Persistence, Runs and Recurrence of Precipitation

IVER A. LUND AND DONALD D. GRANTHAM

Air Force Geophysics Laboratory, Hanscom A.F.B., Mass. 01731

(Manuscript received 1 March 1976, in revised form 15 March 1977)

ABSTRACT

A total of 511 056 hourly observations of precipitation, taken over a 13-year period at nine stations, were studied to obtain a better understanding of the characteristics of persistence, runs and recurrence. Each hourly precipitation observation was categorized as either none, light, moderate or heavy. Probabilities of each category, except heavy, were estimated from relative frequencies determined from this large data sample and were compared with some theoretical models. The heavy category occurred too infrequently at all of the nine stations to obtain statistically stable relative frequencies from a 13-year period of record. The sample provided sufficient information about the other categories to confidently fit models to the data. The models can be applied to estimate probabilities that precipitation will be observed for sequences of x hours, or more; for exactly x hours; or at time t and also at time $t+x$ hours.

1. Introduction

Duration, persistence, runs and recurrence are all interrelated. For this study, they have been defined as follows: duration—continuous successes; persistence—consecutive successes separated by 1 h; runs—consecutive successes separated by intervals of 1 h beginning and ending with a failure; and, recurrence—successes occurring at time t and also at time $t+x$ hours.

This study is part of an investigation conducted for the purpose of obtaining a better understanding of persistence, runs and recurrence of weather events. Duration could not be studied because the data were observed at hourly intervals. Of major interest are those weather events which are usually recorded in categories, e.g., precipitation recorded as none, light, moderate or heavy, or sky cover recorded as clear, scattered, broken or overcast.

This paper includes tables of observed relative frequencies of precipitation occurrences and models for estimating probabilties of precipitation occurrences. The models provide answers to questions such as: What is the probability of observing sequencies of more than 5 h of precipitation; of observing a run of exactly 5 h of precipitation; and, of observing precipitation at time t and also at time $t+5$ h? The models require a knowledge of the unconditional probability of the event, in this case precipitation, and a measure of the temporal correlation between occurrences of precipitation.

2. Data

Records of hourly precipitation occurrences, observed in winter (December, January, February) and summer (June, July, August) during the 13-year period 1951–63, at the following nine stations, shown in Fig. 1, were studied:

- LGA LaGuardia Airport, New York, N. Y.
- JFK Kennedy International Airport, New York, N. Y.
- EWR Newark Airport, N. J.
- PHL Philadelphia International Airport, Pa.
- BAL Baltimore-Washington International Airport, Md.

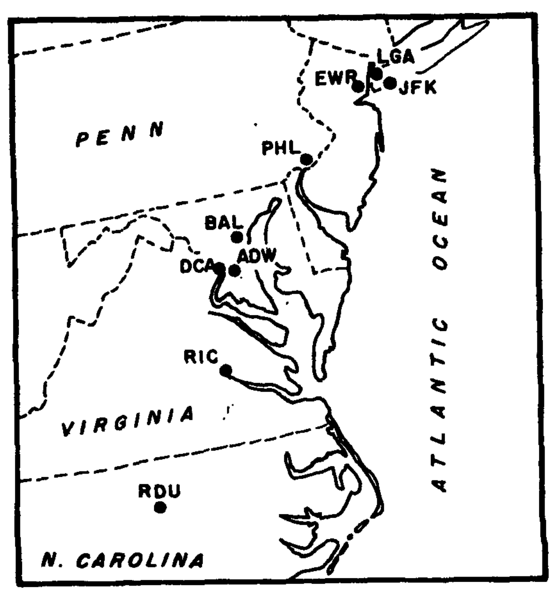

FIG. 1. Location of the nine stations whose winter and summer hourly observations of precipitation were studied.

DCA National Airport, Washington, D. C.
ADW Andrews AFB, Md.
RIC Byrd Field, Richmond, Va.
RDU Raleigh-Durham Airport, N. C.

Each hour, approximately on the hour, a weather observer at each of the above stations went outdoors to make his regular observation. One of the weather elements that he recorded was precipitation, including type and intensity. The *Federal Meteorological Handbook* (1975) describes how the observations are taken. For this paper, the precipitation events were somewhat arbitrarily categorized as light (L), moderate (M) and heavy (H) as shown in Table 1.

The meaning of the letters and symbols in Table 1 are as follows: light (−), moderate (no intensity indicated), heavy (+), hail (A), sleet (E), grains (G), drizzle (L), pellets (P), squalls (Q), rain (R), snow (S), showers (W) and freezing (Z).

3. Data processing

Each hourly observation was coded according to precipitation category as follows: none (0), light (1), moderate (2) and heavy (3). Some of the stations had no missing observations, others only a very few. These few observations were assumed to be precipitation-free and coded as none (0). There were 28 080 [(24 observations day^{-1})×(90 days $season^{-1}$)×(13 seasons)] observations, in winter, and 28 704 [(24 observations day^{-1}) ×(92 days $season^{-1}$)×(13 seasons)] observations, in summer, processed for each station.

4. Persistence

a. Observed

We denote the occurrence of a given precipitation category as a success S and the nonoccurrence as a failure F. The relative frequency of S, $RF(S_1)$, is found from the data by dividing the number of times the precipitation category occurred, $n(S_1)$, by the sample size N. The relative frequency of two successes in a row, $RF(S_2)$, is found from the data by dividing the number of times a success was followed by a success, $n(S_2)$, by the sample size N minus the end effect, in this case 13, because there were 13 years when the next season's data were not used to determine whether precipitation occurred on the first hour of the next season. The relative frequency of x successes in a row, $RF(S_x)$, is found by dividing the number of times x consecutive successes was observed, $n(S_x)$, by the sample size N minus the end effects, in this case $13(x-1)$, i.e.,

$$RF(S_x)=\frac{n(S_x)}{N-13(x-1)}\approx\frac{n(S_x)}{N}. \quad (1)$$

This processing of the data was done for all categories for all nine stations in both winter and summer. The precipitation category heavy occurred too infrequently to obtain statistically stable relative frequencies from a 13-year period of record; therefore, this category was included with the moderate category.

TABLE 1. Precipitation categories (see text).

Light (L)		Moderate (M)	Heavy (H)
R−	RW−	R	R+
ZR−	RQ−	RW	RW+
L−	L	ZR	ZR+
L+	ZL−	RQ	RQ+
ZL	ZL+	S	S+
S−	SP−	SW	SW+
SP	SP+	SQ	SQ+
SW−	SQ−	E	E+
SG−	SG	EW	EW+
SG+	E−	A	A+
EW−	A−	AP	AP+
AP−			

It can be shown that the relative frequency of a success given that x consecutive successes have occurred, $RF(S|S_x)$, is equal to the relative frequency of $x+1$ consecutive successes, $RF(S_{x+1})$, divided by the relative frequency of x consecutive successes, $RF(S_x)$, i.e.,

$$RF(S|S_x)=\frac{RF(S_{x+1})}{RF(S_x)}. \quad (2)$$

The conditional relative frequencies $RF(S|S_x)$ were computed for periods up to 72 h. Selected values for the first 13 h are shown for all nine stations and all three precipitation categories in Tables 2, 3 and 4. The median relative frequencies are indicated with asterisks.

The first column in each of the tables gives $RF(S|S_0)$ which is defined as $RF(S)$, the unconditional relative frequency of the given precipitation category. Although both the unconditional and conditional relative frequencies vary from station to station, there is often no consistent pattern to the variations. It was subjectively decided to assume that the data from all stations were drawn from the same sample and to use the median values to obtain estimates of the conditional probabilities $\hat{P}(S|S_x)$ required for obtaining estimates of the joint probabilities $\hat{P}(S_{x+1})$.

The median values of $RF(S|S_x)$ for winter and summer are shown in Fig. 2 for periods up to 12 h when sufficient data existed to obtain a median. The median relative frequency of no precipitation (none) was 0.865 in winter and 0.930 in summer. Because this is a frequently occurring category there were many long sequences of successes when no precipitation was considered a success. The median conditional relative frequencies, given in Tables 2 and 5 and shown as dots in Fig. 2, increase in the magnitude for 10 h in winter and 12 h in summer. They never vary by more

TABLE 2. Relative frequencies of success given that x consecutive successes have occurred, $RF(S|S_x)$, obtained from the data sample when no precipitation (none) is considered a success. Median values are identified with asterisks.

