Copyright © IFAC, Water and Related Land Resource Systems
Cleveland, Ohio 1980

HYDROLOGIC NETWORK DESIGN METHODS AND SHANNON'S INFORMATION THEORY

T. Husain* and W. F. Caselton**

**Associate Research Engineer, Research Institute, University of Petroleum and Minerals, Dhahran, Saudi Arabia*
***Assistant Professor, Department of Civil Engineering, University of British Columbia, Canada*

Abstract. Various existing methods of hydrologic network design are reviewed. These methods are grouped into mean-square error, inter-station correlation and iso-correlation, regionalization, simulation, systems analysis, rational and the simplified form of Fisher's information approach. The applicabilities and limitations of each of the above approaches are discussed in context with the hydrologic network design and estimation objectives. The relative performance of each of the above approaches is then evaluated in selecting the optimum station locations and estimating hydrologic events at ungauged locations using simulated rainfall data. It is found that due to normality and linearity assumptions in above approaches, the existing methods are not very effective in identifying true optimum station locations. A broader and more universal methodology based on Shannon's information measure is proposed which is able to select the optimum station locations without the assumptions of normality and linearity. It also treats the network in true multivariate form. The methodology is then extended to estimate the events at ungauged locations on the basis of multivariate optimum information transmission criterion.

Keywords. Correlation methods; design; entropy; information theory; multivariable systems; network; probability; water resources.

INTRODUCTION

Although a number of approaches of hydrologic network design are cited in the literature but most of these approaches can not be generalized on scientific and rational basis due to the following unresolved problems:

(a) The cost of installing and operating a hydrologic network and the probable future benefits from the data transmitted by the network in planning water resources schemes are the two main factors affecting the design of a hydrologic network system. The problem of attaining optimal balance between these two factors, mainly due to difficulties in evaluating the probable benefits, have not yet been resolved.

(b) Hydrologic network is an organized system for the collection of information of specific kind. It transmits information on hydrologic phenomena which vary in space and time domains. Since the hydrologic phenomena in a region, which is the ensemble of the hydrologic events at various point locations, is space and time dependent, the data collection network, in true sense, is of multivariate in nature. Due to mathematical complexities involved in dealing with the multivariate techniques, the research on hydrologic network methodology has been ignored in the past.

(c) Another challenging problem, the hydrologic network design is faced with, is of combinatorial type. Suppose we want to select a set of 'n' optimum station locations from a region with 'm' feasible locations. To obtain a true optimum station set, it is necessary to evaluate $\frac{m!}{(m-n)!n!}$ possible combinations. It is computationally infeasible for values of 'm' and 'n' greater than, for example, n=5 and m=50. Therefore, for effective design of hydrologic network, it is essential to investigate possible techniques to reduce the combinatorial search to acceptable computational load.

A REVIEW OF THE EXISTING METHODS

A number of methods of hydrologic network design are cited in the literature. Some representative approaches are briefly described here.

Hydrologic network design was initially based on the relationship of the mean square error, between the observed and estimated values of the hydrologic events, to the distance between the station points (Drozdov, 1936; Ganguli and others, 1951; Horton, 1923). This approach was first established by Horton (1923), later by Drozdov (1936), for two station points. Ganguli and others (1951) generalized this approach and proved that the number of stations in a region would be inversely proportional to the square of the coefficient of variation of the hydrologic variables. The generalized form was expressed as:

$$N_r = N_e.(C_{ve}/C_{vr})^2 \qquad (1)$$

where

N_e is the existing number of stations;

C_{ve} is the coefficient of variation of the variables in the existing networks;

C_{vr} is the coefficient of variation desired in a network of N_r stations.

A second approach was based on the inter-distance correlation relationship. It was assumed that the correlation between the hydrologic events at two stations decreased with the distance between the stations. The decay function of the coefficient of correlation between two stations was assumed to be expotential (Stol, 1972). The relative efficiency was then obtained from the following relationship:

$$\eta_r = [r(x,t)/\rho(t)]^2 \qquad (2)$$

where

η_r is the relative efficiency;

$r(x,t)$ is the correlation coefficient for the inter-station distance x at time t;

$\rho(t)$ is the correlation coefficient between two point locations for very dense network at time t.

The method of isocorrelation proposed by Hershfield (1965) was also based on the inter-station correlation approach. The correlations between one or more key stations and all other stations in a network were calculated and lines of equal correlations plotted around each of the key stations.

