A climatological study of persistency and probability of precipitation in Sweden

By BERTIL ERIKSSON, *Swedish Meteorological and Hydrological Institute, Stockholm, Sweden*

(Manuscript received March 30, 1965)

ABSTRACT

Time series consisting of wet and dry days have been studied regarding persistency properties. The conditional probabilities calculated show that the time series approximately behave like simple Markov chains. Using the Markov model some probabilities regarding precipitation have been computed and the results are compared with empirical probabilities.

1. Introduction

The variation of meteorological variables may be considered as the sum of a systematic and a stochastic variation. The first part can be expressed by a function. For example the variation of temperature with latitude, altitude and the diurnal and seasonal variation can be described by a Fourier series. The latter variation must be described with the aid of probability distributions. Symbolically we may write:

$$W_i = F_i(s, t_y, t_d) + f_i(s, t_y, t_d), \qquad (1)$$

W_i denotes a certain weather element i, observed at a certain location s at a certain time of the day t_d and time of the year t_y. F_i is a function, for instance a trigonometric one and f_i is a probability distribution function, e.g. a binomial, normal or χ^2-function. The appearance and properties of these distributions have received insufficient attention in climatological monographs. The climate of a particular place is inadequately described if only the first term on the right-hand side of the above equation is given. The random variations f_i around the mean state are often considerable and should be expressed in mathematical terms or by a number of statistical parameters which express the most important character of the distribution.

Apart from this approach meteorological observations should also be regarded as time series. This ought to be done in order to clarify the degree of information that is given by a certain meteorological element about future values of this element or others. Properties of persistency exist more or less among all meteorological variables. It is an expression for the inertia of the atmosphere and is especially strong in certain latitudes for such variables as temperature and pressure. A knowledge of the peristency of weather has even a certain prognostic value. So-called persistency forecasts are a kind of blind forecasts that ought to be studied when verifying the skill of weather forecasts.

It is also important to attempt the construction of mathematical models giving the variation of meteorological elements which agree as far as possible with empirical data. In this study the so-called Markov chain model has been used. This model implies that the value of a certain variable x at time t is dependent only on the value of the same variable at time $t-1$. The knowledge of previous x-values thus give no additional information in this model. We can express the value of x at time t called x_t as a linear function of x_{t-1}.

$$x_t = ax_{t-1} + \varepsilon_t, \qquad (2)$$

where ε_t is the residual. The autocorrelation coefficient $\varrho\nu$ is quite easy to calculate as it is given by the expression:

$$\varrho\nu = a^{\nu}. \qquad (3)$$

The meteorological element, which will now be studied is the qualitative quantity "precipitation day–dry day". A precipitation day or a wet day has been defined as a day with precipitation amount ≥ 0.1 mm.

The time series that will be studied is thus of quite a simple type. The variable can only have two states, here denoted by 0 respectively 1. Zero means that precipitation has not occurred and the figure one that it has. The probability for state 1 is simply estimated with the aid of the relative frequency of wet days, that is to say the number of wet days divided by the total number of days in the period investigated. This probability is denoted by $p(1)$ and thus $p(0) = 1 - p(1)$. These estimates are maximum-likelihood estimates of the population values $\pi(1)$ and $\pi(0)$. Conditional probabilities are indicated by indices according to the following example. $p_{00}(1)$ shall be interpreted as the conditional probability of a precipitation day when the two preceding days had none. Similarly $p_{11}(0)$ means the probability for state 0 when state 1 occurred the two previous days. If the Markov model is valid the following conditions are fulfilled.

$$\left.\begin{aligned} p_1(1) &= p_{11}(1) = p_{111}(1) = \ldots \text{ etc.,} \\ p_0(0) &= p_{00}(0) = p_{000}(0) = \ldots \text{ etc.} \end{aligned}\right\} \quad (4)$$

These expressions mean that all information about the state on day d is given by the state on day $d-1$. If this could be verified by empirical data concerning the meteorological attribute "precipitation day–dry day", it is possible to derive and apply several interesting formulas.

2. Purpose of the investigation

The investigation was primarily started in order to gain an appreciation of the character of precipitation persistency. For instance how many days ahead the persistency information reaches, whether there exists any seasonal variation of the persistency of precipitation, and if it is realistic to apply the Markov chain model mentioned in the introduction. In order to simplify the investigation no consideration was taken to precipitation amounts.

3. The material

The primary data material for the study was taken from extenso tables published by the Swedish Meteorological and Hydrological Institute in "Meteorological observations in Sweden". As investigation period the so-called normal period 1931–60 was chosen. The calculations were done most detailed for Stockholm. After a number of results had been obtained for this locality some calculations were also carried out for a few other Swedish stations.

4. The extent of the first calculations

The conditions during one month were considered as stationary. For this reason, data from one and the same month were taken together to one sample. In order to gain an idea of the random variations of the calculated parameters, each of the three decades of the investigation period was treated separately. For the precipitation data from Stockholm the following computations were carried out.

1. The distribution of the number of precipitation days.
2. The means and standard deviations of these distributions.
3. The mean errors of the arithmetic mean values.
4. Mean precipitation amount and average precipitation per precipitation day.
5. The probability $p(1)$ and its mean error.
6. Transitional probabilities with up to six conditions.
7. Coefficients of persistence according to the formulas:

$$r(1) = 1 - \left[\frac{1-p_1(1)}{1-p(1)}\right]^2$$

respectively

$$r(0) = 1 - \left[\frac{1-p_0(0)}{1-p(0)}\right]^2, \quad (5)$$

r will receive the value 0, if $p(1)$ and $p_1(1)$ are the same and the value 1, if $p_1(1) = 1$, that is to say, the value today gives complete information of the state tomorrow.

For other stations the calculations made were less complete. Only a certain part of the parameters mentioned above were computed.

5. Results and inferences from the first computations

In Table 1 some statistics for Stockholm for the period 1931–60 for the different months of the year are given.

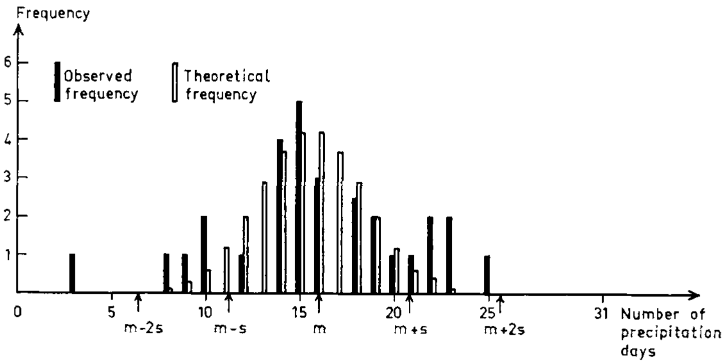

FIG. 1. The distribution of the number of precipitation days in Stockholm during January. For comparison a binomial distribution with π (1) = π (0) = $\frac{1}{2}$ has been drawn beside the observed frequencies.