Season	Station	0	1	3	5	7	9	11	≥12
					x (hours)				
Winter	LGA	0.853	0.968	0.977	0.979	0.979	0.980	0.980	0.980
	JFK	0.860	0.968	0.977	0.979	0.981	0.981	0.981	0.981
	EWR	0.851	0.967	0.976	0.978	0.979	0.981	0.981	0.980
	PHL	0.862	0.967	0.977	0.980	0.981	0.981	0.981	0.982
	BAL	0.865*	0.973	0.980	0.981*	0.982*	0.982*	0.983*	0.983*
	ADW	0.874	0.971	0.979*	0.981*	0.982*	0.983	0.983*	0.983*
	DCA	0.871	0.971	0.979*	0.981*	0.982*	0.982*	0.983*	0.983*
	RIC	0.870	0.971	0.980	0.982	0.983	0.983	0.983*	0.984
	RDU	0.884	0.969*	0.979*	0.982	0.983	0.984	0.985	0.985
Summer	LGA	0.924	0.970*	0.979	0.982	0.982	0.983	0.984	0.984
	JFK	0.924	0.970*	0.977*	0.980*	0.981*	0.981	0.982*	0.983*
	EWR	0.924	0.970*	0.977*	0.980*	0.982	0.982*	0.983	0.983*
	PHL	0.928	0.971	0.979	0.980*	0.982	0.982*	0.983	0.983*
	BAL	0.931	0.973	0.979	0.981	0.982	0.983	0.983	0.983*
	ADW	0.935	0.972	0.978	0.980*	0.981*	0.982*	0.982*	0.982
	DCA	0.933	0.970*	0.977*	0.979	0.981*	0.981	0.982*	0.982
	RIC	0.930*	0.970*	0.977*	0.980*	0.981*	0.982*	0.982*	0.983*
	RDU	0.930*	0.969	0.975	0.977	0.979	0.980	0.980	0.980

than 0.001 after hour 12. Therefore, the conditional probabilities are regarded as constant after hour 12.

b. *Modeled*

The probability of a sequence of x hours of successes can be estimated with the first-order Markov expression

$$\hat{P}(S_x)=P(S)[P(S|S_1)]^{x-1}, \qquad (3)$$

where $P(S|S_1)$ is the probability of a success given a success has occurred and x equals the number of hours.

The relative frequencies $RF(S)$ and $RF(S|S_1)$ obtained from the data are given in columns one and two, respectively, in Tables 2–4. The estimated conditional probabilities $\hat{P}(S|S_1)$, given in Table 5, were used to test Eq. (3). The model fit the observed values within a few percent for the first few hours but there were large differences between the model estimates and corresponding sample relative frequencies when probabilities of sequences of successes longer than a few hours were estimated. Figs. 3–5 illustrate differences between model estimates and observed relative frequencies when the data from all nine stations are pooled into one sample. These figures illustrate the failure of Eq. (3) to adequately estimate long sequences

TABLE 3. As in Table 2 except when precipitation of any intensity (L+M+H) is considered a success.

Season	Station	0	1	3	5	7	9	11	≥12
					x (hours)				
Winter	LGA	0.147	0.815	0.861	0.869	0.870	0.855	0.855*	0.855*
	JFK	0.140	0.803*	0.848	0.859	0.850	0.841	0.838	0.847
	EWR	0.149	0.813	0.862	0.877	0.870	0.863	0.853	0.862
	PHL	0.138	0.791	0.850*	0.860*	0.861*	0.857*	0.856	0.855*
	BAL	0.135*	0.827	0.864	0.877	0.875	0.875	0.864	0.855*
	ADW	0.126	0.802	0.850*	0.860*	0.869	0.866	0.869	0.860
	DCA	0.129	0.803*	0.846	0.856	0.850	0.850	0.842	0.839
	RIC	0.130	0.803*	0.853	0.857	0.852	0.859	0.867	0.859
	RDU	0.116	0.764	0.818	0.827	0.849	0.845	0.831	0.822
Summer	LGA	0.0758	0.633*	0.760	0.805	0.815*	0.802	0.807	0.828
	JFK	0.0764	0.636	0.755	0.799	0.809	0.796	0.807	0.800
	EWR	0.0762	0.639	0.762	0.802	0.805	0.819	0.789	0.794
	PHL	0.0723	0.634	0.744*	0.782	0.825	0.801	0.803	0.822
	BAL	0.0688	0.637	0.752	0.812	0.823	0.803	0.840	0.855
	ADW	0.0649	0.604	0.744*	0.802*	0.812	0.814	0.833	0.836
	DCA	0.0673	0.590	0.742	0.809	0.826	0.813*	0.815*	0.827*
	RIC	0.0695	0.601	0.732	0.796	0.822	0.840	0.880	0.872
	RDU	0.0696*	0.584	0.672	0.733	0.788	0.824	0.821	0.826

TABLE 4. As in Table 2 except when moderate and heavy (M+H) precipitation is considered a success.

Season	Station	x (hours) 0	1	3	5	7	9	11	≥12
Winter	LGA	0.00773	0.396*	0.476	0	0	0	0	0
	JFK	0.00573	0.354	0.615	0.818	0.714	0.333	0	0
	EWR	0.00705	0.439	0.477	0.400	0.667	0	0	0
	PHL	0.00534	0.373	0.461	0.400	0	0	0	0
	BAL	0.00484	0.346	0.454	0.833	0.750	0.500	0	0
	ADW	0.0118	0.468	0.641	0.595*	0.667	0.714	0.750	0.667
	DCA	0.00630*	0.350	0.500*	0.600	0.333*	0	0	0
	RIC	0.00865	0.399	0.511	0.400	0	0	0	0
	RDU	0.00573	0.453	0.540	0.692	0.833	0.750	0.500	0
Summer	LGA	0.00728	0.282*	0.517*	0.429	0	0	0	0
	JFK	0.00547	0.267	0.500	0.500*	0	0	0	0
	EWR	0.00582	0.216	0.400	0.333	0	0	0	0
	PHL	0.00502	0.215	0.300	0	0	0	0	0
	BAL	0.00704	0.351	0.564	0.615	0.833	0.750	0.500	0
	ADW	0.00829	0.336	0.667	0.880	0.900	0.875	0.833	0.800
	DCA	0.00676*	0.304	0.552	0.556	0.750	0.500	0	0
	RIC	0.00892	0.340	0.706	0.692	0.786	0.778	0.600	0.333
	RDU	0.00488	0.207	0.250	0	0	0	0	0

of successes. For example, note the large departures of the first-order Markov model estimates [Eq. (3)] from the sample relative frequencies shown in Fig. 3.

Agreement between model estimates and sample relative frequencies can be improved by applying the following higher order model to estimate 1, 2 and x hours of consecutive successes, respectively:

$$\hat{P}'(S_1)=P(S), \tag{4}$$

$$\hat{P}'(S_2)=P(S)P(S|S_1), \tag{5}$$

$$\hat{P}'(S_{x+1})=P(S)P(S|S_1)\ldots P(S|S_x), \tag{6}$$

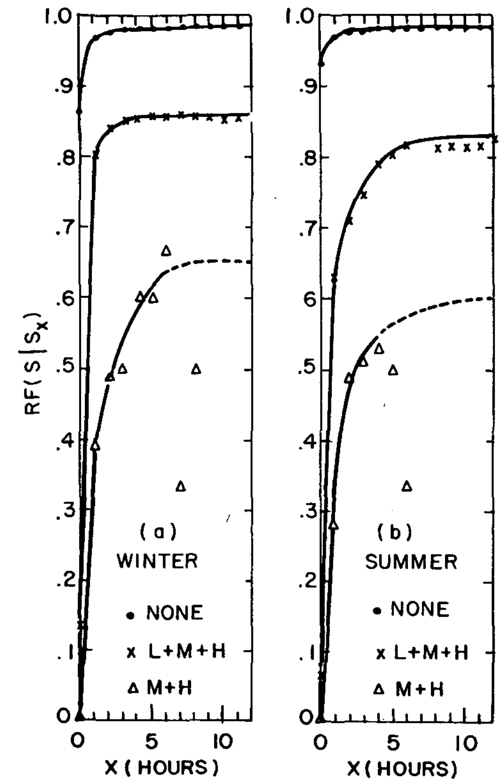

FIG. 2. Relative frequencies of success, given x hours of consecutive successes have occurred, in winter (a) and summer (b). The curves were subjectively drawn. All points to the right of the solid portions of the curves are based on fewer than 30 cases.

