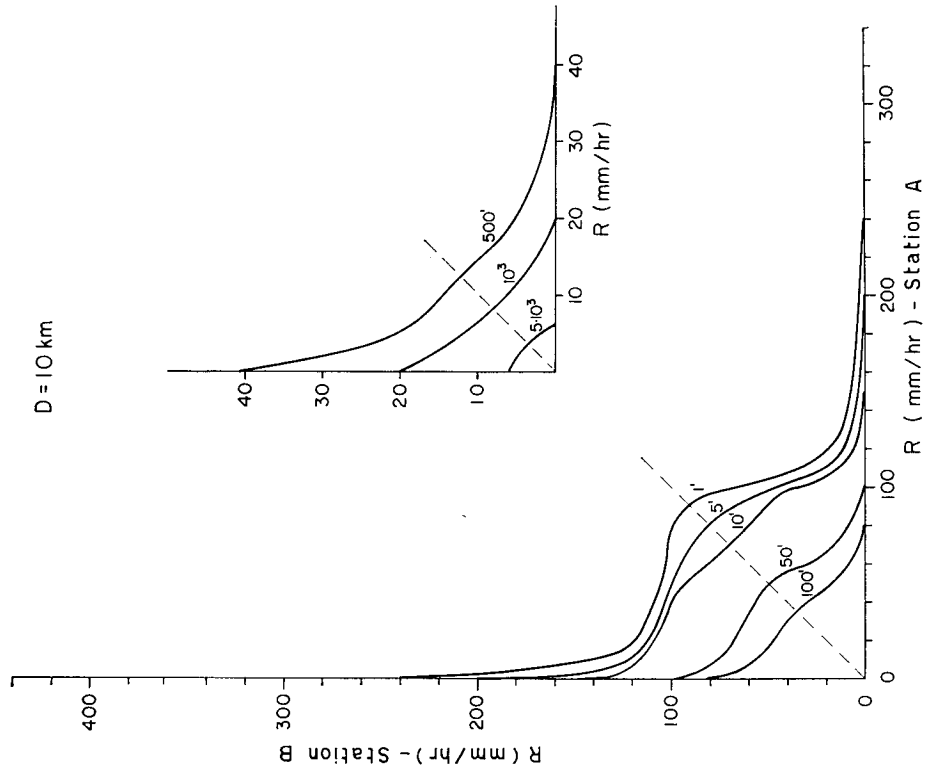

FIG. 10. As in Fig. 7, for $D=10$ km.