The distribution of the number of precipitation days per month is very flat as can be seen from the high values of the standard deviation. Only one distribution is shown, the one for January, which is given in Table 2 and Fig. 1. From the figure can be drawn the conclusion that 30 years is too short a period to get the shape of the distribution with any confidence. The random variations are very great. For example we may note that there has been no year during the investigated period when a January month had 17 precipitation days, though this number is quite close to the arithmetical mean.

To test the hypothesis that the distribution of precipitation days is a binomial one, the equivalent theoretical frequencies were calculated on the assumption that the universe probabilities $\pi(1) = \pi(0) = \frac{1}{2}$. These frequencies are given both in Table 2 and in Fig. 1. A χ^2-test showed that the hypothesis could not be accepted. This is not an unexpected result as the binomial distribution can be valid only if the probability of precipitation tomorrow is independent of the condition today. The dispersion of the observed distribution is much greater than that of binomial distributions with the same value of $p(1)$ as observed. The standard deviation of a binomial is about 2.6 ($\sigma = \sqrt{30\,p(1)\,p(0)}$. The high value for χ^2 is mainly due to the high deviations in the tails of the distribution. The deviations are explained by the existence of persistency in precipitation.

The mean number of precipitation days varies with season. The lowest value coincides with precipitation minimum during March. The highest values occur during November, December and January. The mean precipitation amount per precipitation day, which is given in Table 1, has a yearly variation, which resembles the annual trend of temperature, with minimum in February, but with maximum some-

TABLE 1. *The mean number of precipitation days* (m), *the standard deviation of the distribution of precipitation days* (s), *the mean error of the means* (s_m), *mean precipitation amount in mm* (N), *average precipitation per precipitation day in mm* (g), *the probability of a precipitation day* $p(1)$ *in per cent and the mean error of this probability* s_p.

Months . . .	J	F	M	A	M	J	J	A	S	O	N	D
Parameters												
m	16.0	13.5	10.1	11.2	10.8	12.7	12.9	13.6	14.0	14.8	16.5	16.8
s	4.8	3.8	4.0	3.8	3.6	3.4	4.3	4.5	4.2	5.4	3.9	4.8
s_m	0.9	0.7	0.7	0.7	0.7	0.6	0.8	0.8	0.8	1.0	0.7	0.9
N	43	30	26	31	34	45	61	76	60	48	53	48
g	2.7	2.2	2.6	2.8	3.2	3.5	4.7	5.6	4.3	3.3	3.2	2.9
$p(1)$	51.5	47.9	32.7	37.4	34.8	42.2	41.5	43.8	46.6	47.7	54.9	45.3
s_p	2.9	2.3	2.4	2.3	2.2	2.1	2.6	2.7	2.6	3.3	2.4	2.9

TABLE 2. *The distribution of the number of precipitation days during January 1931–60 in Stockholm, observed and theoretical.*

No. of wet days	Observed frequency	Theoretical frequency
3	1	0
4	0	0
5	0	0.01
6	0	0.01
7	0	0.04
8	1	0.11
9	1	0.28
10	2	0.62
11	0	1.18
12	1	1.97
13	0	2.88
14	4	3.70
15	5	4.20
16	3	4.20
17	0	3.70
18	3	2.88
19	2	1.97
20	1	1.18
21	1	0.62
22	2	0.28
23	2	0.11
24	0	0.04
25	1	0.01
26	0	0.01
	30	30.00

$\chi^2 = 64$ with 4 degrees of freedom.
$P < 0.001$.

what later for precipitation intensity than for temperature. From the mean errors of the mean number of precipitation days is seen that a 30-year period may be long enough to give fairly good estimates of the true population means. It can be assumed with a confidence of 95% that the true means are situated within an interval of somewhat less than 4 days, with 2 days on either side of the sample means m. The seasonal variations of the probability for a precipitation day coincides with the variation of the means m. The highest probability for a wet day in Stockholm occurs during November and December. The lowest values of $p(1)$ is found in March. The confidence of the estimate of $p(1)$ is given in Table 1 by s_p. The mean error reaches 2 to 3% and the decimals that are given have no significance.

A great number of conditional probabilities were calculated for a precipitation respectively a dry day when preceded by certain states. The following probabilities were computed: $p_1(1)$, $p_{11}(1)$, ... $p_{111111}(1)$ and $p_0(0)$, $p_{00}(0)$, ... $p_{000000}(0)$. The uncertainty of these estimates increases with a rise in the number of conditions as the

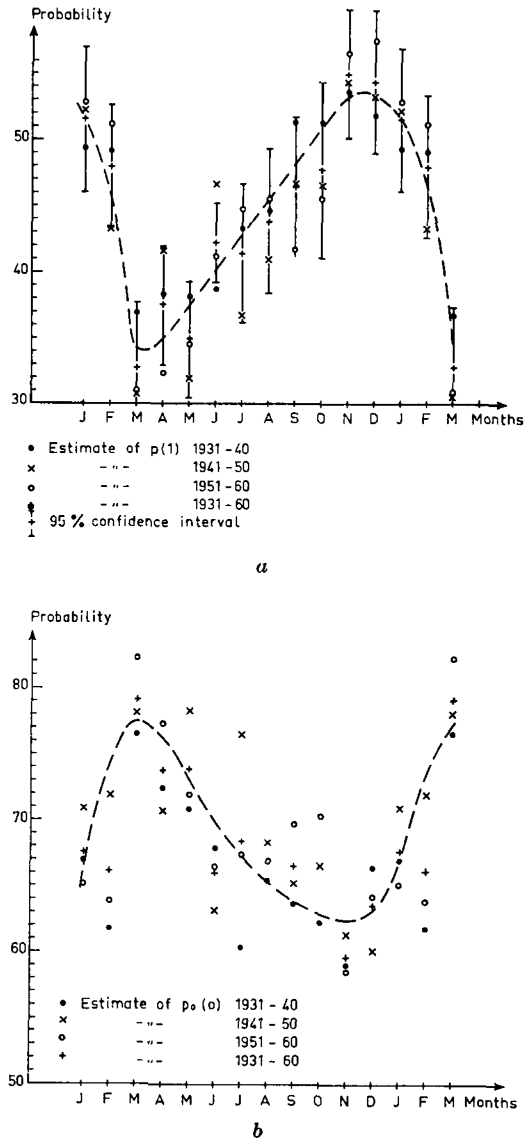

FIG. 2. (*a*) Estimates of the probability of a wet day in Stockholm during the different months of the year. The calculations are based on different decades. (*b*) Estimates of the probability of a dry day when last day was dry. The probabilities are estimated from different periods and are valid for Stockholm.