A third approach to network design, called the simulation approach, was based on the improvement of the estimates of selected statistical parameters, such as the mean and variance of the hydrologic time series data, by the use of primary and secondary stations (Brass and Rodriguez-Iturbe, 1976; Fiering, 1965; Karlinger, 1974). Primary stations, often referred to as time sampling stations, were used to ascertain the hydrologic time series relationships. Secondary stations, referred to as space sampling stations, were used for a short time to establish spatial relationships. A decision could be made to either continue or discontinue a station in a region based on the requirement of the accuracy of the statistical estimates of the parameters (Fiering, 1965b).

The regionalization approach, an advancement in mapping techniques by dividing the whole area into square grids, proposed by Solomon and others, is also applicable to hydrologic network design (Solomon and others, 1968; Solomon, 1972). The proposed method was used to process the hydrologic information from a large area and was used to relate hydrologic variables to physical characteristics. The square grid system was applied to estimate the runoff distribution in a large area using meteorologic and hydrologic informations. Estimates of the hydrologic parameters at the ungauged sites could also be obtained on the basis of estimates of the parameters of rainfall and runoff at the gauged points.

Another approach to network design has been named the rational approach and this takes into account the demographic, economic, meteorologic and basin characteristic factors (Desi and others, 1965; Uryvaeu, 1965). Multiple linear regression was used in assessing the network performance on the basis of the estimation error at the ungauged points in the basin (Benson, 1972). In addition to the above factors, the drainage area could also be accounted for in the multiple regression analysis (Benson, 1965). The variance of the territorial mean of the hydrologic sampling was used as the basic criterion for determining the density of the precipitation network rationally (Desi and others, 1965). The approach is empirical in nature and is not universally applicable.

The latest development in network planning is the application of systems analysis and decision theory (Davis and others, 1972; Dawdy and others, 1970; Jacobi, 1975; Moss, 1972; Moss and Dawdy, 1973). Langbein (1954) has been able to formulate the network design problem and determined the number of primary as well as secondary stations by taking into account the desired

741

accuracy and a budgetary constraint. The method proposed by Fiering (1965a) is similar to Langbein's except for the introduction of an objective function and deciding whether the stations are to be continued or not.

Jacobi (1975) applied a decision theoretical approach to determine the optimum economic record lengths using the concept of the economic balance between the information value and the information cost. The optimum economic record length was defined as a function of the size of the data sample. The methodology, which utilized a Bayesian decision framework, was then applied to a sediment deposition problem.

Decision theory has also been applied in the optimum design of a mountainous rain-gauge network (Davis and others, 1972). A conceptual framework interrelating the uncertainty of a model parameter with the record length was established. Moss and Dawdy (1973) attempted to improve this methodology by combining it with a Monte Carlo simulation. Initially the statistical properties were assumed to be known but, at a later stage, the condition of assumed statistical characteristics was relaxed by incorporating a prior distribution on the unknown statistics.

Very recently Fisher's information criterion has been introduced in the field of hydrologic network design (Matalas and Gilory, 1968). In the early 1920's Fisher defined the information content in a sequence of observations as the reciprocal of the variance of the estimates of the parameter of interest. Matalas (1973) proposed two basic approaches for designing the optimum gauging schemes in a region, both utilizing Fisher's information criterion. The first approach was based on the identification of those stations which were to be discontinued from a dense network due to budget curtailments. It was based on the principle of maximizing the total information content. The second approach was based on a marginal information concept. In this case the stations in a region were discontinued in such a manner that the decrease in the value of Fisher's information was minimized.

An information transfer criterion from gauged to ungauged points was discussed by Wallis and Matalas (1972). The characteristics of the data series were assumed to be Markovian. The model thus developed was an improvement over a previous model by Maddock (1977) by relaxing the assumptions of linearity and stochastic independency through the application of a Monte Carlo technique. Maddock (1977) and Carrigan and Golden (1975) applied the above methodologies to the real world situations using a mixed integer programming formulation along with a number of budgetary and information transferability constraints.

LIMITATIONS OF THE EXISTING METHODS

The principal limitations of the network design methods discussed above can be summarized as follows:

1. The approach based on the mean square error and its relation to interstation distance is a useful approach to the preliminary design of a network. However, it can not be used for the accurate determination of the network density since Eq. (1) does not take into account the heterogeneity of the hydrologic variables which exists in mountainous regions. It is also difficult to decide upon the value of the required coefficient of variation in Eq. (1). Furthermore, this approach is not able to discriminate between primary and secondary stations in a region.

2. In methods of isocorrelation and interstation distance correlation there is no rule to provide the value of the relative efficiency and the correlation coefficient for a very dense network. The term dense network is itself a relative term which is difficult to assess. Since the correlation coefficient is a function of time and space it may vary from month to month, hence the stationary time series assumption may not be valid.

