Jal Vigyan Sameeksha—a publication of
High Level Technical Committee on Hydrology
Vol. II No. 2, December 1987

Precipitation Network Design

K. Mukherjee and Surinder Kaur
India Meteorological Department, New Delhi

ABSTRACT : *For proper assessment and utilization of available water resources of the country, optimum design of network of raingauges assumes primary importance. The objectives of network in specific physical terms have to be identified for research and design purposes. These become complicated when one has to quantify these physical parameters for mathematical formulation taking into account the cost and benefit aspects of the project. In this paper, important aspects of precipitation network design have been discussed. Errors in estimation of areal precipitation have been explained. Need for scientists and engineers to harmonically synchronize the demands of the scientific procedure and financial resources to derive optimum network density has been emphasised.*

1. Introduction

For proper utilisation of the available water resources of the country, its precise assessment over an area for a predetermined period of time is an absolute necessity. Both these, period of time and area vary according to the purpose for which the water management is required. This assessment can be finalised by an optimum design of network of gauges from which the relevant data are to be collected.

The variation in time and space of the network of gauges for different projects are numerous. Generally, this can be classified into two broad categories depending on the purpose for which the network of gauges is to be utilised. These are (a) diagnostic and (b) prognostic. Diagnostic studies involve water balance measurements, computation of areal precipitation etc. whereas, precipitation prediction for varying periods comprises prognostic studies. Generally speaking these are the objectives of network design. However, for actual design purposes scientists and engineers have to identify the objectives in specific physical terms which is very difficult. These become more complicated when one has to quantify these physical parameters for mathematical formulation. This also has to take into account the cost and benefit aspect of the project. The last requirement though an extremely desirable property, is not always amenable to straight forward logic. Even the data required for this purpose may not be easily available. Hence, the process of network design is largely limited to obtaining reliable observations from optimally distributed stations for estimating various meteorological parameters.

2. General Consideration

Since precipitation varies both in space and time, in most general form it may be represented at any point is P (x, y, z, t). The variation of P with altitude is important in mountainous regions and may be neglected if the area under consideration lies in a more or less flat-terrain. A precipitation event at a point of duration.

$t = t_0$ is defined as

$$P(x, y) = \int_0^{t_o} P(x, y, t)\, dt \quad (1)$$

The areal means of this precipitation event may be given by

$$\overline{P_A} = \frac{1}{A} \int\int_A P(x, y)\, dx\, dy \quad (2)$$

This has to be estimated from the average of point observations recorded by n gauges installed in the catchment, the estimated, value of P_A is given by

$$\hat{P}_A = \frac{1}{n} \sum_{\lambda=1} P_i \quad (3)$$

where P_i is the precipitation recorded by the ith gauge in the network.

If $n \rightarrow \infty$, $\hat{P}_A \rightarrow \overline{P_A}$

The mean square error of this estimate will be

$$e^2 = E(\overline{P_A} - \hat{P}_A)^2 \quad (4)$$

The aim of optimum network design is to minimise e.

When we are concerned with the long term rainfall process, the long terms mean precipitation (monthly, seasonal or annual) over the areas may be written as

$$\overline{P} = \frac{1}{AT} \sum_{t=o}^{T} \int_A p(x, y, t)\, dA \quad (5)$$

where in principle $T \rightarrow \infty$

This is estimated by precipitation data of n observing stations in the area for a period of T years, seasons or months. The estimate of P_A is given by

$$\hat{P}_A = \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=0}^{T} P_{it} \quad (6)$$

where P_{it} is the precipitation of i^{th} recording station in t^{th} period of times such as day, month year etc.

The variance of the estimate of $\overline{P}$ defined by $V(\overline{P}) = E(\overline{P} - \hat{P})^2$ (7).

This has to be reduced to that level at which the desired accuracy is achieved.

The earliest attempt in network design was by Benton (1920). He presumed that number of gauges required for network design is a function of area alone. Later, Horton (1923) proposed the network design in which he expressed error (δ) of the areal estimate as

$$\delta = \frac{KR}{\sqrt{n}} \quad (8)$$

where R is the range of precipitation, n is the number of observing points and K is constant. Ganguly et. al. (1951) introduced the standard deviation in place of range of precipitation. They assumed that this precipitation distribution in space over the catchment is not far from normal and proposed that

$$n_r = n_e \left\{ \frac{\text{Coefficient of variation of mean rainfall in the catchment}}{\text{desired coefficient of variation}} \right\}^2 \quad (9)$$

where n_r is the required density of gauges for achieving the desired coefficient of variation and n_e is the existing no. of gauges.