TABLE 3. *Unconditional and conditional probabilities for a dry or a wet day.*

Months ...	J	F	M	A	M	J	J	A	S	O	N	D	Year
Prob.													
$p(1)$	52	48	33	37	35	42	42	44	47	48	55	54	45
$p_1(1)$	69	61	58	55	52	53	56	59	61	64	67	70	61
$p_{11}(1)$	68	63	56	55	54	51	55	59	61	64	68	73	62
$p_{111}(1)$	66	62	50	58	52	53	54	62	61	67	70	75	63
$p_{1111}(1)$	64	61	55	59	51	56	55	54	63	69	73	74	64
$p_{11111}(1)$	67	63	63	63	54	58	51	60	68	66	76	73	67
$p_{111111}(1)$	67	65	47	64	46	61	39	71	73	67	76	73	68
$p(0)$	48	52	67	63	65	58	58	56	53	52	45	46	55
$p_0(0)$	68	66	79	74	74	66	68	67	67	67	60	64	69
$p_{00}(0)$	70	68	80	77	76	68	73	72	73	74	61	67	73
$p_{000}(0)$	70	70	83	75	78	70	72	74	76	74	66	69	74
$p_{0000}(0)$	71	75	84	76	77	70	73	76	74	76	70	69	76
$p_{00000}(0)$	73	77	84	78	78	69	76	75	76	78	71	67	77
$p_{000000}(0)$	74	82	84	81	77	71	77	76	80	77	75	69	78

number of cases will be thereby reduced. The number of cases with two successive precipitation days was, for example, in January for the whole 30-year period 329, while the number of occasions with seven consecutive precipitation days was 42.

In Fig. 2*a* estimates of $p(1)$ have been plotted and in Fig. 2*b* of $p_0(0)$ for different decades. The great dispersion between decades is clearly seen. The double mean error $2s_p$ has been marked on both sides of $p(1)$ calculated from the 30-year period. It is seen that about 5% of the estimates from 10-year periods fall outside the confidence intervals, which agrees well with sampling theory. An attempt has been made to draw a smooth curve for the yearly variation of $p(1)$ and $p_0(0)$.

Table 3 gives estimates of the different conditional probabilities together with $p(1)$ and $p(0)$. The most striking feature is the great increase from $p(1)$ to $p_1(1)$ respectively from $p(0)$ to $p_0(0)$. After that the increase is rather modest. Due to random variations certain months can show a lower value for the conditional probability of higher degree than for one of lower degree. This is certainly not in accordance with the true state. Looking upon the results for the year as a whole, it can be seen that the probability of precipitation increases from 45% to 68% when precipitation has occurred during the past 6 days. The probability of a dry day increases with the same amount, 23%, from 55% to 78% probability for a dry day tomorrow, when preceding 6 days have been dry. The conditions vary of course during different seasons. The periods with stationary anticyclones during the month of March are reflected in the high figure 84% probability for a dry day, when there has been a sequence of 6 dry days.

The calculation of the transition probability figures shown in Table 3 reveals that the simple Markov model cannot give a perfect description of observed conditions. A small amount of additional information is gained if also the state on day $d-2$, $d-3$, ... $d-N$ is considered. Investigations by BJÖRGUM & FJORDHOLM (1946) have shown an increase in the conditional probability at least up to $N = 14$ days.

An index giving a measure of the strength of the persistency was calculated according to formula 5. The result of these calculations is shown in Table 4. As seen, this index can give different values for precipitation days and for dry days, depending upon sampling effects. This can be shown in the following way. The probabilities $p(01)$ and $p(10)$, which are the probabilities for both a dry and a wet day in sequence, can be written:

$$p(10) = p(01) = p(1)\, p_1(0) = p(0)\, p_0(1) \quad (5a)$$

from which is derived:

$$\frac{p(1)}{p(0)} = \frac{p_0(1)}{p_1(0)}. \quad (5b)$$

TABLE 4. *Indices of persistency.*

Months ...	J	F	M	A	M	J	J	A	S	O	N	D	Year
Index													
$r(1)$	0.58	0.45	0.61	0.48	0.45	0.35	0.43	0.46	0.33	0.52	0.47	0.57	0.51
$r(0)$	0.60	0.50	0.59	0.51	0.48	0.35	0.42	0.43	0.48	0.51	0.46	0.55	0.51

Using this relationship in formula 5 gives:

$$r(1) \equiv r(0). \tag{5c}$$

The highest value for the persistency index is found in March and the lowest values during the period June–September. To what extent the difference between months is significant has not been proved.

6. Calculations with the use of the simple Markov chain model

The following calculations will try to show how realistic it is to assume that the time series of dry and wet days behave like a simple Markov chain.

As all information given by the time series about future conditions with regard to the likelihood of precipitation is assumed to pass through the last value, several formulas can be derived, giving different probabilities for precipitation respectively dry weather. For example:

1. The probability of dry or wet weather a certain day, n days after a given state 0 or 1.
2. The probability of precipitation one or more days during an interval of N days.
3. The probability for spells with dry or wet weather of different length.
4. The probability of a certain number of precipitation days occurring during a certain period.

6.1. The probability of precipitation respectively dry weather the day n, when the state at day 0 is known

Let us denote the probability for precipitation on day n with Q_n, if precipitation was observed day 0, with q_n if no precipitation occurred on day 0. The following recurrence formulas can be written:

$$\left.\begin{aligned}
&Q_n = Q_{n-1}p_1(1) + (1 - Q_{n-1})p_0(1),\\
\text{where}\quad &Q_1 = p_1(1) \quad \text{and}\\
&q_n = q_{n-1}p_1(1) + (1 - q_{n-1})p_0(1),\\
\text{where}\quad &q_1 = p_0(1),\\
&\lim_{n\to\infty} Q_n = \lim_{n\to\infty} q_n = p(1).
\end{aligned}\right\} \tag{6}$$

Similarly, formulas can be written giving the probability for dry weather on day n, P_n, if no precipitation occurred on day 0 and denoted by p_n the same probability if day 0 was wet.

$$\left.\begin{aligned}
&P_n = P_{n-1}p_0(0) + (1 - P_{n-1})p_1(0),\\
\text{where}\quad &P_1 = p_0(0) \quad \text{and}\\
&p_n = p_{n-1}p_0(0) + (1 - p_{n-1})p_1(0),\\
\text{where}\quad &p_1 = p_1(0),\\
&\lim_{n\to\infty} P_n = \lim_{n\to\infty} p_n = p(0).
\end{aligned}\right\} \tag{7}$$

The probability numbers, given in Table 5, very soon approach the unconditional climatological probabilities $p(1)$ and $p(0)$. This gives a certain control that the formulas 6 and 7 on the whole give a fairly correct picture of observed conditions. Some small deviations appear for some months between the limit values of Q_n, q_n, P_n, and p_n and the estimated values of $p(1)$ and $p(0)$. But these deviations do not exceed 1%. After 4–6 days the persistency effect has ebbed out. But already after 2 to 3 days the difference between P_n and p_n and between Q_n and q_n is so small that it is of no practical value. This shows that precipitation has a much lower degree of persistency than temperature or pressure. With regard to temperature the persistence information extends much further ahead. Not before n is about 30 days the autocorrelation coefficient is zero.