3. The data synthesis and simulation approach to network design is based on the assumption that the hydrologic parameters must be known. Hence the accuracy of this approach depends upon the extent of the prior knowledge of the hydrologic parameters. During the network planning stage very little prior knowledge concerning these parameters is available and, as a consequence, the methodology is not effective.

4. The Bayesian decision approach accounts for the uncertainty in the estimates of the hydrologic parameters through the use of the joint and conditional probability distribution of the parameters being estimated. Although in a simplified form it is a straightforward approach, it can only be used for single time steps into the future to make the decision whether or not to collect data. For multiple time step decisions, the mathematical complexity increases drastically due to the combinatorial nature of the problem. Hence the method is too restrictive for network design which needs more foresight than the methodology provides.

5. The concept of applying Fisher's information measure to the design of a hydrologic network is based on optimizing the reciprocal of the variance of the statistical parameters of the variable being estimated. One criticism of this measure is that Fisher's information content always has an infinite upper bound regardless of the specific

circumstances. In the context of hydrologic network design, Fisher's information relies upon linearity and normality assumptions.

Each of the existing approaches to network design described above, address themselves to very specific problems and depend upon very restrictive assumptions. Hence, a broader, more universal basis for the performance, evaluation, and design of spatial data collection networks is needed.

The applicability of the existing methods as discussed above depends upon the factors such as the nature and the extent of the resource development; the required precision in estimating the hydrologic characteristics; the physical size of the basin; and topographical variation etc. Among these approaches, the methodology based on Fisher's information measure has recently been applied in hydrologic network design although in a limited manner due to the linearity and nomality assumptions in its derivation. The proposed methodology based on Shannon's information measure is, however, not restricted to such assumptions. It has been recognized as being relevant in physical as well as statistical contexts in a number of disciplines and is also interpretable in stochastic processes. Due to well defined multivariate form of Shannon's information theory, it has effectively been applied in precipitation network design (Husain, 1979).

The following section deals with a comparative study of Shannon's and Fisher's information measures.

COMPARISON OF SHANNON'S AND FISHER'S INFORMATION MEASURES IN HYDROLOGIC NETWORK DESIGN APPLICATIONS

Bivariate Case

Consider two hydrologic variables, $X_j = (x_{j,1}, x_{j,2}, \ldots, x_{j,N_1+N_2})$ and $X_i = (x_{i,1}, x_{i,2}, \ldots, x_{i,N_1})$ with (N_1+N_2) and N_1 observations respectively. After collecting N_1 observations concurrently on X_j and X_i, the data collection scheme for the variable X_i is discontinued but is extended to N_2 additional measurements on X_j. Based on the concurrent N_1 observations, a linear regression, with X_i as dependent variable and X_j as independent variable, is carried out and yields the linear model:

$$\hat{x}_{i,k} = \bar{X}_i + b(x_{j,k} - \bar{X}_j) \qquad (3)$$

Using the above equation, N_2 estimates of X_j based on the N_2 additional observations of X_i are obtained. Combining the N_1 observations and N_2 additional estimates, the weighted estimates of the mean $\hat{\mu}_i$ is derived as follows:

$$\hat{\mu}_i = (N_1\bar{X}_i + N_2\hat{\bar{X}}_i)/(N_1+N_2)$$

where,

$$\bar{X}_i = \sum_{k=1}^{N_1} x_{i,k}/N_1 ; \quad \hat{\bar{X}}_i = \sum_{k=N_1+1}^{N_1+N_2} \hat{x}_{i,k}/N_2$$

The variance of $\hat{\mu}_i$, denoted by σ_m^2 is as follows:

$$\sigma_m^2 = \frac{\sigma_{x_i}^2}{N_1}\left[1 - \frac{N_2}{N_1+N_2}\{\rho_{ij}^2 - (1-\rho_{ij}^2)/N_1 - 3\}\right] \qquad (4)$$

where,

$\sigma^2_{x_i/N_1}$ is the variance of the estimates of the mean for the random sequence of X_i with N_1 observations.

ρ^2_{ij} the correlation coefficient between X_i and X_j based on N_1 concurrent observations.

Fisher's information $I_f(X_i;X_j)$, which is equal to the ratio of the variance of the mean for a purely random sequence and the variance of the mean for the sequence of interest, is defined as follows:

$$\left.\begin{aligned} I_f(X_i;X_j) &= 1\Big/\left[1 + \frac{N_2}{N_1}\{(1-\rho_{ij}^2)(\frac{2-N_1}{3-N_1})\}\right] \quad \text{for } i \neq j \\ &= 1 \quad \text{for all } i=j \end{aligned}\right\} \qquad (5)$$

The information transmission by X_j about X_i using Shannon's information concept is defined as (Shannon and Weaver, 1949):

$$\left.\begin{aligned} T(X_i;X_j) &= H(X_i)+H(X_j)-H(X_i,X_j) \quad \text{for all } i \neq j \\ &= H(X_i) \quad \text{for } i=j \end{aligned}\right\} \qquad (6)$$

where

$H(X_i)$ and $H(X_j)$ are entropies of the variables X_i and X_j respectively.