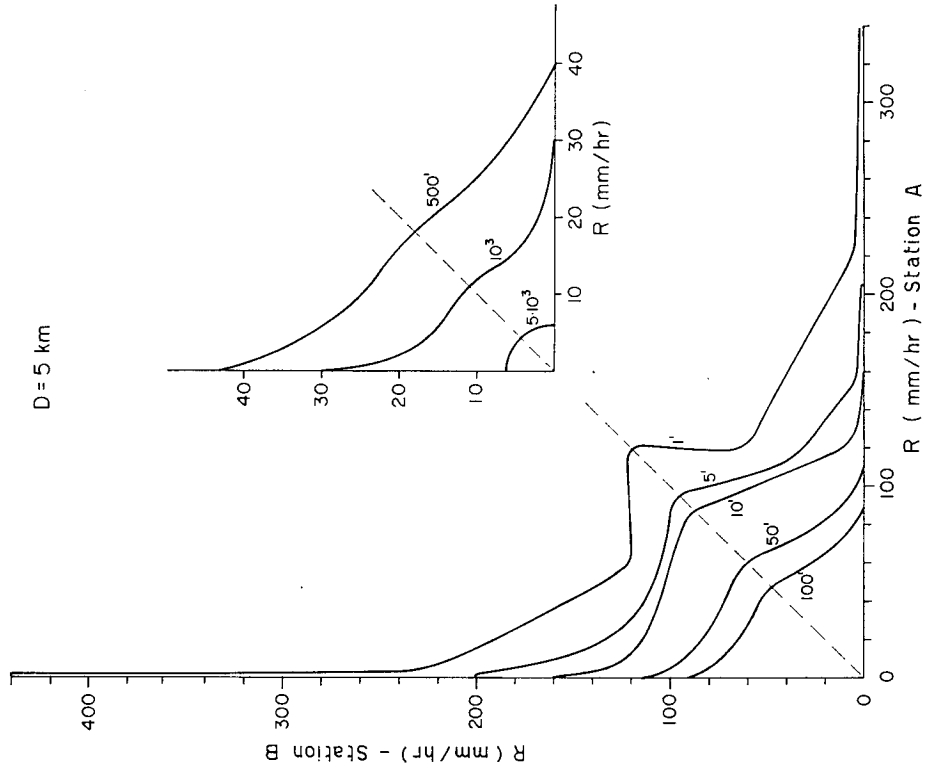

FIG. 9. As in Fig. 7, for $D=5$ km.

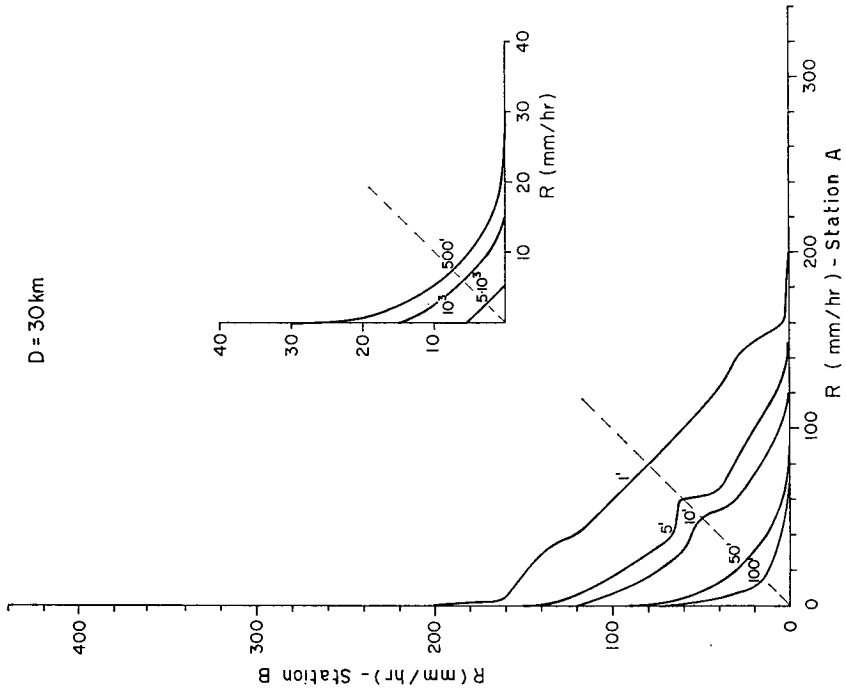

FIG. 12. As in Fig. 7, for $D=30$ km.