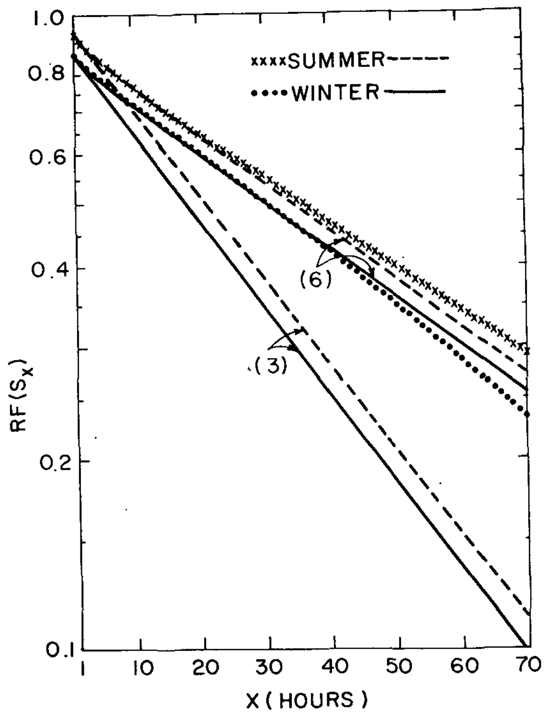

FIG. 3. Relative frequencies of x hours of consecutive successes, in winter (dots) and in summer (X's), when no precipitation (none) is regarded as a success. The solid lines are solutions to Eqs. (3) and (6) for winter and the dashed lines are for summer (see text).

TABLE 5. Median values of $RF(S|S_x)$ obtained from the data sample (Tables 2, 3 and 4) probability estimates $\hat{P}(S|S_x)$ determined from subjectively drawn curves of the medians shown in Fig. 2.

Season		x (hours) 0	1	2	3	4	5	6	7	8	9	10	11	≥12
Winter	None													
	Median	0.865	0.969	0.976	0.979	0.980	0.981	0.981	0.982	0.982	0.982	0.983	0.983	0.983
	$\hat{P}(S\|S_x)$	0.865	0.969	0.976	0.979	0.980	0.981	0.981	0.982	0.982	0.982	0.983	0.983	0.983
	Gringorten		0.935	0.953	0.959	0.963	0.966	0.968	0.971	0.972	0.974	0.975	0.976	0.976
	L+M+H													
	Median	0.135	0.803	0.840	0.850	0.856	0.860	0.859	0.861	0.858	0.857	0.853	0.855	0.855
	$\hat{P}(S\|S_x)$	0.135	0.803	0.840	0.850	0.856	0.860	0.860	0.860	0.860	0.860	0.860	0.860	0.860
	Gringorten		0.485	0.738	0.734	0.802	0.808	0.810						
	M+H													
	Median	0.00630	0.396	0.488	0.500	0.600	0.595	0.667	0.333	0.500	0	0	0	0
	$\hat{P}(S\|S_x)$	0.00630	0.396	0.475	0.540	0.580	0.620	0.630	0.640	0.650	0.650	0.650	0.650	0.650
Summer	None													
	Median	0.930	0.970	0.975	0.977	0.979	0.980	0.981	0.981	0.982	0.982	0.982	0.982	0.983
	$\hat{P}(S\|S_x)$	0.930	0.970	0.975	0.977	0.979	0.980	0.981	0.981	0.982	0.982	0.982	0.982	0.983
	L+M+H													
	Median	0.0696	0.633	0.710	0.744	0.788	0.802	0.816	0.815	0.810	0.813	0.812	0.815	0.827
	$\hat{P}(S\|S_x)$	0.0696	0.633	0.715	0.755	0.785	0.805	0.815	0.825	0.830	0.830	0.830	0.830	0.830
	M+H													
	Median	0.00676	0.282	0.491	0.517	0.531	0.500	0.333	0	0	0	0	0	0
	$\hat{P}(S\|S_x)$	0.00676	0.282	0.450	0.520	0.545	0.560	0.575	0.585	0.590	0.595	0.600	0.600	0.600

where $P(S|S_1)$ is the probability of a success given that a success occurred the previous hour, and $P(S|S_x)$ is the probability of a success following x hours of unbroken successes.

Curves were drawn for the points in Fig. 2 and estimates of $P(S|S_x)$ for use in Eqs. (4)–(6) were obtained from the curves. These values are given in Table 5 in the rows labeled $\hat{P}(S|S_x)$. The upper curves follow the median values exactly because the medians are based on large samples of observations of no precipitation. The middle curves, depicting conditional probabilities of precipitation, are based on smaller numbers of observations. Their shapes are more uncertain but they most likely represent good estimates of true conditional probabilities. The lowest curves (M+H) in Fig. 2 do not provide precise estimates of the true conditional probabilities because even a 13-year sample of hourly observations at nine stations is too small to accurately estimate conditional probabilities of moderate or heavy precipitation. Extremely large samples of data are required to estimate probabilities of such rare events from relative frequencies. These curves were drawn subjectively on the assumption that their shape would not depart significantly from those based on the more frequently occurring categories.

The curves in Fig. 2 are dashed where relative frequencies are based on less than 30 cases. For example the (M+H) curve through the first seven triangles plotted in Fig. 2a is solid because these points are based on more than 30 unbroken periods of moderate or heavy precipitation. The curve is dashed beyond the seventh point because there were less than 30 cases of more than seven consecutive hours of moderate or heavy precipitation, in winter, in the 13-year data sample. The dashed portion of the curves in Fig. 2 are questionable but necessary to the solution of Eq. (6).

Solutions of the higher order model [Eq. (6)] are shown by the curves in Figs. 3–5. They are in better agreement with the relative frequencies than curves based on Eq. (3). This is bound to be the case, because the probability estimates are based more closely on the relative frequencies. Modeling is involved only in the smoothing of the relative frequencies and in the assumption that $P(S|S_x)$ is constant beyond 12 h.

Gringorten (1968) employed Monte Carlo simulation to generate correlated values by an Ornstein-Uhlenbeck process. Distributions of the values were presented graphically for use in estimating probabilities. Some estimates determined from his graph are shown in Table 5 for winter (none) and winter (L+M+H). His estimates are lower than those obtained from the data but this is consistent with expectations because they represent duration probabilities, not persistence probabilities.

5. Runs

a. Observed

Another way of examining persistence is to consider the number of runs of exactly x hours in length, i.e.,

$n(FS_1F), n(FS_2F), \ldots, n(FS_xF)$. The relative frequency of runs is given by

$$RF(FS_xF) = \frac{n(FS_xF)}{N-13(x-1)} \approx \frac{n(FS_xF)}{N}, \quad (7)$$

where $n(FS_xF)$ is the observed number of runs of exactly x hours in length and N is the total number of hours in the data sample.

The observed number of runs, based on 28 080 h of winter observations, and 28 704 h of summer observations, at each of the nine stations are given, for selected hours, in Tables 6, 7 and 8. The median values are indicated with asterisks. Although the frequencies are based on more than 28 000 observations at each station and season, there are large sampling variations. For example, in Table 6 in winter, LGA had 64 runs of 2 h and the nearby station of JFK had 91; BAL had 18 runs of 3 h while the nearby stations of ADW and DCA had 41 and 42, respectively; and PHL had 4 runs of 18 h but $2\frac{1}{2}$ times as many (10 runs) of 36 h.

All nine stations contributed at least one median number of runs indicating some justification for applying the same model at all stations. The sample variations indicate that model estimates may yield at least as good estimates of run lengths as a 13-year data sample.

A model was considered that is very similar to Eq. (6). This model requires estimates of the conditional probabilities $\hat{P}(S|FS_x)$. Relative frequencies of success given a failure and x hours of successes were determined

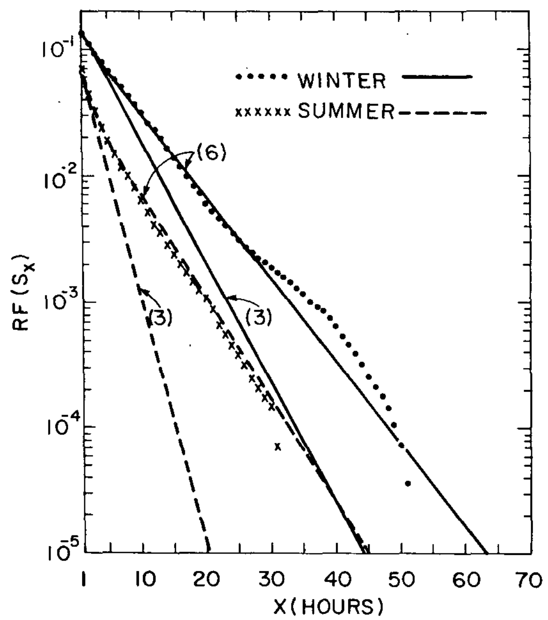

FIG. 4. Relative frequencies of x hours of consecutive successes, in winter (dots) and in summer (×'s), when precipitation (L + M + H) is regarded as a success. The solid lines are solutions to Eqs. (3) and (6) for winter and the dashed lines are for summer (see text).