Lee et. al (1956) applied the concept of Student's t-statistic to evaluate the sample size n (no. of gauges) in order to minimise the difference $| x - \mu |$, where μ, the true areal mean, is estimated by $\overline{x}$, student t statistic is given by

$$t = \frac{|\overline{x} - \mu| \sqrt{n}}{s} \quad (10)$$

where s is the unbiased estimate of the population standard deviation given by

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{x})^2 \quad (11)$$

in this case, the network should be so designed that

$$\frac{|\overline{x} - \mu|}{\overline{X}} \leqslant e \quad (12)$$

where e is allowable error

From eqn. (10) and (12), it may be seen that

$$n \geqslant \left(\frac{C_v \cdot t}{e}\right)^2 \quad (13)$$

where $C_v = \frac{s}{\overline{X}}$

which gives the optimum no. of gauges for desired accuracy. Ahuja (1980) assumed t = 1, in which case, the above equation reduced to

$$n \geqslant \left(\frac{Cv}{e}\right)^2 \quad (14)$$

The above equations are derived on assumption that the distribution of $\overline{X}$ is normal or near normal.

Taking different factors which influence rainfall distribution over an area, WMO (1975) has recommended minimum density of raingauges for determination of areal rainfall in different regions which is given in Table-1.

Table 1 : Minimum density of precipitation stations

Region	Minimum density range (Km²/gauge)
1. Temperate, Mediterranean & Tropical zones	
(i) Flat areas	600—900
(ii) Mountainous areas	100—250
2. Small Mountainous islands ($\nless$ 20,000 km²).	25
3. Arid & Polar zones	1500—10,000

In previous paragraphs, all the derivations have been made based on the implicit assumption that the observations at different gauges are unrelated. However, a more realistic approach to this problem would be to take into account the intergauge rainfall variability.

In a relatively homogeneous area, the correlation coefficient between precipitation series at the gauge point i and j may be defined as

$$r_{ij} = \frac{\text{Cov}\,(P_i, P_j)}{\sqrt{\text{var}\,(P_i)\ \text{var}\,(P_j)}} \quad (15)$$

where P_i and P_j denote the time series of precipitation at points i & j respectively. r_{ij} varies with the distance between the gauges. Kagan (1972) expressed this relation as

$$r\,(s) = r_o \exp(-s/s_o) \quad (16)$$

where r (s) represents the correlation coefficient between precipitation of gauges situated at mean distance s.

Hershfield (1965) was perhaps the first to use the concept of correlation between gauges to design their proper spacing. Kagan (1966) approached the problem in a more systematic way. According to him the random error in the precipitation measurement may be determined by

$$V(e) = [1-r_o]\ V(p) \qquad (17)$$

where $V(p) = ((\overline{p^2} - (\overline{p})^2)$ is the variance of precipitation time series at a fixed point in the area.

When the average precipitation over an area 'a' is estimated by a raingauge located at its centre, Kagan (1966) expressed the variance of the error in this estimate as

$$V(\epsilon) = V(e) + 0.23\ \frac{\sqrt{a}}{s_o}\ V(p) \qquad (18)$$

It is derived on the assumption that r (s) exists in the area and is described by the Eqn. (16). When there are n gauges evenly distributed in the area under study, such that A = na, the variance of error ϵ in the average rainfall over A is given by

$$V_e(\epsilon) = \frac{V_{(e)}}{n} + \frac{0.23\ \sqrt{A}}{s_o\ n^{3/2}}\ V(p) \qquad (19)$$

Thus for any prefixed value of $V_n(\epsilon)$, n can be evaluated. It may be mentioned here that the first term of eqn. (19) is due to random errors and the second term is distributed to spatial variation in the precipitation field. The relative root mean square error which directly gives the accuracy of the estimate, may be written as

$$Z = \frac{\sqrt{Vn\ (\epsilon)}}{\overline{p}} = Cv \sqrt{\frac{1-r_o}{n} + \frac{0.23\ \sqrt{A}}{s_o\ n^{3/2}}} \qquad (20)$$

where $C_v = \frac{\sigma_p}{\overline{p}}$ and $\overline{p}$ is the average precipitation over the area.