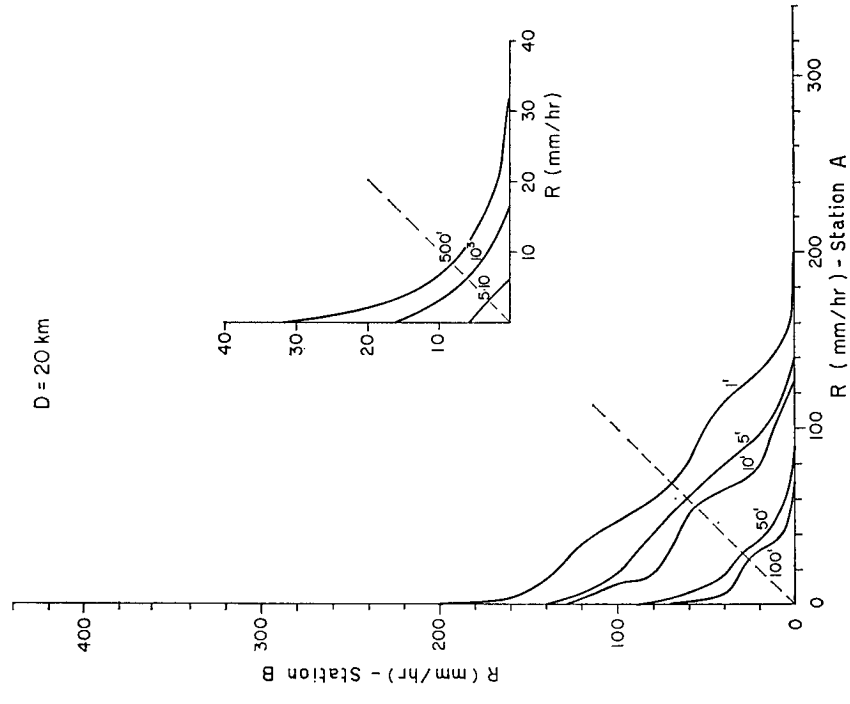

FIG. 11. As in Fig. 7, for $D=20$ km.

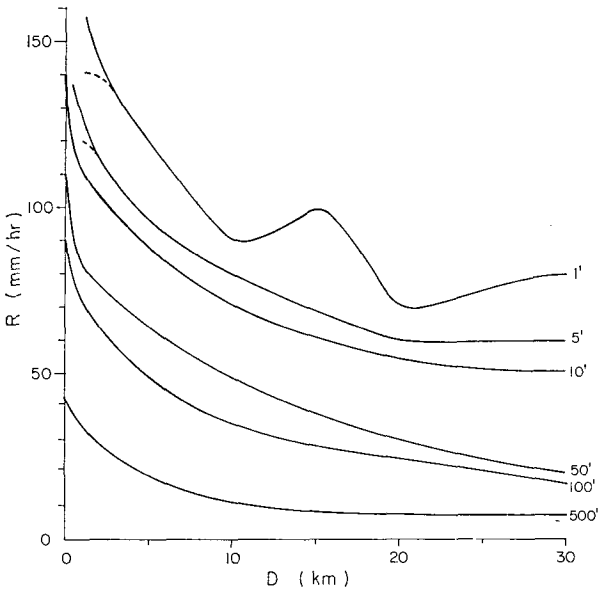

FIG. 13. Rainfall rate exceeded contemporaneously at two points for various probability levels (in minutes per 10 years) as a function of separation distance.

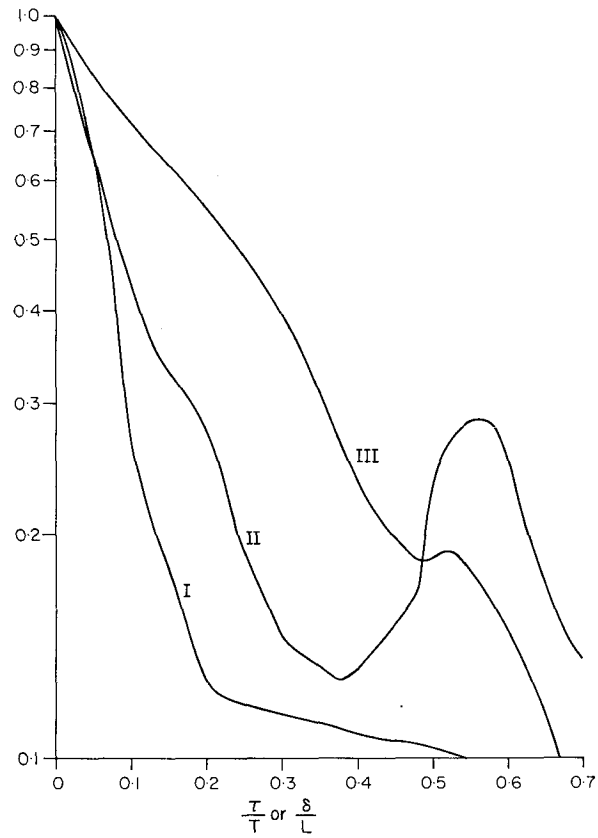

FIG. 14. Normalized autocorrelation function as a function of time relative to total length for three different rain events. The parameters of the three cases are listed in Table 6.