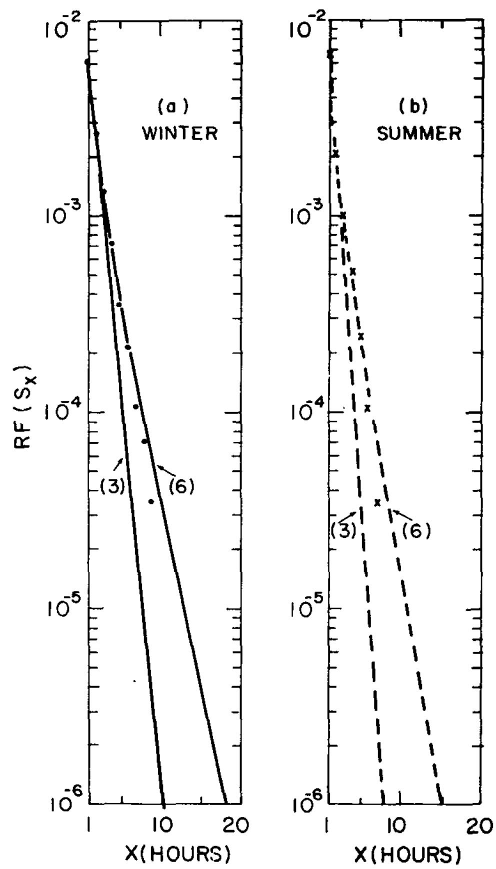

FIG. 5. Relative frequencies of x hours of consecutive successes, (a) in winter (dots) and (b) in summer (×'s) when moderate or heavy precipitation (M + H) is regarded as a success. The solid lines are solutions to Eqs. (3) and (6) for winter and the dashed lines are for summer (see text).

from the data with the expression

$$RF(S|FS_x) = \frac{n(S|FS_x)}{N} = \frac{n(FS_{x+1})}{n(FS_x)}. \quad (8)$$

These relative frequencies, a selection of which are given in Tables 9–11, were used to obtain the required conditional probabilities.

b. Modeled

The probability of a run of x hours in length $P(FS_xF)$, is the probability that there will be a failure followed by x successes followed by another failure. This might be estimated for runs of length 1, 2 and x hours, respectively, as follows:

$$\hat{P}(FS_1F) = P(F)P(S|F)P(F|FS_1), \quad (9)$$

$$\hat{P}(FS_2F) = P(F)P(S|F)P(S|FS_1)P(F|FS_2), \quad (10)$$

$$\hat{P}(FS_xF) = P(F)P(S|F)P(S|FS_1)P(S|FS_2) \ldots P(S|FS_{x-1})P(F|FS_x), \quad (11)$$

TABLE 6. Observed number of runs $n(FS_xF)$ of x hours in length observed in the data sample and estimated through the use of Eq. (12) when no precipitation (none) is considered a success. Median values are identified with asterisks.

		Run length (hours)										
Season	Station	1	2	3	4	5	6	12	18	24	30	36
Winter	LGA	169	64	38	29*	25	11	5	7	5	4	6
	JFK	172*	91	42*	23	25	20	6*	4	5	2	7
	EWR	179	73	43	24	28	16*	7	5	4*	3*	5
	PHL	203	83	47	36	19	16*	6*	4	5	7	10
	BAL	133	63	18	22	20*	12	6*	6*	4*	3*	3
	ADW	140	74*	41	27	20*	17	3	6*	4*	3*	3
	DCA	161	68	42*	23	14	17	10	7	1	6	4*
	RIC	172*	75	35	34	20*	15	2	7	4*	3*	3
	RDU	188	91	51	36	30	17	6*	1	4*	4	0
	Median	172	74	42	29	20	16	6	6	4	3	4
	$\hat{n}(FS_xF)$	169	77	41	27	22	17	6	5	5	4	4
Summer	LGA	171	93	49*	44	21	12	10	8	2	8	5
	JFK	152	80	44	32	29	17	16	3	3	2	5
	EWR	152	74	58	25	29	30	12*	7*	1	5	0
	PHL	154	70	32	29	28*	19	12*	3	8	2	1
	BAL	131	65	39	22	21	22*	6	4	6	3*	2*
	ADW	118	73	52	20	15	23	9	3	8	2	2*
	DCA	143	61	49*	44	21	30	14	8	7*	5	0
	RIC	147*	72*	56	33*	32	26	15	8	8	5	6
	RDU	143	70	44	35	38	19	14	9	7*	2	7
	Median	147	72	49	33	28	22	12	7	7	3	2
	$\hat{n}(FS_xF)$	136	70	46	34	26	20	6	5	5	4	4

where $P(S|F)$ is the probability of a success given that a failure occurred the previous hour, $P(S|FS_1)$ is the probability of a success given that a success occurred and a failure occurred two hours earlier..., and $P(F|FS_x)$ is the probability of a failure given that x successes occurred the previous x consecutive hours and

TABLE 7. As in Table 6 except when precipitation of any intensity (L+M+H) is considered a success.

		Run length (hours)										
Season	Station	1	2	3	4	5	6	12	18	24	30	36
Winter	LGA	258*	106*	65	52	38	32	19	4	1	0	0
	JFK	258*	118	64	68	33	26	15	7	1	0	0
	EWR	272	111	68*	65	30	21	16	4	3	0	0
	PHL	304	120	68*	54	36*	34	8	9	3	0	0
	BAL	205	88	70	47	28	27	20	6	2*	0	2
	ADW	244	96	67	49	38	36	11	1	2*	2	0
	DCA	234	106*	72	50*	30	30*	12*	3	5	1	2
	RIC	258*	98	59	44	41	21	8	6	2*	1	0
	RDU	272	136	84	44	56	37	10	5*	0	0	0
	Median	258	106	68	50	36	30	12	5	2	0	0
	$\hat{n}(FS_xF)$	263	108	69	51	40	32	13	5	2	0.8	0.3
Summer	LGA	412	149	81	41	31	12	3*	4	0	0	0
	JFK	400	153*	81	44	30	16	0	0	1	0	0
	EWR	392	160	77	42*	26*	17	4	1*	0	0	0
	PHL	378	143	77	46	37	20	5	0	0	0	0
	BAL	359	135	82	40	23	15*	2	1*	0	0	0
	ADW	390	149	77	32	23	11	6	1*	0	0	0
	DCA	424	169	78*	35	20	14	3*	1*	0	0	0
	RIC	399*	181	77	46	26*	14	1	0	0	0	0
	RDU	410	176	93	59	35	16	4	0	0	0	0
	Median	399	153	78	42	26	15	3	1	0	0	0
	$\hat{n}(FS_xF)$	399	159	82	45	28	19	4	1	0.4	0.1	0.04

Table 8. As in Table 6 except when moderate and heavy (M+H) precipitation is considered a success.

Season	Station	Run length (hours) 1	2	3	4	5	6	12	18	24
Winter	LGA	87	22	8	8	6	0	0	0	0
	JFK	73*	21	5	3	0	0	0	0	0
	EWR	68	20*	12	5	5	0	0	0	0
	PHL	64	16	7	4*	2*	0	0	0	0
	BAL	64	13	8	3	0	0	0	0	0
	ADW	113	30	11	7	8	2	0	0	0
	DCA	85	14	10	2	1	1	0	0	0
	RIC	94	30	9*	7	3	2	0	0	0
	RDU	52	19	10	3	1	2	1	0	0
	Median	73	20	9	4	2	0	0	0	0
	$\hat{n}(FS_xF)$	74	20	8	4	2	1	0.1	0.009	0.0007
Summer	LGA	120	16*	6*	4	2	1*	0	0	0
	JFK	87	21	2	4	0	1*	0	0	0
	EWR	110	12	6*	1	1*	1*	0	0	0
	PHL	92	14	5	1	1*	0	0	0	0
	BAL	99	15	8	4	3	1*	1	0	0
	ADW	126	16*	9	4	1*	0	0	1	0
	DCA	105*	17	6*	3*	3	0	0	0	0
	RIC	133	21	5	2	4	1*	1	0	0
	RDU	94	8	6*	3*	0	0	0	0	0
	Median	105	16	6	3	1	1	0	0	0
	$\hat{n}(FS_xF)$	105	15	6	3	1	0.7	0.03	0.001	0.00008

a failure occurred $x+1$ hours earlier. The unconditional and conditional probabilities can be estimated from the relative frequencies but very large samples of data are required to obtain statistically stable relative frequencies of long runs, because they are rare events.