TABLE 5. *The probabilities Q_n, q_n, P_n and p_n together with $p(1)$ and $p(0)$ expressed in per cent.*

		n							
Month	Prob.	0	1	2	3	4	5	6	$p(0)/p(1)$
January	Q_n	100	69	57	53	52	51		51
	q_n	0	32	44	48	50	51		
	P_n	100	68	56	52	50	50	49	49
	p_n	0	31	43	47	48	49	49	
February	Q_n	100	61	51	48	47	47		48
	q_n	0	34	43	46	46	47		
	P_n	100	66	57	54	54	53		52
	p_n	0	39	49	52	53	53		
March	Q_n	100	58	42	37	34	34	33	33
	q_n	0	21	29	32	33	33	33	
	P_n	100	79	71	69	67	67	67	67
	p_n	0	42	58	63	66	66	67	
April	Q_n	100	55	42	38	37			37
	q_n	0	26	34	36	37			
	P_n	100	74	66	64	63			63
	p_n	0	45	58	62	63			
May	Q_n	100	52	39	36	35			35
	q_n	0	26	33	35	35			
	P_n	100	74	67	66	65			65
	p_n	0	49	61	64	65			
June	Q_n	100	53	44	43	42			42
	q_n	0	34	41	42	42			
	P_n	100	66	59	58	58			58
	p_n	0	47	56	57	58			
July	Q_n	100	56	45	43	42			42
	q_n	0	32	39	41	42			
	P_n	100	68	61	59	58			58
	p_n	0	44	55	57	58			
August	Q_n	100	59	48	45	45	44		44
	q_n	0	33	42	44	44	44		
	P_n	100	67	59	56	56	56		56
	p_n	0	42	52	55	55	56		
September	Q_n	100	61	50	47	46			47
	q_n	0	34	43	45	46			
	P_n	100	67	57	55	54			53
	p_n	0	39	50	53	54			
October	Q_n	100	64	53	50	49	48		48
	q_n	0	33	44	47	47	48		
	P_n	100	67	57	53	52			52
	p_n	0	36	47	51	52			
November	Q_n	100	67	58	56	55			55
	q_n	0	40	51	54	55			
	P_n	100	60	49	46	45			45
	p_n	0	33	42	44	45			
December	Q_n	100	70	60	57	56	55		54
	q_n	0	37	49	53	54	55		
	P_n	100	64	51	47	46	45		46
	p_n	0	30	40	43	44	45		

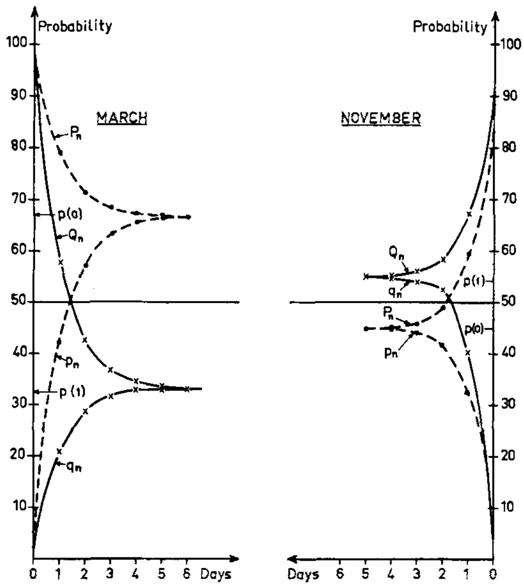

FIG. 3. The probability of a wet (dry) day in Stockholm during March and November a certain day in the future, when the last known day is wet (dry). Full drawn lines indicate wet weather, stretched lines dry weather.

The persistency of precipitation, according to Table 5, shows, in accordance with the values in Table 4, a tendency to higher values during the winter season than during summer.

Figure 3 shows how the precipitation persistency loses its importance with increasing time interval from the last known state.

6.2. PROBABILITY OF PRECIPITATION DURING PERIODS OF DIFFERENT LENGTH

The probability to get either 1, 2, 3,... or N days with precipitation occurring during a period of N days can be expressed by the following formula if the assumption of Markov series is valid:

$$R_N = 1 - p(0)p_0(0)^{N-1}. \quad (8)$$

The proof of this expression is given by CASKEY (1963). R_N approaches certainty when the length of the period increases. In Table 6 these probabilities are given for different months of the year. To determine the size of the differences obtained due to the assumption made, the really observed probabilities of precipitation for periods up to 7 days were calculated. These relative frequencies have been denoted by r_N in Table 6. The observed probabilities are systematically less than the theoretical except for $N=2$. The explanation for this is not easy to give. But the differences are not very great. Therefore it can be acceptable to compute the probability for precipitation during a period in the simple manner using formula 8. Fig. 4 illustrates for a number of months how the probability for precipitation grows exponentially with increasing length of period.

6.3. THE PROBABILITY OF SPELLS OF WET OR DRY DAYS WITH DIFFERENT DURATION

A spell of dry weather can have the length 1, 2, 3,... m days. The total number of periods is assumed to be N, and the number having length m is called N_m. We have thus $N = \sum_{i=1}^{m} N_i$ and the empirical probability of a dry spell of length m is defined:

$$u_m = \frac{N_m}{N}.$$

For Markov series can be shown that the probability for a dry spell of length m can be computed from the following formula:

$$U_m = p_0(1)p_0(0)^{m-1}. \quad (9)$$

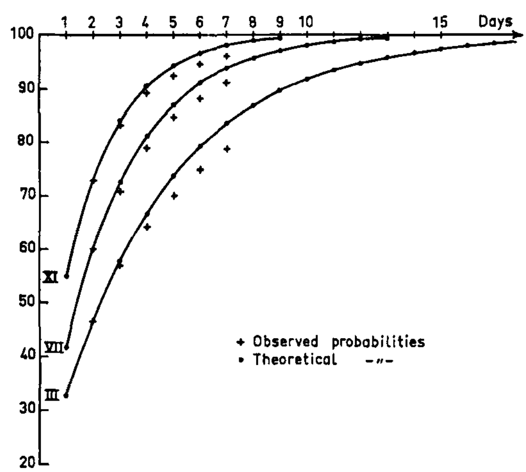

FIG. 4. The probability to get precipitation during intervals of different length. Values for Stockholm and for the months of March, July and November.