rate is exceeded at a fixed point they have a spatial interpretation. In fact, these curves can be used as estimates of the spatial distribution of rainfall rate in a precipitation pattern. For instance, for a storm of average rainfall rate between 2 and 4 mm h^{-1} we estimate that on the average 10% of the length of the storm has a rainfall rate in excess of 6 mm h^{-1}, 1% in excess of 28 mm h^{-1}, and 0.1% in excess of 50 mm h^{-1}.

How are these various regions of different rainfall rate actually distributed inside the pattern? It is a very difficult question. Some statistical information about this problem can be obtained by a knowledge of the autocorrelation function and the joint proba-

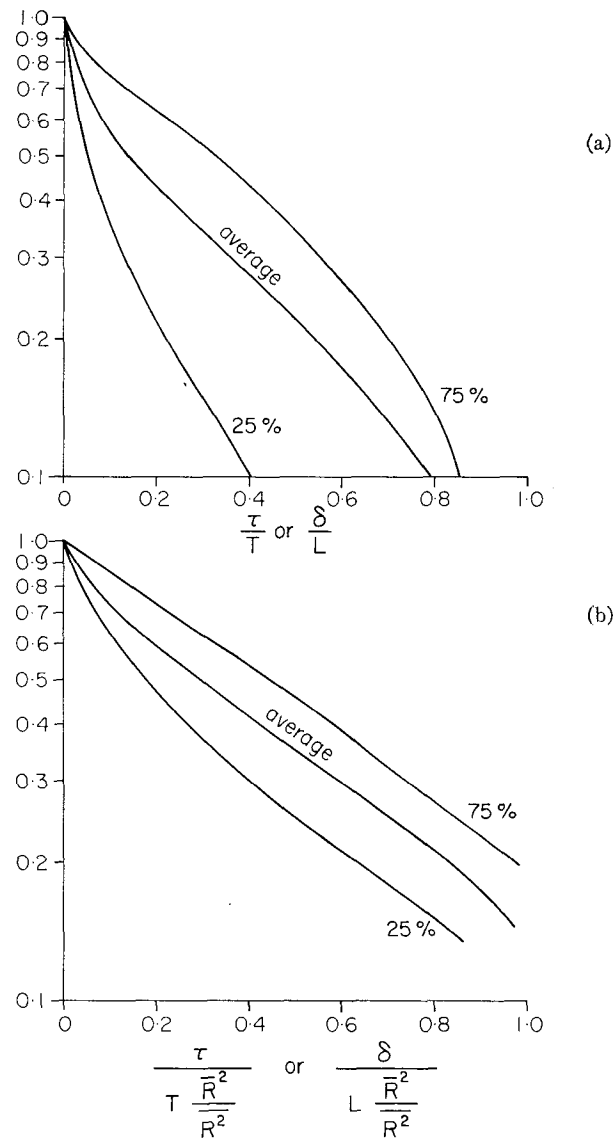

FIG. 15. Average normalized autocorrelation function as a function of time or space lag relative to total duration or length (a) and relative to average uniformity (b). The 25% and 75% percentiles contain 50% of the individual values.

bilities of exceeding certain values of rainfall rate for two points at various distances.

For this purpose the "synthetic storms" have been processed to obtain histograms of the amount of time that values of rainfall rate were exceeded at points separated by fixed distances. Naturally these points lie along the direction of motion of the precipitation pattern but it can be assumed that the properties measured in this way can be applied to a couple of points in any direction, as has been discussed in Section 2.

Figs. 7–12 show joint probabilities that a given rainfall rate is exceeded for various separation distances. The probabilities are expressed as minutes in 10 years. For example a rainfall rate of 25 mm h^{-1} at Station A and 10 mm h^{-1} at Station B, separated by 1 km, occurs 10^3 min over 10 years, yielding a probability of only 1.9×10^{-4}.

