Agri. Res. J. Kerala, 1977 15 (1) 47 — 55

PREDICTION OF WEEKLY RAINFALL OF A PLACE

P. U. SURENDRAN, K L SUNNY, and P. V. PRABHAKARAN

College of Veterinary and Animal Science, Mannuthy, Kerala

Knowledge of weekly rainfall of a place would undoubtedly be useful to the cultivators in their farm operations. As such, if the weekly rainfall could be forecast with reasonable confidence the cultivators could adjust their cultivation accordingly. While these predictions will be more helpful to rain-fed agriculture, irrigated agriculture could adjust the utilization of irrigation facility in consonance with the expected rainfall. Interest is therefore shown in predicting the weekly rainfall of a place. Two methods for predicting the weekly rainfall of a place have been given by Patro *et al.* (1975). They are based on California and Hazen's approaches for studying the flow of streams or for selecting flood frequency. It is easy to see that the two methods are identical except for the confidence percentages they attach to the minimum amounts of rainfall. In essence the general approach can be summarised thus. Let there be annual rainfall data of a place for Y years and assume that the arrangement of the data in descending order of magnitude is

(1) $X_1, X_2, \ldots\ldots X_y$

If now C is the percentage confidence in getting a minimum rainfall X_i, i = 1, 2, ..., Y, it is evident that

(2) $C_1 < C_2 < \ldots\ldots < C_y$

The points (3) $(C_1, X_1), (C_2, X_2), \ldots\ldots, (C_y, X_y)$

are then plotted in a semilogarithmic graph paper by measuring C_i along the horizontal axis and *X.* along the vertical axis. The best fitting line to these points drawn to the eye is then used to estimate the minimum rainfall for any chosen confidence percentage. The same procedure applied to the rainfall data of a week for Y years could be used to estimate the rainfall during that week for any chosen confidence percentage. In all cases Patro *et al.* (1975) found the best fitting lines of the points of the graph as straight lines. In California approach Cn has been taken as F_n and in Hazen's approach it has been assigned the value F_n' where

(4) F_n

(5) $F_n' = \frac{100\,(2n}{}$

Further, the data are considered adequate if Y is not less than

(6) $Y_m = (4.3\ t_{10} \log R)^2$

where t_{10} is the value of Students' t at 10 per cent level of significance with Y—6 degrees of freedom and R the ratio of the rainfall at 1% confidence to the same at 50% confidence both being estimated from the graph.

The method outlined above has its spirit the smoothening of the relation between the dependent and independent variable with the help of a mathematical function and then exploiting the same for finding the values of the dependent variable for some chosen values of the independent variable. It is expected to work extremely nicely if the rainfall data, after its arrangement in the descending order, falls approximately into a geometric series. In cases where this is not true it may lead to estimates which are disappointing. Further, drawing the best fitting line to the eye renders the estimation procedure arbitrary as the closeness of the points to a line depends not only on the scale of the graph but also on the subjective judgement of the person concerned with it. In short the method lacks rigour. A more competent method is, therefore, needed to tackle the problem effectively. In this paper the authors have suggested two general methods of estimation of weekly rainfall of a place based on probability considerations. California and Hazen's approaches have been skown to be their special cases. It was easy to suggest other methods closely related to those treated in this paper. However, as all these methods will have the same correlation with the data, it was considered rational to suggest only those which are generalizations of the two known methods, namely California and Hazen's approaches. A new criterion for judging the adequacy of the data has also been suggested.

Materials and Methods

As before it will be assumed throughout this paper that $X_1; X_2 \ldots, x_Y$ is the arrangement in descending order of magnitude of annual rainfall data for Y years. Similar arrangement can be made of the data for any week. Since X_i are observed rainfalls it is only reasonable to expect uniform probability for rainfall between X_{i+1} and X_i and equal probability forgetting rainfall in the interval X_{i+1} to X_i, $i = 1, 2, \ldots, (Y-1)$ in any future year. Further, though the X_i are the only observed rainfalls, it is not certain that the rainfall in a future year will not fall outside the range X_Y to X_1. Thus, if X standards for the rainfall during an arbitrary year. We shall assume that

i) for any given X_1 and X_{i+1} probability is P that $X_{i+1} < X < X_i$, $i = 1, 2, \ldots, Y-1$

ii) p_1 is the probability that $X > X_1$, and

(iii) p_2 is the probability that $X < X_y$
Clearly we should have then

$$(7)\quad p_1 + p_2 + (Y-1)\, p = 1$$

Also, if P(n) is the probability that minimum rainfall of a year is X it should satisfy the relation

$$(8)\quad P(n) = p_1 + (n-1)\, p$$

Assigning different sets of values to p_2 and p subject to the relation (7), when p_1 will be automatically determined, different expressions for P(n) can be obtained. Of these only the following two have been adopted:

(9) $p_1 = \frac{2}{k(Y+1)}$ and $p = \frac{}{Y+1}$ for $k > 2$, and

(10) $p_1 = \frac{1}{} \quad \frac{1}{Y}$ and $P = \frac{1}{Y}$

Making use of (9) and (10) successively in (8) and denoting the resulting forms of p(n) respectively by P_n and P_n' it is easy to see that

(11) $P_n = \frac{}{k(Y+1)} \left\{ 2 + k(n-1) \right\}$, and

(12) $P_n' = \frac{}{rY} \left\{ (r-1) + r(n-1) \right\}$

The corresponding confidence percentages and are obtained by multiplying the right hand sides of (11) and (12) by 100. Thus

(13) $F_n = \frac{100}{k(Y+1)} \left\{ 2 + k(n-1) \right\}$ and

(14) $F_n' = \frac{100}{rY} \left\{ (r-1) + r(n-1) \right\}$

On putting k = 2 in (13) and r = 2 in (14) they reduce respectively to (4) and *(5)* thus showing that California and Hazen's approaches are particular cases of the general results established in this paper. Since (13) and (14) can be rewritten as

$$\frac{100\, n}{Y+1} \quad \frac{100}{Y+1}\left(\frac{2}{k} - 1\right) \text{ and}$$

(16) $F_n' = \frac{100\, n}{Y} - \frac{100}{rY}$

it follows that each can be expressed as a linear function of the other for any given set of values of k, r and Y. This means that the coefficients of correlation between the F_n or F ' and X_n ill b ' Hence, fitti linear functions in which rainfall is considered as the dependent variable and either F_n or F ' as the independent variable the closeness of fit obtained will be the same.

Now, as uniform probability has been attributed to any rainfall between X_i and $X_{i\ 1}$ whose confidence percentages have been taken as C_i and, C. the minimum rainfall X corresponding to a confidence coefficient C for which C < C < C_1 can be estimated from the relation

(17, X— X_{i+1} I \ $\frac{X. - X._1}{C_{i,1} - C_i}$ > ($C_{i,1}$ C), i—1, 2 ..... , Y—1).

In order to obtain estimates in tune with the confidence percentages developed

n' are given in

The pairing of X_i and C. discussed in the introduction is meant to introduce a negative correlation between the two variables. This relation can be exploited to determine the adequency of the data. If r is the observed coefficient of correlation between X_i and $C_i = F_i'$ as given in (14)

(18) Z = $\frac{1}{2}$ $\log_e$

can be assumed to be distributed as a normal variable with variance 1/Y— 3 (Fisher, R. A. 1954). The 99 per cent confidence limits of Z are then given by

(19) (Z-a) √

where 'a' is the mean of the distribution of Z and on an average the value of (Z — a) will be zero. Since negative correlation has been induced between X_i and F_i', the likelihood is more for the average value of Z to be different from zero. Hence, if Z [stands for the numerical value of Z, on the average

(20) Z $\sqrt{Y - 3} \geq 3$

that is,

(21) Y • 3 + -

The right hand side of (21) rounded to the next integer will denote the minimum number of years of rainfall data necessary for the purpose in view.

The procedure described above can be applied to weekly data for obtaining estimates of weekly rainfalls.