TABLE 6. *Theoretical (R) and empirical (r) probabilities for precipitation during different time intervals.*

N	R/r	J	F	M	A	M	J	J	A	S	O	N	D
2	R	67	66	47	54	52	62	60	62	65	65	73	71
	r	67	66	47	54	52	62	60	62	65	65	73	71
3	R	78	77	58	66	65	75	73	75	76	77	84	82
	r	77	77	57	65	63	74	71	73	74	74	74	80
4	R	85	85	67	75	74	83	81	83	84	85	90	88
	r	84	84	64	73	72	82	79	80	80	81	81	87
5	R	90	90	74	82	81	89	87	89	90	90	94	93
	r	89	88	70	80	78	87	85	85	85	86	86	91
6	R	93	94	79	86	86	93	91	92	93	93	97	95
	r	92	90	75	84	83	91	88	89	89	89	89	94
7	R	95	96	84	90	90	95	94	95	95	95	98	97
	r	94	92	79	87	87	94	91	91	91	91	91	96
8	R	97	97	87	93	92	97	96	97	97	97	99	98
9	R	98	98	90	95	94	98	97	98	98	98	99.3	99
10	R	99.1	99	92	96	96	99	98	99	99.7	99	99.6	99.2
11	R	99.1	99.1	94	97	97	99.1	99	99.0	99.1	99.1	99.8	99.5
12	R	99.4	99.4	95	98	98	99.4	99.1	99.3	99.6	99.4	99.9	99.7
13	R	99.6	99.6	96	98	98	99.6	99.4	99.5	99.6	99.6		99.8
14	R	99.7	99.7	97	99	99	99.7	99.6	99.7	99.7	99.7		99.9
16	R	99.9	99.9	98	99.3	99.3	99.9	99.8	99.9	99.9	99.9		
18	R			99	99.6	99.6							
20	R			99.2	99.8	99.8							

The probability of a wet spell of length k denoted by N_k can be shown to be:

$$N_k = p_1(0)p_1(1)^{k-1}. \tag{10}$$

For the proof see FELLER (1957). The abovementioned probabilities are given in Table 7a and b together with the observed, denoted by small letters. The decimals given have no significance.

A study of Table 7 may induce some comments. The greatest probability for long dry spells in Stockholm occurs during March followed by April and May. However, the longest period with consecutive dry days was found during an autumn for the 30-year period investigated. In 1951 no precipitation was measured from 20/9 to 30/10, wich makes 41 days. One 30-day period with dry weather occurred during August–September and a spell of 27 days duration was found in March–April. During the 30 years examined, no dry spells longer than 13 days occurred in November. Concerning long and short wet spells can be noted that during March and July the longest observed periods are 9 days. This is in accordance with the theory that the probability of long wet spells has its lowest values from March until August. The longest wet spell has been observed during November and December with 18 wet days in sequence. This is in good agreement with the model tested, which gives highest probability for long wet spells from November to January.

The agreement between observed and theoretical frequencies is in general worse for dry spells than for wet spells. A χ^2-test giving figures of how good the agreement is between observed and theoretically calculated number of cases was performed. For January the χ^2-value was 2.9 for dry spells and 6.4 for wet spells. With 7 degrees of freedom these values are too low to give rise to doubts that the differences between theoretical and observed frequencies are due to chance. Formula 9 gives the worst agreement for $m = 1$, where the empirical probabilities are higher.

From the tables it is clear that 30 years is too short a period to give good estimates of the probability for different spells of wet or dry days.

Figures 5 and 6 show the theoretical frequencies of dry and wet spells for a number of months.

TABLE 7*a*. *The probability* (%) *of dry weather during different number* (*m*) *of days iu sequence, theoretical* (*U*) *and observed* (*u*).

m	*U*/*u*	J	F	M	A	M	J	J	A	S	O	N	D
1	*U*	32.4	33.9	20.9	26.3	26.2	34.0	31.6	33.1	33.5	33.3	40.4	36.5
	u	37.7	38.0	26.0	35.1	32.7	38.4	41.3	43.9	46.6	46.9	42.7	43.2
2	*U*	21.9	22.4	16.5	19.4	19.3	22.4	21.6	22.1	22.3	22.2	24.1	23.2
	u	18.5	22.7	23.7	10.8	19.5	21.5	14.5	17.3	18.0	14.2	26.2	20.0
3	*U*	14.8	14.8	13.1	14.3	14.3	14.8	14.8	14.8	14.8	14.8	14.4	14.7
	u	14.4	16.0	10.7	14.9	10.1	12.4	13.4	13.3	6.8	13.0	13.4	11.6
4	*U*	10.0	9.8	10.3	10.5	10.5	9.8	10.1	9.9	9.9	9.9	8.6	9.3
	u	9.6	7.3	5.3	12.2	8.8	7.9	11.1	4.6	8.7	8.0	5.5	7.1
5	*U*	6.8	6.5	8.2	7.8	7.8	6.5	6.9	6.6	6.6	6.6	5.1	5.9
	u	6.2	6.0	6.1	8.8	6.3	6.8	5.2	6.4	7.5	3.1	4.9	6.5
6	*U*	4.6	4.3	6.5	5.7	5.7	4.3	4.7	4.4	4.4	4.4	3.0	3.8
	u	3.4	2.7	2.3	6.1	5.0	3.4	2.3	5.2	4.4	3.7	1.8	5.2
7	*U*	3.1	2.8	5.1	4.2	4.2	2.8	3.2	3.0	2.9	2.9	1.8	2.4
	u	4.1	2.0	4.6	2.0	4.4	3.4	1.2	2.9	1.2	3.1	0.6	1.3
8	*U*	2.1	1.9	4.0	3.1	3.1	1.9	2.2	2.0	1.9	2.0	1.1	1.5
	u	1.4	0.7	4.6	0.7	1.9	2.3	2.9	1.2	1.2	3.7	1.2	0.7
9	*U*	1.4	1.2	3.2	2.3	2.3	1.2	1.5	1.3	1.3	1.3	0.6	1.0
	u	0	0.7	1.5	0.7	3.1	0	2.9	1.7	1.9	0.6	0.6	1.3
10	*U*	1.0	0.8	2.5	1.7	1.7	0.8	1.0	0.9	0.9	0.9	0.4	0.6
	u	1.4	2.0	2.2	0.7	2.5	0	1.2	0.6	1.9	0.6	0.6	0
11	*U*	0.6	0.5	2.0	1.2	1.3	0.5	0.7	0.6	0.6	0.6	0.2	0.4
	u	0.7	0	1.5	0.7	1.9	2.8	1.2	0.6	0	0.6	1.2	1.3
12	*U*	0.4	0.4	1.6	0.9	0.9	0.4	0.5	0.4	0.4	0.4	0.1	0.2
	u	0.7	0	2.3	2.7	0	0.6	0.6	0	0	0.6	0	1.3
13	*U*	0.3	0.2	1.3	0.7	0.7	0.2	0.3	0.3	0.3	0.3	0.1	0.2
	u	0	0	1.5	1.4	0.6	0	1.2	0.6	0.6	0	1.2	0
14	*U*	0.2	0.1	1.0	0.5	0.5	0.2	0.2	0.2	0.2	0.2	0.1	0.1
	u	1.4	0	0.8	1.4	0.6	0	0.6	0	0	0	0	0.7
15	*U*	0.1	0.1	0.8	0.4	0.4	0.1	0.2	0.1	0.1	0.1	0	0.1
	u	0	0	2.3	0	0	0	0	0	0	0.6	0	0
16	*U*	0.1	0	0.6	0.3	0.3	0.1	0.1	0.1	0.1	0.1	0	0
	u	0	0	0.8	0.7	1.9	0	0.6	0.6	0.6	0	0	0
>16	*U*	0.1	0.2	2.2	0.6	0.6	0	0.1	0.1	0	0.1	0	0
	u	0.7	2.1	3.8	1.3	0.6	0.6	0	1.2	0.6	1.2	0	0