Fig. 13 shows the rainfall rate exceeded simultaneously on both points for various probability levels and versus separation distance.

The autocorrelation function as a function of space and time has been calculated for each event and the decorrelation time and distance have been presented in the previous section. In Fig. 14 three examples of the normalized autocorrelation function are presented which correspond to three different types of storms. The parameters characterizing the storms are listed in Table 6. Event I was of the convective type presenting large variability in the precipitation, event III was of the widespread type with uniform and light precipitation throughout a long duration, while event II corresponds to an intermediate case. In all cases the autocorrelation functions behave nearly exponentially for an extended time or space lag. Because of the correlation between the decorrelation time and duration (Table 2) it was inferred that plotting the autocorrelation functions versus time lag relative to duration would bring together curves which otherwise would be widely separated. In an attempt to facilitate the estimation of the autocorrelation function of a rain event of which only the duration (or length) is known, all the normalized autocorrelation functions were averaged with the time lag normalized to the duration. Fig. 15a shows the average autocorrelation function and the 25% and 75% percentiles, that is, 50% of the individual values fall between the two extreme curves. A better estimate is obtained if the time lag is normalized to $T(\bar{R})^2/\overline{R^2}$ (Fig. 15b), which follows from the fact that the correlation between decorrelation time and average uniformity is particularly high. An attempt to classify events by mean rainfall rate did not reduce further the scatter between the individual values.

TABLE 6. Parameters of three storms (Fig. 14).

	Event I	Event II	Event III
Date	21 Aug 1962	30 Aug 1963	11 Nov 1968
T (min)	50	128	500
L (km)	33	23	300
T_0 (min)	4.3	40.5	63
Q (mm)	10.7	12.7	7.9
$\overline{R^2}$ (mm^2 h^{-2})	572	69.7	2.16
$\bar{R}$ (mm h^{-1})	12.8	6	0.95
R_M (mm h^{-1})	73	30.5	7.6

5. Summary and conclusions

Ten years of raingage data have been processed to obtain a statistical description of the precipitation process at the ground as function of time at a point and as function of space using the concept of "synthetic storm."

Probability distributions of various parameters are given: mean rainfall rate, total amount, maximum rate, mean square rate, duration length, decorrelation time and distance. The joint distributions of mean rate, maximum rate and total amount are also determined. The probability distribution of rainfall rate is calculated and the effect of time smoothing of the data is studied. This distribution has also been calculated for various types of precipitations as represented by the mean rainfall rate.

Some information about the spatial structure of the precipitation process is given as joint probability of rainfall rate at two points for various separation distances and by the space autocorrelation function.

The high correlation between some of the parameters indicates regularity in the behavior of precipitation, thus suggesting a statistical model.

Studies such as the one presented here have direct applications in hydrology (for example Zawadzki, 1973b), radar meteorology (Zawadzki, 1973c), precipitation attenuation of microwaves, etc. Also they represent the kind of information which can be used to test convection models on a synoptic scale.

Acknowledgments. Financial support for this work was provided by the Communication Research Centre of the Department of Communication of Canada. Mr. D. Salomon and Ms. P. Sharp did all the computations.

REFERENCES

Canadian Department of Transport, Meteorological Branch, 1952: The tipping bucket raingauge. Instrument Manual 41.5.

Drufuca, G., 1974: Rain attenuation statistics for frequencies above 10 GHz from raingauge observations. *J. Rech. Atmos.*, **1-2**, 399–411.

Zawadzki, I. I., 1973a: Statistical properties of precipitation patterns. *J. Appl. Meteor.*, **12**, 459–472.

——, 1973b: Errors and fluctuations of raingauge estimates of areal rainfall. *J. Hydrol.*, **18**, 243–255.

——, 1973c: The loss of information due to finite sample volume in radar-measured reflectivity. *J. Appl. Meteor.*, **12**, 683–687.