The points plotted in Fig. 6 show median relative frequencies of success after a failure and x hours of successes have occurred. The median values are given in Table 12. Smooth curves were subjectively drawn through the points in Fig. 6. The probabilities required for the solution of Eqs. (9), (10) and (11) were estimated from these curves. Once again, as in Fig. 2, the curves are dashed where relative frequencies are based on less than 30 cases to indicate that these portions of the (M+H) curves are possibly poor estimates of the true probabilities.

Table 9. Relative frequency of success $RF(S|FS_x)$ given a failure and x hours of success obtained from the data sample when no precipitation (none) is considered a success. Median values are identified with asterisks.

Season	Station	x (hours) 0	1	3	5	7	9	11	⩾12
Winter	LGA	0.185	0.778	0.928	0.946	0.958	0.970*	0.973	0.986
	JFK	0.198*	0.779	0.918*	0.944	0.975	0.971	0.975*	0.983*
	EWR	0.188	0.773	0.920	0.940	0.948	0.977	0.992	0.981*
	PHL	0.209	0.749	0.910	0.957	0.970*	0.976	0.970	0.983*
	BAL	0.174	0.799	0.961	0.953	0.964	0.957	0.976	0.982
	ADW	0.198*	0.800	0.915	0.952	0.973	0.955	0.976	0.990
	DCA	0.198*	0.775*	0.913	0.967	0.974	0.959	0.968	0.970
	RIC	0.197	0.760	0.925	0.950*	0.978	0.985	0.955	0.993
	RDU	0.236	0.756	0.896	0.926	0.955	0.967	0.984	0.980
Summer	LGA	0.367*	0.786	0.908	0.952	0.948	0.968	0.980	0.971
	JFK	0.364	0.810	0.922	0.941*	0.971	0.954	0.968*	0.956
	EWR	0.361	0.807	0.897	0.940	0.955	0.972	0.972	0.966*
	PHL	0.367*	0.797	0.940	0.941*	0.958	0.972	0.959	0.966*
	BAL	0.364	0.818	0.925	0.954	0.947	0.971*	0.981	0.983
	ADW	0.397	0.840	0.905	0.968	0.959*	0.978	0.982	0.976
	DCA	0.410	0.819	0.916*	0.957	0.961	0.971*	0.964	0.963
	RIC	0.399	0.815*	0.903	0.934	0.979	0.956	0.961	0.959
	RDU	0.415	0.827	0.928	0.929	0.960	0.984	0.968*	0.965

TABLE 10. As in Table 9 except when precipitation of any intensity (L+M+H) is considered a success.

Season	Station	x (hours) 0	1	3	5	7	9	11	≥12
Winter	LGA	0.0318	0.661	0.836	0.864*	0.933	0.868	0.858	0.826
	JFK	0.0322	0.668	0.840	0.877	0.881	0.863	0.791	0.835
	EWR	0.0327	0.652*	0.830	0.887	0.902	0.871	0.801	0.853
	PHL	0.0333	0.623	0.822*	0.862	0.880*	0.883	0.860*	0.913
	BAL	0.0271	0.688	0.808	0.887	0.855	0.899	0.919	0.823
	ADW	0.0284	0.650	0.813	0.843	0.893	0.869	0.926	0.875
	DCA	0.0291	0.672	0.807	0.880	0.859	0.870*	0.857	0.867
	RIC	0.0293	0.640	0.837	0.841	0.837	0.809	0.917	0.909
	RDU	0.0311*	0.647	0.769	0.762	0.838	0.906	0.877	0.859*
Summer	LGA	0.0301	0.484	0.660	0.733	0.863	0.804	0.719	0.870
	JFK	0.0301	0.498*	0.668	0.748	0.822	0.745	0.839	0.769
	EWR	0.0298	0.504	0.676	0.782	0.776	0.913	0.771*	0.852
	PHL	0.0285	0.503	0.678	0.681	0.898	0.756	0.724	0.762
	BAL	0.0268	0.499	0.632	0.772	0.873	0.766	0.760	0.895
	ADW	0.0275	0.471	0.613	0.744*	0.857	0.722	0.818	0.667
	DCA	0.0296	0.465	0.608	0.767	0.846*	0.789*	0.760	0.842*
	RIC	0.0298*	0.498*	0.642*	0.717	0.769	0.813	0.941	0.938
	RDU	0.0311	0.507	0.620	0.624	0.738	0.864	0.800	0.667

Table 12 shows values of $P(S|FS_x)$ that were estimated from the curves shown in Fig. 6. The conditional probabilities always increase for at least 7 h and most of the values increase for at least 12 h.

The values found in Table 12 were used to solve Eq. (11). By substituting $\hat{P}(FS_xF)$ from Eq. (11) for $RF(FS_xF)$ in Eq. (7) the following expression is obtained for estimating $n(FS_xF)$:

$$\hat{n}(FS_xF)=\hat{P}(FS_xF)N. \qquad (12)$$

Solutions to Eq. (12) are given in Tables 6–8. The agreements between the observed number of runs and those calculated from Eq. (12) are very good. It should be understood that this is not an independent test of Eq. (12) but rather a subjective fitting to the data to obtain conditional probabilities and an objective method for finding the desired probability estimates.

6. Recurrence

a. Observed

The relative frequency of the recurrence of a success l hours later given that a success occurred, $RF(S_l|S)$, can be determined from the data by dividing the number of occurrences of successes spaced l hours apart, $n(SS_l)$, by the total number of successes $n(S)$, i.e.,

$$RF(S_l|S)=\frac{n(SS_l)}{n(S)}. \qquad (13)$$

TABLE 11. As in Table 9 except when moderate or heavy (M+H) precipitation is considered a success.

Season	Station	x (hours) 0	1	2	3	4	5	6	7
Winter	LGA	0.00470	0.335*	0.500	0.636	0.428	0	0	0
	JFK	0.00373	0.298	0.322	0.500*	0.400	0.500	0	0
	EWR	0.00398*	0.387	0.534	0.478	0.545	0.166	0	0
	PHL	0.00337	0.319	0.466	0.500*	0.428	0.333	0	0
	BAL	0.00318	0.280	0.480*	0.333	0.250	0	0	0
	ADW	0.00634	0.357	0.523	0.666	0.681	0.466*	0.714	0.600
	DCA	0.00412	0.260	0.533	0.375	0.666	0.750	0.666	0.500
	RIC	0.00524	0.356	0.423	0.590	0.461*	0.500	0.333	0
	RDU	0.00315	0.409	0.472	0.411	0.571	0.750	0.333	0
Summer	LGA	0.00526	0.200	0.467	0.571	0.500*	0.500	0.500	0
	JFK	0.00403	0.243	0.250	0.714	0.200	0	0	0
	EWR	0.00459	0.160	0.428	0.333	0.667	0.500	0	0
	PHL	0.00396	0.185	0.333	0.285	0.500*	0	0	0
	BAL	0.00460*	0.244	0.531	0.529*	0.556	0.400*	0.500	0
	ADW	0.00555	0.202*	0.500	0.437	0.428	0.666	0.500	0
	DCA	0.00474	0.222	0.433*	0.538	0.571	0.250	0	0
	RIC	0.00594	0.213	0.416	0.666	0.800	0.500	0.750	0.666
	RDU	0.00389	0.153	0.529	0.333	0	0	0	0

TABLE 12. Median values of $RF(S|FS_x)$ obtained from the data sample (Tables 9, 10 and 22) and probability estimates $\hat{P}(S|FS_x)$ determined from subjectively drawn curves of the medians shown in Fig. 6.