The probability for a certain number of wet or dry days occurring during a certain period has not been computed. Formulas for such probabilities have been given by GABRIEL (1959).

7. Some results from other stations

From the computations of data for Stockholm for the period 1931–60 it became fairly evident that there is no particular reason for calculating for other stations other conditional probabilities than $p_1(1)$ and $p_0(0)$.

The stations for which certain computations have been carried out are as follows: (1) Lund, (2) Gothenburg, southerly stations with a maritime climate, (3) Riksgränsen, a northerly one with very strong maritime influence, (4) Haparanda, another more continental station in northern Sweden and (5) Knon, a station in central Sweden with a very continental climate. In Fig. 7 a small map shows the position of these stations.

Unfortunately the computations covered only a 10-year period 1951–60, with the exception of Lund, where the normal period 1931–60 has

TABLE 7b. *The probability (%), theoretical (N) and empirical (n) for wet spells of different length (k).*

k	N/n	J	F	M	A	M	J	J	A	S	O	N	D
1	N	31.3	38.7	42.1	45.0	48.5	46.6	44.0	41.5	39.4	36.0	32.8	29.9
	n	29.3	40.8	39.1	44.7	51.0	43.5	42.9	42.0	40.6	35.6	34.6	37.7
2	N	21.5	23.7	24.4	24.8	25.0	24.9	24.6	24.3	23.9	23.0	22.0	21.0
	n	19.3	21.0	22.7	27.6	21.7	29.4	24.7	26.6	22.4	27.5	23.5	19.9
3	N	14.8	14.5	14.1	13.6	12.9	13.3	13.8	14.2	14.5	14.7	14.8	14.7
	n	16.7	14.7	21.1	11.8	12.7	13.6	15.3	7.7	15.8	13.1	15.4	8.6
4	N	10.1	8.9	8.2	7.5	6.6	7.1	7.7	8.3	8.8	9.4	10.0	10.3
	n	14.0	9.6	9.4	7.2	7.6	6.2	7.1	12.4	9.7	5.6	9.3	8.0
5	N	7.0	5.5	4.7	4.1	3.4	3.8	4.3	4.9	5.3	6.0	6.7	7.2
	n	6.7	5.7	1.6	3.3	2.6	3.4	3.5	6.5	4.9	6.9	4.3	6.6
6	N	4.8	3.3	2.7	2.3	1.8	2.0	2.4	2.8	3.2	3.9	4.5	5.1
	n	2.7	3.2	3.1	1.3	3.2	1.1	4.1	1.8	0.6	3.8	3.1	5.3
7	N	3.3	2.1	1.6	1.2	0.9	1.1	1.4	1.7	2.0	2.5	3.0	3.5
	n	5.3	1.9	2.3	2.6	0	1.1	1.2	0	2.4	1.2	2.5	3.3
8	N	2.3	1.3	0.9	0.7	0.5	0.6	0.8	1.0	1.2	1.6	2.0	2.5
	n	1.3	0.6	0.8	0	0.6	1.1	0.6	0.6	0.6	2.5	1.9	4.0
9	N	1.6	0.8	0.5	0.4	0.2	0.3	0.4	0.6	0.7	1.0	1.4	1.7
	n	2.7	0.6	0.8	0	0	0	0.6	0.6	1.2	1.9	1.2	1.3
10	N	1.1	0.5	0.3	0.2	0.1	0.2	0.2	0.3	0.4	0.6	0.9	1.2
	n	0	0.6	0	0.7	0.6	0	0	0.6	0	1.3	0.6	1.3
11	N	0.7	0.3	0.2	0.1	0.1	0.1	0.1	0.2	0.3	0.4	0.6	0.9
	n	0.7	0.6	0	0	0	0.6	0	0.6	1.2	0	0.6	0.7
12	N	0.5	0.2	0.1	0.1	0	0	0.1	0.1	0.2	0.3	0.4	0.6
	n	0.7	0	0	0.7	0	0	0	0.6	0	0	1.2	0.7
13	N	0.3	0.1	0.1	0	0	0	0	0.1	0.1	0.2	0.3	0.4
	n	0.7	0.6	0	0	0	0	0	0	0	0	0	0.7
14	N	0.2	0.1	0	0	0	0	0	0	0.1	0.1	0.2	0.3
	n	0	0	0	0	0	0	0	0	0.6	0	0.6	1.3
15	N	0.2	0	0	0	0	0	0	0	0	0.1	0.1	0.2
	n	0	0	0	0	0	0	0	0	0	0	0.6	0
16	N	0.1	0	0	0	0	0	0	0	0	0	0.1	0.1
	n	0	0	0	0	0	0	0	0	0	0.6	0	0
>16	N	0.2	0	0	0	0	0	0	0	0	0	0.1	0.2
	n	0	0	0	0	0	0	0	0	0	0	0.6	0

been investigated. The means and mean frequencies calculated from samples consisting of data from 10 years, have mean errors that are $\sqrt{3}$ times as great as those calculated from a period of 30 years.