The uniform spacing of stations can be achieved either on the basis of square grid or triangular grids. In the case of square grid, the spacing between the stations will be given by

$$l = \sqrt{\frac{A}{n}} \qquad (21)$$

In the case of triangular grid which is considered to be more convenient for triangular catchments, the spacing l is given by

$$l = \sqrt{\frac{2A}{n\sqrt{3}}} = 1.07 \sqrt{\frac{A}{n}} \qquad (22)$$

3. Error in estimation of areal precipitation

3.1 For a rainfall event

The mean rainfall of a rainfall event over an area 'A' may be estimated by the formula.

$$\overline{p_A} = \frac{1}{A} \int_A p\ (x_i)\ d_{xi} \qquad (23)$$

where $p\ (x_i)$ is the precipitation recorded at the i^{th} station.

In the case of n observations in the area 'A' being used for estimation of $\overline{p_A}$, we may write the estimate $\hat{p}_A$ as

$$\hat{p}_A = \frac{1}{n} \sum_{i=1}^{n} p\ (x_i) \qquad (24)$$

The performance of the network is judged by the magnitude of $V\ (\overline{p_A})$ which decreases as the gauge density increases. In this case $V\ (\overline{p_A})$ is given by

$$V\ (\overline{p_A}) = \sigma^2\ \psi\ (n, \overline{r}), \qquad (25)$$

where $\psi\ (n, \overline{r}) = \frac{1-\overline{r}}{n}$

and $\overline{r}$ is the mean correlation over the area 'A'

3.2 For long term areal precipition

Rodriguez & Mejia (1974) have given the expression for variance of long term areal precipitation given by Eqn (6) as

$$V\ (\overline{p_A}) = \sigma^2\ J\ (T)\ \psi\ (n, \overline{r}) \qquad (6)$$

where $J\ (T) = \frac{1}{T}\left(\frac{1+e}{1-e}\right)$ (temporal reduction factor), e is lag 1 autocorrelation in time series and

$$\psi\ (n, \bar{r}) = \frac{1 + (n-1)\bar{r}}{n}$$ (spatial reduction factor),

$\bar{r}$ is the mean correlation over the area A

4. The network density (n)

The optimum number of gauges required for the estimation of areal precipitation depends upon the accuracy desired. If the correlation structure $\bar{r}$ is not considered, then e may be computed from the expression

$$e = \sqrt{\frac{V(\bar{P}_A)}{\bar{P}_A}}$$

$$n = \left(\frac{C.V.}{e}\right)^2 \qquad (27)$$

If the correlation structure of precipitation field ($\bar{r}$) is considered, then for a rainfall event,

$$n = \left(\frac{C.V.}{e}\right)^2 (1 - \bar{r}) \qquad (28)$$

It may be seen from eqn. (27) & (28) that if the correlation structure of precipitation field ($\bar{r}$) is taken into account, a lower number of precipitation gauges are required for estimation of areal precipitation with same degree of accuracy (e).

5. General remarks

From discussions made in foregoing paragraphs it will be evident that network design ultimately is a process of minimising error that is inherent in any such exercise. Scientists and engineers should harmonically synchronize the demands of the scientific procedure and the financial resources at their disposal to derive optimum density which is economically viable and technically acceptable.

1. Ahuja, P.R. (1960), "Planning of precipitation network for Water Resources Development in India". United Nations, ECAFE and WMO Hydraulic Network and methods Flood Control series No. 15, Bangkok, Thailand.

2. Benton, J. (1920), "Irrigation Pocket Book" E & F.N. Spon. Ltd., London.

3. Ganguly, M.K. et. al. (1951), "Accuracy of Mean Rainfall Estimates of Damodar Catchment" Irrigation, Power J.B., pp 278-284.

4. Hershfield, M. (1965), "On the Spacing of Raingauges" Sym. on Design of Hydrological networks. Int. Assoc. Sci. Hyd. Publ. 67, pp 72-75.

5. Horton, R.F. (1923), "Accuracy of Areal Rainfall Estimates" Monthly Weather review.

6. Kagan R.L. (1966), "An Evaluation of the Representativeness of Precipitation data" Works of the Geophysical Observatory, U.S.S.R. Vol. 191.

7. Kagan, R.L. (1972), "Precipitation-Statistical Principles". WMO case book on Hydrological network design practice No. 324.

8. Lee, S.C. (1956), "Length of Record and Number of Stations Required in Statistical Analysis of Hydrologic Data." Proc. of Reg. Tech. conf. on Water Res. Dev. in Asia & Far East, U.N. — ECAFE Flood Control Series No. 9, pp-184-191.

9. Rodriguez, Iturbe, & Mejia, J.M. (1974), "The Design of Rainfall Network in Time and Space", Water Resources Research, Vol. 10, No. 4, pp 713-727.