Season		$P(F)$	x (hours) 0	1	2	3	4	5	6	7	8	9	10	11	≥12
Winter	None														
	Median	0.135	0.198	0.775	0.867	0.918	0.941	0.950	0.960	0.970	0.971	0.970	0.980	0.975	0.983
	$\hat{P}(S\|FS_x)$		0.198	0.775	0.867	0.918	0.941	0.950	0.960	0.970	0.973	0.977	0.979	0.981	0.983
	L+M+H														
	Median	0.865	0.0311	0.652	0.782	0.822	0.834	0.864	0.868	0.880	0.867	0.870	0.844	0.860	0.859
	$\hat{P}(S\|FS_x)$		0.0311	0.652	0.780	0.820	0.837	0.850	0.858	0.860	0.860	0.860	0.860	0.860	0.860
	M+H														
	Median	0.99370	0.00398	0.335	0.480	0.500	0.461	0.466	0	0	0	0	0	0	0
	$\hat{P}(S\|FS_x)$		0.00398	0.335	0.450	0.515	0.555	0.590	0.610	0.630	0.640	0.645	0.650	0.650	0.650
Summer	None														
	Median	0.070	0.367	0.815	0.884	0.916	0.939	0.941	0.950	0.959	0.967	0.971	0.971	0.968	0.966
	$\hat{P}(S\|FS_x)$		0.367	0.815	0.884	0.913	0.930	0.943	0.953	0.960	0.967	0.973	0.977	0.980	0.983
	L+M+H														
	Median	0.930	0.0298	0.498	0.598	0.642	0.716	0.744	0.808	0.846	0.810	0.789	0.816	0.771	0.842
	$\hat{P}(S\|FS_x)$		0.0298	0.498	0.598	0.655	0.710	0.748	0.771	0.780	0.804	0.815	0.823	0.829	0.830
	M+H														
	Median	0.99324	0.00460	0.202	0.433	0.529	0.500	0.400	0	0	0	0	0	0	0
	$\hat{P}(S\|FS_x)$		0.00460	0.202	0.433	0.495	0.528	0.548	0.562	0.572	0.582	0.591	0.596	0.598	0.600

Conditional recurrence relative frequencies based on 13 years of hourly observations taken at each of the nine stations are given, for selected hours, in Tables 13–15. The median values for each season are also given in the tables and plotted in Figs. 7–9. The relative frequencies given in Table 13 are based on from 26 103 to 20 180 recurrences of no precipitation. Because the data sample is large these values are judged to be good estimates of the true probabilities. With few exceptions, there is little variation from the median values for the no precipitation (Table 13) and all precipitation categories (Table 14). Conditional relative frequencies of the less frequent category of moderate or heavy precipitation are much more variable. The moderate

TABLE 13. Relative frequency of the recurrence of a success $RF(S_l|S)$ l hours after a success has occurred observed in the data sample and estimated through the use of Eq. (14) when no precipitation (none) is considered a success. Median values are identified with asterisks.

Season	Station	$P(S)$	l (hours) 1	2	3	4	5	6	9	12	18	24	30	36	48	60	71
Winter	LGA	0.853	0.968	0.954	0.943	0.933	0.924	0.916	0.896	0.881	0.864	0.855	0.850	0.849	0.851	0.850	0.849
	JFK	0.860	0.968	0.953	0.943	0.934	0.926	0.918	0.899	0.886	0.870	0.862	0.857	0.856	0.858	0.856	0.854
	EWR	0.851	0.967	0.953	0.942	0.932	0.924	0.916	0.896	0.881	0.864	0.852	0.847	0.846	0.847	0.848	0.846
	PHL	0.862	0.967	0.954	0.944	0.936	0.928	0.921	0.903	0.890	0.873	0.865	0.861	0.860	0.861	0.861	0.860
	BAL	0.865*	0.973	0.960	0.949	0.940*	0.932*	0.925*	0.908*	0.895*	0.876*	0.869*	0.865*	0.864*	0.864*	0.863*	0.861*
	ADW	0.874	0.971	0.959	0.948*	0.940*	0.934	0.926	0.911	0.901	0.883	0.877	0.875	0.873	0.873	0.870	0.869
	DCA	0.871	0.971	0.958	0.948*	0.940*	0.933	0.926	0.910	0.900	0.882	0.875	0.872	0.870	0.871	0.868	0.868
	RIC	0.870	0.971	0.959	0.949	0.939	0.932*	0.925*	0.909	0.898	0.882	0.874	0.871	0.869	0.867	0.866	0.863
	RDU	0.883	0.969*	0.956*	0.948*	0.941	0.934	0.930	0.918	0.908	0.895	0.888	0.884	0.880	0.879	0.879	0.878
	Median	0.865	0.969	0.956	0.948	0.940	0.932	0.925	0.908	0.895	0.876	0.869	0.865	0.864	0.864	0.863	0.861
	Eq. (14)		0.961	0.945	0.934	0.926	0.919	0.913	0.901	0.893	0.883	0.877	0.873	0.871	0.868	0.867	0.866
Summer	LGA	0.924	0.970*	0.962*	0.957*	0.953*	0.950*	0.947*	0.939	0.934	0.929	0.925	0.924	0.924	0.921	0.922	0.923
	JFK	0.924	0.970*	0.961	0.955	0.952	0.948	0.945	0.938	0.934	0.929	0.925	0.923	0.923	0.921	0.922	0.922
	EWR	0.924	0.970*	0.961	0.955	0.952	0.948	0.946	0.939	0.934	0.928	0.925	0.923	0.923	0.921	0.922	0.923
	PHL	0.928	0.971	0.963	0.958	0.953*	0.950*	0.947*	0.942*	0.937	0.932	0.929	0.928	0.928	0.926	0.927*	0.926
	BAL	0.931	0.973	0.965	0.960	0.956	0.953	0.951	0.945	0.941	0.936*	0.933*	0.932	0.931	0.931	0.931	0.930
	ADW	0.935	0.972	0.964	0.959	0.957	0.954	0.952	0.948	0.944	0.940	0.937	0.936	0.936	0.935	0.934	0.933
	DCA	0.933	0.970*	0.962*	0.957*	0.954	0.952	0.950	0.944	0.941	0.937	0.934	0.933	0.933	0.932	0.932	0.931
	RIC	0.930*	0.970*	0.961	0.955	0.952	0.950*	0.948	0.945	0.940	0.936*	0.933*	0.931*	0.929*	0.931	0.929	0.928*
	RDU	0.930*	0.969	0.958	0.952	0.948	0.946	0.945	0.941	0.938*	0.936*	0.934	0.930*	0.930	0.929*	0.927*	0.928*
	Median	0.930	0.970	0.962	0.957	0.953	0.950	0.947	0.942	0.938	0.936	0.933	0.930	0.929	0.929	0.927	0.928
	Eq. (14)		0.976	0.967	0.961	0.956	0.953	0.950	0.944	0.940	0.936	0.934	0.933	0.932	0.931	0.931	0.931

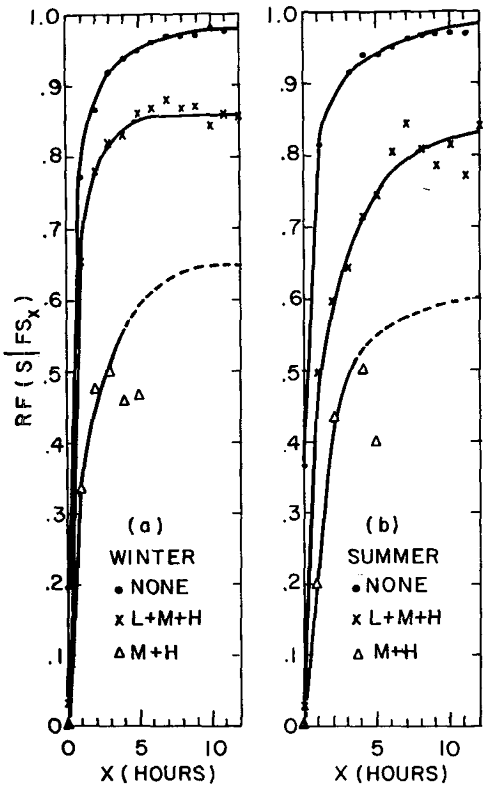

FIG. 6. Relative frequencies of success given exactly x hours of consecutive success have occurred, in winter (a) and summer (b). The curves were subjectively drawn. All points to the right of the solid portions of the curves are based on fewer than 30 cases.

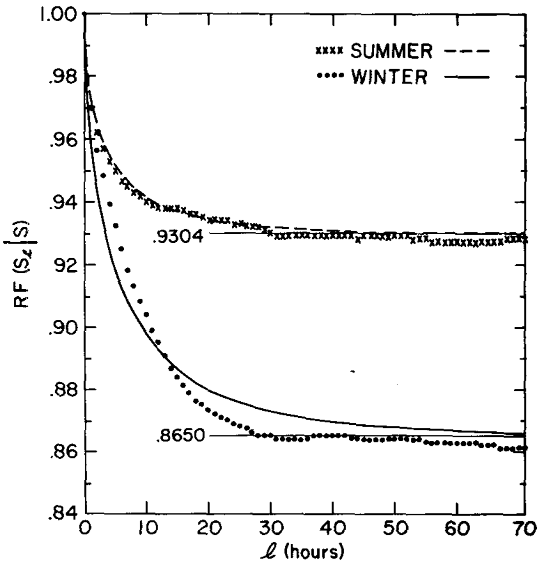

FIG. 7. Relative frequencies of a success, no precipitation (none), l hours later given a success has occurred, in winter (dots) and in summer (X's). The solid curve is the solution to Eq. (14) with $a = 0.337$ for winter, and the dashed curve is for summer with $a = 0.419$.