For the above-mentioned stations the following parameters are given in Table 9:

1. Mean number of precipitation days per month (m).
2. Standard deviation of the number of precipitation days (s).
3. Mean error of the mean (s_m).
4. The probability $p(1)$ and its mean error (s_p).
5. The conditional probabilities $p_1(1)$ and $p_0(0)$.
6. Persistency indices $r(1)$ and $r(0)$.

Some characteristics concerning the persistency of precipitation at different stations can be observed from a study of Table 9. The most striking is that stations with maritime influence have the highest values of the persistency index. Among the stations investigated, Riksgränsen, besides having strong maritime influence, has such a location that orographical influences make the persistency extraordinarily high. The mean for the whole year is about 0.7. The persistency of precipitation is connected with the persistency of the pressure distribution or, in other words, with the persistency of weather types, which cause a persistency of the wind. To compare the persistency index for Riksgränsen with other places, it can be mentioned that the figure 0.4 has been stated for Paris, Bergen and Kew.

The calculations seem to verify that a yearly trend of persistency of precipitation does exist. The values from Gothenburg, Haparanda and Riksgränsen show, as with Stockholm, the lowest values during June. The period with the highest values varies between stations. Lund and Gothenburg have maximum values in March–April, Haparanda in May. Riksgränsen shows the same values from November to March. The computations for Knon from a 10-year period show equal persistency conditions during the whole year.

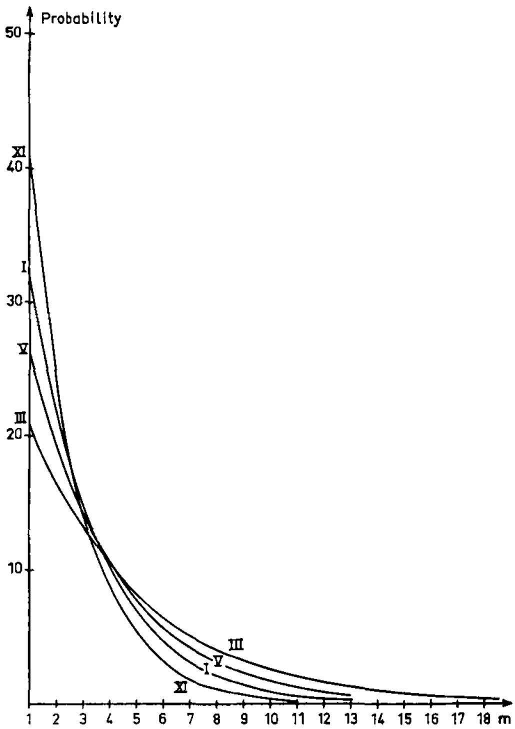

FIG. 5. The probability of dry spells of length m in Stockholm during the months of January, March, May and November.

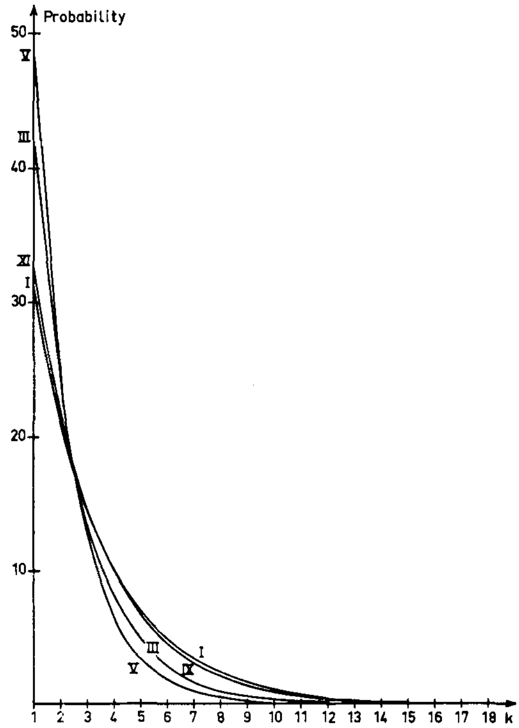

FIG. 6. The probability of wet spells of length k in Stockholm during the months of January, March, May and November.

8. Conclusions

The investigation has shown that a 30-year period can be too short to give good estimates of certain precipitation parameters. Though no consideration has been taken to precipitation amounts, it is evident that 30 years is insufficient to give distributions and certain relative frequencies with any high degree of accuracy. It would be possible to increase the accuracy of the estimates of certain parameters by taking together those months showing similar proper-

TABLE 9. *Mean number of precipitation days (m), standard deviation of the precipitation days (s), mean error of m (s_m), probability of a wet day $p(1)$, mean error of $p(1)$, the conditional probabilities $p_1(1)$ and $p_0(0)$ and persistency coefficients $r(1)$ and $r(0)$.*