Results and Discussion

The estimation of weekly rainfall of a place will be attempted only on the basis of adequate number of years of data as determined either by (20) or by some other relation of the type (6). In any case such data would cover the normal range of rainfall of the place. But it would be unwise to think that the rainfall will not fall below the observed lowest rainfall X_y or that it will not go above the highest recorded rainfall X_i. Each of these situations will have only little chance. The obvious implication is that, without the chances for either of them being zero, neither should equal or exceed the chance for obtaining a rainfall in any observed interval X_i to X_j, j = 1, ..., Y — 1. Observing this principle only (14) with r > 1 will be suitable for assigning confidence percentage to X_n. Thus, while Patro *et al* 1975 leaves the farmer in the dark as to which method of attributing confidence percen-tages is to be pursued the present discussion has succeeded in suggesting a definite procedure for it. The method suggested here has the additional ad-vantage of being flexible. At times it may be known from the data the existence of some limit point either at the higher end or at the lower end of the observed rainfall. On such occasions, in accordance with the confidence of concerned person or persons the value of r could be adjusted to attribute the needed chance for either exceeding X_1 or for receding from X . For instance in the rainfall data for Trichur, on arrangement in the descending order of magnitude, two lower values have been found to be close together thus indicating probably a limit point at the lower end. In the computation of the weekly rainfall of the place p_1 and p_2 were taken in the ratio 2:1. The estimated weekly rain-falls on the basis of 15 years of data by the method outlined in this paper are given in table 1. The corresponding estimates by the least square equivalent of the method due to Patro *et al* 1975 were in many cases disappointing in the sense that the minimum rainfall with 10 per cent confidence far exceeded the highest observed rainfall for the week. Similar incongruity was found with regard to the estimates with 90 per cent confidence. As these incongruities were quite frequent a regression of the square root of the rainfall on the confidence percantage was tried but without substantial improvement. Linear regression was also not found to be giving satisfactory results in all cases. A reason for this is that the distributions of annual rainfalls and weekly rain-falls do not have the same form throughout. The method outlined in this paper assumes uniform distribution for each observed interval and therefore will not lead to estimates which are incongruous in any situation.

One of the important contributions of this paper is a new procedure for determining the adequacy of the data for the purpose of estimation A general comparison of this procedure with the one given in Patro *et al* and reproduced in (5) is not attempted here. In the case of the rainfall data of

Table 1

Estimated weekly Rainfall (m. m.) at Trichur on the basis of 15 years of data at various confidence Levels

Sl. No	Week Date		$Fn = \frac{100 (rn-1)}{ry}$ r = 3 Confidence levels 10%	50%	80%	90%
1	2		3	4	5	6
1.	January	1—7	0	0	0	0
2.		8—14	0	0	0	0
3.		15—21	0	0	0	0
4.		22—28	0	0	0	0
5.		29—4 February	0	0	0	0
6.	February	5—11	0	0	0	0
7.		12—18	19.9	0	0	0
8.		19—25	43.1	0	0	0
9.		26—4 March	20.7	0	0	0
10.	March	5—11	0	0	0	0
11.		12—18	0	0	0	0
12.		19—25	27.9	0	0	0
13.		26—1 April	30.1	0	0	0
14.	April	2—8	59.0	2.4	0	0
15.		9—15	27.0	4.0	0	0
16.		16—22	52.3	5.0	1.7	6
17.		23—29	36.9	8.0	0	0
18.		30—6 May	102.8	33.2	6.2	0
19.	May	7—13	79.8	20.5	0	0
20.		14—20	240.0	35.8	0	0
21.		21-27	253.9	52.2	11.7	3.4
22.		28—3 June	202.8	56.0	29.8	21.2

1	2			3	4	5	6
23.	June	4—10		357.0	84.7	24.6	23.7
24.		11—17		276.6	122.4	75.6	37.6
25.		18—24		342.4	139.7	89.2	23.0
26.		25— 1	July	308.7	157.8	90.2	58.9
27.	July	2— 8		458.5	216.9	91.3	76.7
28.		9-15		407.5	236.9	135.7	93.8
29.		16—22		407.0	117.3	87.1	80.2
		23—29		322.8	136.0	53.7	46.7
31.		30— 5	August	217.6	109.7	72.3	59.6
32.	August	6—12		218.7	120.9	69.4	42.9
33.		13—19		264.9	103.0	46.3	20.1
34.		20-26		115.6	78.4	50.0	32.2
35.		27— 2	September	135.2	80.1	19.9	4.0
36.	September	3— 9		104.8	39.3	19.4	8.8
37.		10—16		200.0	21.4	4.5	0
38.		17--23		196.4	46.9	0	0
39.		24—30		193.2	67.0	4.3	0
40.	October	1— 7		130 4	39.1	4.0	0
41.		8—14		1469	63.6	22.2	12.5
42.		15-21		132.6	85.1	95	4.6
43.		22—28		160 5	38.0	5.8	0
44.		29 —• 4	November	149.6	22.7	.5.3	1.6
45.	November	5—11		97.2	4.6	0	0
46.		12—18		56.7	5.1	0	0
47.		19—25		81.2	92	0	0
48.		26— 2	December	31.9	0	0	0
49.	December	3— 9		93.8	4.5	0	0
50.		10—16		42.9	0	0	0
51.		17—23		34.1	0	0	0
52.		24—31		13.9	0	0	0