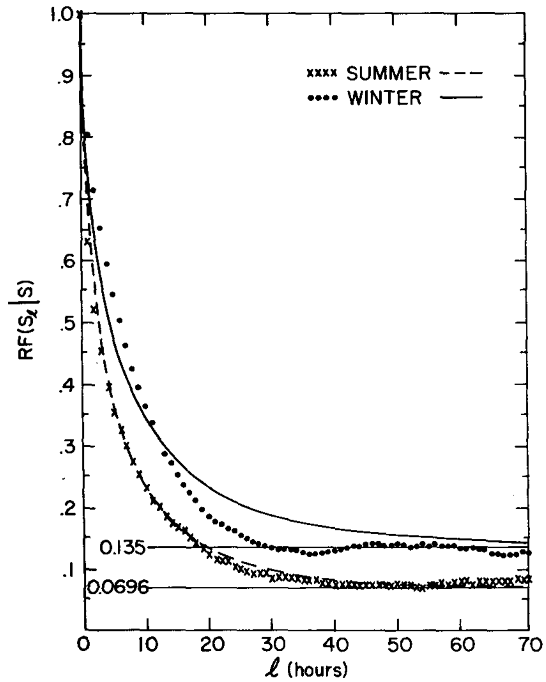

FIG. 8. Relative frequencies of a success, precipitation (L + M + H), l hours later given a success has occurred, in winter (dots) and in summer (X's). The solid curve is the solution to Eq. (14) with $a = 0.337$ for winter, and the dashed curve is for summer with $a = 0.419$.

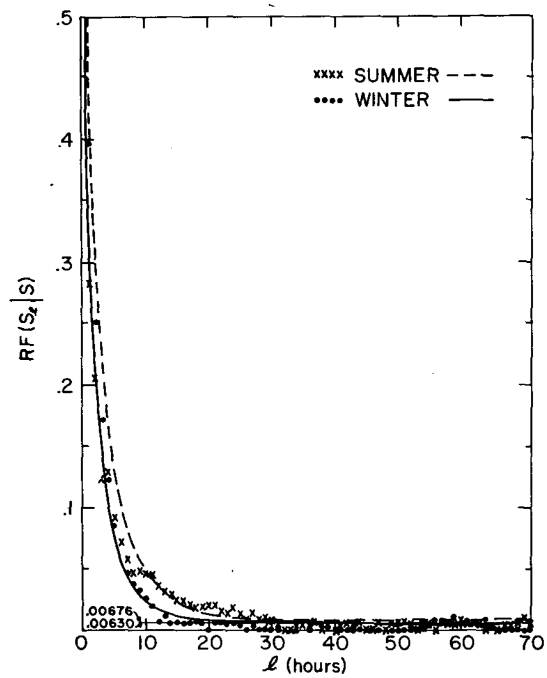

FIG. 9. Relative frequencies of a success, moderate or heavy precipitation, l hours later given a success has occurred, in winter (dots) and in summer (X's). The solid curve is the solution to Eq. (14) with $a = 0.969$ for winter, and the dashed curve is for summer with $a = 0.761$.

TABLE 14. As in Table 13 except when all precipitation (L+M+H) is considered a success.

Season	Station	$P(S)$	l (hours) 1	2	3	4	5	6	9	12	18	24	30	36	48	60	71
Winter	LGA	0.1469	0.815	0.734	0.670	0.611	0.560	0.515	0.395*	0.314	0.213	0.162*	0.134	0.130	0.143	0.141	0.137
	JFK	0.1404	0.803*	0.715*	0.652*	0.596	0.547*	0.502*	0.384	0.308	0.212*	0.161	0.134	0.127	0.141	0.136*	0.127
	EWR	0.1493	0.813	0.734	0.670	0.616	0.570	0.523	0.410	0.326	0.230	0.165	0.137	0.133	0.140*	0.146	0.138
	PHL	0.1377	0.792	0.714	0.648	0.597*	0.553	0.506	0.395*	0.312	0.211	0.162*	0.138*	0.134	0.144	0.145	0.143
	BAL	0.1353	0.827	0.742	0.677	0.620	0.570	0.524	0.412	0.330	0.214	0.169	0.147	0.141	0.145	0.138	0.131
	ADW	0.1258	0.803*	0.714	0.641	0.582	0.534	0.487	0.388	0.315*	0.196	0.153	0.138*	0.127	0.128	0.116	0.111
	DCA	0.1287	0.803*	0.715*	0.648	0.594	0.547*	0.500	0.394	0.317	0.209	0.159	0.144	0.128*	0.139	0.121	0.128*
	RIC	0.1297	0.803*	0.724	0.656	0.592	0.542	0.495	0.391	0.319	0.215	0.165	0.139	0.128*	0.117	0.116	0.100
	RDU	0.1164	0.764	0.669	0.608	0.554	0.502	0.466	0.378	0.306	0.206	0.156	0.123	0.101	0.0958	0.0948	0.0915
	Median	0.1353	0.803	0.715	0.652	0.597	0.547	0.502	0.395	0.315	0.212	0.162	0.138	0.128	0.140	0.136	0.128
	Eq. (14)		0.753	0.650	0.580	0.525	0.482	0.446	0.367	0.315	0.250	0.212	0.189	0.174	0.156	0.147	0.143
Summer	LGA	0.0758	0.633*	0.534	0.473	0.424	0.387	0.356	0.255*	0.199	0.137	0.102	0.0924	0.0846*	0.0533	0.0722	0.0905
	JFK	0.0764	0.636	0.525	0.458	0.417	0.374	0.337	0.259	0.210	0.146	0.109	0.0858	0.0803	0.0634	0.0807*	0.0908
	EWR	0.0762	0.639	0.527	0.459	0.415	0.377	0.344	0.265	0.203	0.134	0.106	0.0850	0.0809	0.0599	0.0777	0.0946
	PHL	0.0723	0.634	0.524	0.457	0.399*	0.355*	0.328*	0.260	0.200	0.139	0.104	0.0867	0.0872	0.0761*	0.0867	0.0819
	BAL	0.0688	0.637	0.522*	0.456*	0.404	0.368	0.341	0.264	0.201*	0.142*	0.107*	0.101	0.0881	0.0871	0.0896	0.0871
	ADW	0.0649	0.604	0.480	0.417	0.382	0.336	0.315	0.249	0.201*	0.145	0.108	0.0988	0.0983	0.0838	0.0816	0.0714
	DCA	0.0673	0.590	0.473	0.405	0.368	0.335	0.306	0.232	0.189	0.129	0.0943	0.0865*	0.0870	0.0782	0.0823	0.0849*
	RIC	0.0695	0.601	0.476	0.405	0.365	0.329	0.303	0.244	0.201*	0.153	0.119	0.0853	0.0677	0.0933	0.0787	0.0727
	RDU	0.0696*	0.584	0.445	0.360	0.310	0.276	0.261	0.212	0.178	0.151	0.125	0.0836	0.0761	0.0746	0.0601	0.0756
	Median	0.0696	0.633	0.522	0.456	0.399	0.355	0.328	0.255	0.201	0.142	0.107	0.0865	0.0846	0.0761	0.0807	0.0849
	Eq. (14)		0.681	0.558	0.476	0.415	0.368	0.330	0.251	0.201	0.149	0.116	0.0991	0.0891	0.0788	0.0742	0.0722

or heavy category recurrence relative frequencies are based on from 155 to 0 recurrences. Many of these values are poor estimates of the true probabilities.

b. Modeled

McAllister (1969) proposed an expression of the form

$$\hat{P}(S_{t+l}|S_t)=P(S_{t+l})+[1-P(S_t)]e^{-al^b} \tag{14}$$

for estimating recurrence probabilities of cloud cover. He used $a=0.263$ and $b=0.632$ as the best estimates of the parameters. Gringorten (1971) showed that Eq. (14) yields probability estimates very close to those obtained from the bivariate normal distribution if the parameter b is fixed at 0.620 and a is allowed to vary with the climatic frequency of the initial event and the basic persistence of the element. Gringorten (1972) proposed an expression of the form

$$y(t+l|t)=\frac{y_{t+l}-\rho y_t}{(1-\rho^2)^{\frac{1}{2}}}, \tag{15}$$

where $y(t+l|t)$ is the normalized variable (mean 0.0, variance 1.0) corresponding to the conditional probability $P(S_{t+l}|S_t)$, Y_{t+l} is the normalized value corresponding to the unconditional probability, $P(S_{t+l})$, and y_t is the normalized value corresponding to the unconditional probability $P(S_t)$. The values estimated by Eqs. (14) and (15) and those obtained from tables of the bivariate normal distribution function are almost identical. Curves obtained from solutions of Eq. (14) with $b=0.620$ are shown in Figs. 7–9. Eq. (14) was used because it is easier to solve than Eq. (15). Because the normal probability integral can be es-

TABLE 15. As in Table 13 except when moderate or heavy (M+H) precipitation is considered a success.