Months . . .	J	F	M	A	M	J	J	A	S	O	N	D
Stations												
Lund												
m	15.3	12.1	9.7	11.1	10.8	11.6	14.6	13.4	14.2	15.0	14.9	15.6
s	4.9	3.9	4.0	3.3	3.5	4.0	4.2	4.8	4.4	5.6	4.3	3.8
s_m	0.9	0.7	0.7	0.6	0.6	0.7	0.8	0.9	0.8	1.0	0.8	0.7
$p(1)$	49.5	42.8	31.3	36.9	34.8	38.6	47.0	43.1	47.3	48.5	49.6	50.2
s_p	3.0	2.4	2.4	2.0	2.1	2.4	2.6	2.1	2.7	3.4	2.6	2.3
$p_1(1)$	63.7	57.9	57.9	59.0	52.5	57.9	61.5	58.6	66.0	67.6	61.2	60.8
$p_0(0)$	65.1	69.1	79.9	76.6	75.4	72.7	67.1	67.3	69.6	69.1	62.1	59.6
$r(1)$	0.48	0.46	0.62	0.58	0.47	0.53	0.47	0.47	0.58	0.60	0.41	0.38
$r(0)$	0.50	0.48	0.59	0.60	0.50	0.50	0.51	0.42	0.59	0.59	0.42	0.35
Gothenburg												
m	17	12	11	13	10	11	15	14	15	17	18	16
s	4	5	4	4	3	4	4	4	5	7	4	4
s_m	1.2	1.7	1.1	1.1	1.0	1.1	1.3	1.3	1.6	2.1	1.2	1.4
$p(1)$	55	42	37	44	33	36	49	45	50	54	60	52
s_p	1	6	4	4	3	4	4	4	5	7	4	5
$p_1(1)$	69	62	60	68	57	52	64	61	69	74	76	69
$p_0(0)$	64	73	77	77	77	73	66	68	69	67	65	68
$r(1)$	0.5	0.6	0.6	0.7	0.6	0.4	0.5	0.5	0.6	0.7	0.6	0.6
$r(0)$	0.6	0.6	0.6	0.7	0.5	0.4	0.5	0.5	0.6	0.6	0.7	0.6
Knon												
m	15	13	8	11	14	15	16	16	15	15	18	18
s	4	5	4	4	4	4	5	5	4	4	5	4
s_m	1.2	1.7	1.4	1.1	1.1	1.4	1.7	1.6	1.2	1.3	91.5	1.3
$p(1)$	47	47	25	37	44	50	53	51	49	48	60	58
s_p	4	6	5	4	4	5	6	5	4	4	5	4
$p_1(1)$	63	61	46	60	61	66	68	66	69	67	73	71
$p_0(0)$	63	67	81	77	69	66	65	64	71	70	59	61
$r(1)$	0.5	0.5	0.5	0.6	0.5	0.5	0.5	0.5	0.6	0.6	0.6	0.5
$r(0)$	0.5	0.5	0.4	0.6	0.5	0.5	0.6	0.5	0.6	0.7	0.5	0.6
Haparanda												
m	17	16	12	10	9	11	12	13	12	13	17	19
s	3	5	4	3	3	3	5	4	4	4	4	4
s_m	1.0	1.6	1.3	0.8	1.0	0.9	1.4	1.3	1.2	1.2	1.4	1.3
$p(1)$	56	56	38	32	29	38	38	41	38	40	55	61
s_p	3	6	4	3	3	3	5	4	4	4	5	4
$p_1(1)$	64	72	59	51	56	48	51	60	52	51	70	73
$p_0(0)$	57	61	77	77	81	69	70	70	73	67	62	56
$r(1)$	0.3	0.6	0.5	0.5	0.6	0.3	0.4	0.5	0.4	0.3	0.6	0.5
$r(0)$	0.4	0.5	0.6	0.5	0.6	0.3	0.4	0.5	0.5	0.3	0.5	0.5
Riksgränsen												
m	16	13	17	15	16	18	18	20	20	17	12	14
s	5	6	4	2	4	4	6	3	4	6	4	5
s_m	1.6	2.0	1.4	0.6	1.3	1.4	1.8	0.9	1.1	1.8	1.1	1.6
$p(1)$	52	46	53	48	52	59	57	63	67	56	39	44
s_p	5	7	5	2	4	5	6	3	4	6	4	5
$p_1(1)$	75	74	73	67	73	72	74	76	77	72	69	69
$p_0(0)$	72	76	72	69	69	61	67	57	56	66	79	77
$r(1)$	0.7	0.8	0.7	0.6	0.7	0.5	0.6	0.6	0.5	0.6	0.7	0.7
$r(0)$	0.7	0.7	0.7	0.6	0.6	0.5	0.7	0.5	0.6	0.6	0.7	0.7

ties, for instance regarding the relative frequencies $p(1)$, $p_1(1)$ and $p_0(0)$. A study of the tables presented may indicate that for instance the months of November, December, January and February could be combined in one group. The months of July to October are in certain respects rather alike, just as April and May. For most of the investigated stations March is an extreme month, but could possibly be combined with April and May. A graphical smoothing of the above-mentioned frequencies as done in Fig. 2*a* and *b* could have been justified in order to eliminate parts of the noise. The use of smoothed frequencies in formulas 6–10 ought to give more reliable estimates of those probabilities which have been given in Tables 5–7.

The persistency of precipitation is not of sufficient potency to be of use for long range purposes. For one day precipitation forecasts persistency prognosis give a hit per cent, which lies about 15% above that of the climatological probability. But for a two days outlook the difference has shrunk to around 5%. When testing precipitation forecasts a comparison should be made with persistency forecasts in order to judge the skill of the forecasts.

This investigation has shown that it is possible, with good approximation, to regard the time series consisting of dry and wet days as having that simple property that the state tomorrow is dependent on the state today and not on the value yesterday, the day before yesterday and so on. For such time series a number of simple formulas are valid, giving different kinds of probabilities. To be able to apply these formulas, it is sufficient to know the three relative frequencies $p(1)$, $p_1(1)$ and $p_0(0)$. These should be calculated from a long sequence of years.

Calculations of the type shown in this study ought to be made for a greater number of stations, which would increase our knowledge of the precipitation climate of Sweden. Information of the climatological probabilities of precipitation could be of as great interest for certain purposes as the probability distribution of precipitation amounts. Agriculture and forestry for example would certainly be interested in probability numbers of the type given in this study. In planning time and place for vacations, climatological information of the type given here would also be of great value. Climatological probability distributions for different weather elements at different places and times seems to be the only safe information meteorology today can offer regarding the weather during coming months, seasons and years.

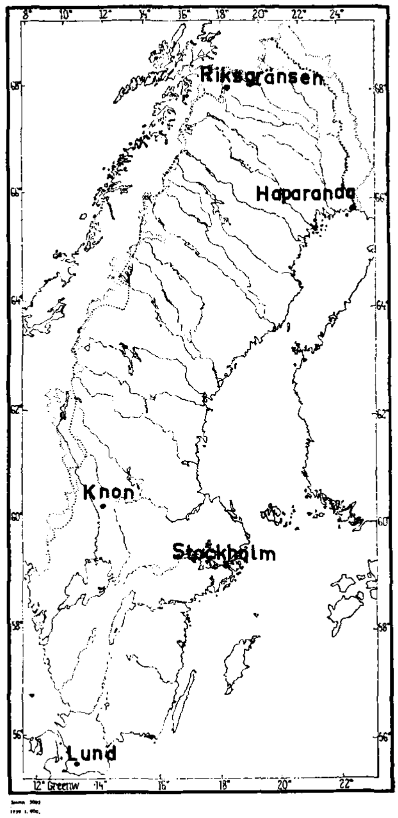

FIG. 7. Position of the stations mentioned in the text.

REFERENCES

BJÖRGUM, O., and FJORDHOLM, J. M., 1946, Statistiske nedbørsspådomar. *Naturen*, pp. 83–93.

CASKEY, J. E., 1963, A Markov chain model for the probability of precipitation occurrence in intervals of various length. *Monthly Weather Review*, pp. 298–301.

FELLER, W., 1957, *An Introduction to Probability Theory and its Applications*. London. John Wiley. pp. 338–373.

GABRIEL, K. R., 1959, The distribution of the number of successes in a sequence of dependant trials. *Biometrica* 1959, pp. 454–460.