Trichur the minimum number of years of data required was found to be 6 by the former method and 8 by the latter. The minimum number of years of rainfall data necessary for estimating weekly rainfall is given in Table 2 for different values of the coefficient of correlation between X. and C.- = F '.

Table 2

Minimum number of years of data necessary for etimating the weekly rainfall for different values of the coefficient of correlation between X and F '

Correlation coefficient r	Minimum number of years of data rounded to next integer
– 0.95	6
— 0.90	8
— 0.85	y
— 0.80	11
— 0.75	13
— 0.70	15
— 0.65	18
— 0.60	22
— 0.55	27
— 0.50	34
— 0.45	42
– 0.40	54
— 0.35	71
— 0.30	98
— 0.25	142
— 0.20	223
-0.15	397
– 0.10	1143
—0 05	3603

Summary

New methods for estimating the weekly rainfall of a place has been developed in this paper. The California and Hazen's methods found in Patro *et al* have been shown to be their special cases. It has further been pointed out that the method, of which Hazen's is a special case, is rationally more appropriate than the other though all of them are linearly related. A new but exquisite criterion for determining the adequacy of the data has also been established.

Acknowledgement

The authors are deeply indebted to Sri. N. Kaleeswaran, Vice-Chancellor, Kerala Agricultural University for inviting their attention to the problem and to Dr. P. Nair, Dean, College of Veterinary and Animal Sciences for his kind interest in the work.

സംഗ്രഹം

ഒരു പ്രദേശത്തു്, ആഴ്ചതോറും fflnjegprycrn മഴയുടെ അളവ് കണക്കാക്കുന്നതിനുള്ള പുതിയ ro"lroi1tft>g06na0 ഈ പ്രബന്ധത്തിലെ പ്രതിപാദ്യം. ഇവയുടെ പ്രത്യേക രൂപങ്ങളാണ്, പേടോ തുടങ്ങിയവർ അവലംബിച്ചിട്ടുള്ള കലിഫോർണിയാ രീതിയും ഹേസൻ രീതിയും എന്നു സ്ഥാപിച്ചിരിക്കുന്നു. രേഖീയ ബന്ധമുള്ള ഈ രീതികളിൽ, ഹേസൻ രീതിയുടെ സാമാന്യവൽകൃതരീതി മറ്റേതിനേക്കാളും യുക്തി സഹമെന്നും സൂചിപ്പിച്ചിട്ടുണ്ട്. ആ കലനത്തിന്ന് ആവശ്യമായ ഡാററയുടെ പര്യാപ്തതയെ നിർണ്ണയിക്കുന്നതിനുള്ള ഒരു നൂതന മാനദണ്ഡവും ആവിഷ്കരിച്ചിരിക്കുന്നു.

REFERENCES

Cramer, 1962 *Mathematical Methods of Statistics*, first Indian Edition, Asia Publishing House. Bombay.

Feller, W 1968. An *introduction to Theory of Probability and its applications*, Vol 1, Third Edition, Wiley International Edition, John Wiley and Sons, Inc. New York.

Fisher, R. A. 1954. *Statistical Methods for Research Workers*, 12th Edition, Oliver and Boyds, London.

Patro, T. G., Srivastava K. L. and Misra, C 1975. Probability Analysis of Weekly Rainfall at Bhubaneswar. *J. Res. 2, 26.*

(M. S. *received: 3-6-1977*)