Season	Station	$P(S)$	l (hours) 1	2	3	4	5	6	9	12	18	24	30	36	48	60	71
Winter	LGA	0.00773	0.396*	0.290	0.171*	0.0968	0.0553	0.0507	0.0138	0	0	0.00922	0.00461	0.00461	0.00461	0.00922	0
	JFK	0.00573	0.354	0.193	0.149	0.112	0.0932	0.0745	0.0124	0.00621	0.00621	0.00621	0	0	0	0.0124	0.00621
	EWR	0.00705	0.439	0.278	0.187	0.131	0.0960	0.0707	0.0354	0.0455	0.0202	0.00505*	0.00505	0	0	0.0152	.00505
	PHL	0.00534	0.373	0.213	0.127	0.0800	0.0467	0.0400	0.0333*	0.0333	0.0200	0	0	0	0	0	0
	BAL	0.00484	0.346	0.140	0.125	0.110	0.0735	0.0588	0.0147	0.00735*	0.00735*	0	0	0.0147	0	0	0
	ADW	0.0118	0.468	0.344	0.251	0.184	0.145	0.133	0.0453	0.0363	0.0121	0.00906	0.00302	0.00302	0.0151	0.00906*	0.00302
	DCA	0.00630*	0.350	0.232	0.175	0.141	0.0847*	0.0621*	0.0169	0.00565	0	0	0	0	0.00565	0	0
	RIC	0.00865	0.399	0.251*	0.169	0.123*	0.0741	0.0412	0.0494	0.0165	0	0.00412	0.0123	0.00823	0	0.0165	0
	RDU	0.00573	0.453	0.273	0.242	0.205	0.143	0.0932	0.0497	0.00621	0.0124	0.00621	0	0	0	0.00621	0.00621
	Median	0.00630	0.396	0.251	0.171	0.123	0.0847	0.0621	0.0333	0.00735	0.00735	0.00505	0	0	0	0.00906	0
	Eq. (14)		0.383	0.230	0.153	0.107	0.0780	0.0587	0.0289	0.0171	0.00926	0.00725	0.00664	0.00643	0.00632	0.00630	0.00630
Summer	LGA	0.00728	0.282*	0.206*	0.124*	0.139	0.0909*	0.0766	0.0670	0.0478	0.0287	0.0287	0.00957*	0	0	0.00957*	0
	JFK	0.00547	0.268	0.134	0.108	0.0637	0.0637	0.0255	0.0382	0.0318	0.0318	0.0318	0.0191	0	0	0.0127	0.00637
	EWR	0.00582	0.216	0.156	0.120	0.0838	0.0719	0.0539	0.0479	0.0359	0.0180*	0.0180	0.00599	0	0	0.0120	0
	PHL	0.00502	0.215	0.118	0.0833	0.0347	0.0417	0.0278	0.0486*	0.0208	0.00694	0.0208	0	0.00694	0	0.00694	0.0139
	BAL	0.00704	0.351	0.223	0.173	0.139	0.124	0.0990	0.0743	0 0446	0.0248	0.0347	0.00495	0.00990	0.0198	0.00990	0.0149
	ADW	0.00829	0.336	0.231	0.189	0.147	0.122	0.113	0.0756	0.0546	0.00840	0.0126*	0.0126	0	0.0126	0.0126	0.0168
	DCA	0.00676*	0.304	0.222	0.149	0.129*	0.0928	0.0722*	0.0464	0.0309	0.0309	0.00515	0.00515	0.0155	0	0.00515	0.0206
	RIC	0.00892	0.340	0.254	0.191	0.180	0.121	0.102	0.0742	0.0352*	0.0156	0.00391	0.0117	0.00391	0.0117	0	0
	RDU	0.00488	0.207	0.114	0.0643	0.0357	0.0500	0.0357	0	0.00714	0.00714	0.00714	0.0143	0.0214	0.00714	0	0.00714*
	Median	0.00676	0.282	0.206	0.124	0.129	0.0909	0.0722	0.0486	0.0352	0.0180	0.0126	0.00957	0.00391	0	0.00957	0.00714
	Eq. (14)		0.471	0.315	0.227	0.171	0.133	0.105	0.0576	0.0352	0.0171	0.0010	0.00864	0.00765	0.00699	0.00682	0.00678

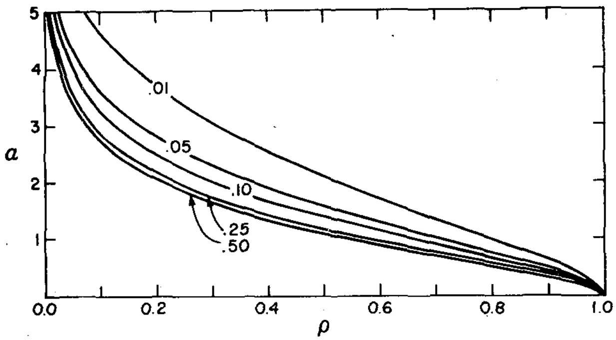

FIG. 10. The relationship between a in Eq. (14) and ρ when $P(s)$ equals 0.01, 0.05, 0.10, 0.25 and 0.50.

timated [see Eq. (26.2.22) from National Bureau of Standards (1964)] using the formula

$$x_p = t - \frac{a_0 + a_1 t}{1 + b_1 t + b_2 t^2} + \epsilon(p), \qquad (16)$$

where $t = [\ln(1/p^2)]^{\frac{1}{2}}$, $a_0 = 2.30753$, $a_1 = 0.27061$, $b_1 = 0.99229$, $b_2 = 0.04481$, $|\epsilon(p)| < 3 \times 10^{-3}$, Eq. (15) is also readily solved.

Because correlation (ρ) is a familiar parameter and a in McAllister's expression is not, some relationships between a and ρ are shown in Fig. 10.

The climatic frequencies of the events were substituted into Eq. (14) using $l = 12$ h and $b = 0.620$ and the equation was solved to find the parameter a. The a values and corresponding hour-to-hour ρ values are given in Table 16.

Eq. (14) was solved for lags from 1 to 71 h using the a values given in Table 16. The resulting curves are shown in Figs. 7–9. The fit to the summer relative frequencies is excellent. The fit to the winter values is not as good but possibly acceptable for many purposes.

TABLE 16. The a values used in Eq. (14) to find the curves shown in Figs. 7–9. The numbers in parentheses are corresponding hour-to-hour correlation coefficients.

Season	Precipitation category None	L+M+H	M+H
Winter	0.337 (0.922)	0.337 (0.922)	0.969 (0.827)
Summer	0.419 (0.911)	0.419 (0.911)	0.761 (0.874)

7. Remarks

The relative frequencies of persistence, runs and recurrence of precipitation along the east coast of the United States between New York and North Carolina, presented in this paper, are based on more than 250 000 hourly observations taken in winter and a similar number taken in summer. They are believed to be good approximations of the true probabilities when occurrences of the events exceed approximately 30. Since moderate or heavy precipitation occurs less than 1% of the time, even a quarter of a million observations are too few to estimate long runs of this event from the relative frequencies alone.

Models are presented for use in estimating joint and conditional probabilities. The estimates are usually in good agreement with the relative frequencies when the parameters are carefully chosen. However, the best choices for the central East Coast area of the United States may not be the best for other geographical areas. Future studies will be extended to other areas. Other weather elements are under study at the present time.

Acknowledgments. The authors are grateful to James F. Atkinson, Leonard J. Natoli and Kenneth C. Zwirble of Analysis and Computer Systems, Inc., and Miss Melinda A. Zouvelos, Student Aid, Lowell University, for expert computational support and to the USAF Environmental Technical Applications Center and the National Weather Service for hourly precipitation data.

REFERENCES

Gringorten, I. I., 1968: Estimating finite-time maxima and minima of a stationary Gaussian Ornstein-Uhlenbeck process by Monte Carlo simulation. *J. Amer. Stat. Assoc.*, **63**, 1517–1521.

——, 1971: Modelling conditional probability. *J. Appl. Meteor.*, **10**, 646–657.

——, 1972: Conditional probability for an exact noncategorized initial condition. *Mon. Wea. Rev.*, **100**, 796–798.

McAllister, C. R., 1969: Cloud-cover recurrence and diurnal variation. *J. Appl. Meteor.*, **8**, 769–777.

National Bureau of Standards, 1964: *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* U. S. Govt. Printing Office, Washington, D.C., 1046 pp.

U. S. Department of Commerce, 1975: *Federal Meteorological Handbook No. 1, Surface Observations.* U. S. Govt. Printing Office, Washington, D. C., 309 pp.