$H(X_i, X_j)$ is the joint entropy of X_i and X_j.

The entropies are defined as follows:

$$H(X_i) = -\sum_{\kappa} P(x_{i,\kappa}) \log P(x_{i,\kappa})$$

$$H(X_i,X_j) = -\sum_{k} P(x_{i,\kappa}, x_{j,k}) \log P(x_{i,k}, x_{j,k})$$

where

$P(x_{i,\kappa})$ is the probability of the concurrence of the κth event of the variable X_i

$P(x_{i,k}, x_{j,k})$ is the joint probability of the occurrence of the kth event of the variables X_i and X_j.

These probabilities are computed using discretization concept.

The information transmittion, as derived by Matalas in Eq.(5), is restricted to the assumptions of linearity and normality. These assumptions may not be applicable in hydrologic network design due to nonlinearity and skewness which commonly exists in time series data. Shannon's information transmission criterion, however, is not restricted to the above assumptions.

Optimum station sets can be obtained by both of these methodologies by defining the objective function and the constraints. The objective for network design, applied here, is to select the optimum station locations for the reduced network so that the maximum possible information concerning the new ungauged locations is transmitted. The objective for a network of "m" stations proposed by Matalas (1973) is:

$$Z = \text{Max} \sum_{i,j=1}^{m} I_f(X_i;X_j)\delta_{i,j} \tag{7}$$

where,

Z is called the objective function

$\delta_{i,j}$ is a decision variable which is zero if the station i=j is to be discontinued or, for i≠j, information cannot be transferred from j to i. On the other hand, $\delta_{i,j}$ is equal to one when the information is transferred from j to i.

Using Shannon's information concept, a similar objective function can be written as:

$$Z = \text{Max} \sum_{i,j=1}^{m} T(X_i;X_j)\delta_{i,j} \tag{8}$$

The following two constraints are common to both forms of network objectives described above:

a - Information transferability constraint:

This indicates that information from only one station j can be transferred to another station i.

$$\sum_{j=1}^{m} \delta_{i,j} = 1 \tag{9}$$

b - Budgetary constraints: if the maximum number of stations which can be retained are n, due to a budgetary constraint, then

$$\sum_{i=1}^{m} \delta_{i,i} = n \tag{10}$$

Multivariate Case

Fisher's information concept has been extended to the multiple regression case where estimation at a point location is based on data from a number of station locations (Matalas, 1973; Carrigan and Golden, 1975). This modified form of Fisher's information equation is simplified, in terms of a multiple correlation coefficient, as follows:

$$I_f(X_i;X_j,X_k,\ldots,X_n) = 1/\left[1+N_2\left\{1-\frac{n-(N_1-2)\,\rho^2_{i\cdot n}}{n-(N_1-2)}\right\}\right] \tag{11}$$

where,

$\rho^2_{i\cdot n}$ is the multiple correlation coefficient between the ith variable X_i and a specific set q of n independent variables $(X_j,X_k,\ldots,X_n)$.

To carry out the analysis to select the optimum n station locations from a dense network of m locations, the objective function is defined as follows:

$$Z = \text{Maximizing} \sum_{i=1}^{m} I_f(X_i;X_j,X_k,\ldots,X_n) \tag{12}$$

$$j,k,l,\ldots,n$$

Subject to $I_f(X_i;X_j,X_k,\ldots,X_n)=0$

if $\rho^2_{i\cdot n} < \dfrac{n}{N_1-2}$

If $(X_1,X_2,\ldots,X_m)$ are variables representing the hydrologic outcomes at various point locations, then the information transmitted by set $(X_j, X_k, \ldots,X_n)$, using Shannon's multivariate information transmission, is as follows:

$$T(X_1,X_2\ldots,X_m;X_j,X_k,\ldots,X_n) = H(X_j,X_k,\ldots,X_n) \tag{13}$$

where

$H(X_j,X_k,\ldots,X_n)$ is the joint entropy of the variable set $(X_j,X_k,\ldots,X_n)$ and is defined as $- \sum_{l} P(x_{j,l}, x_{k,l}, \ldots, x_{n,l}) \log P(x_{j,l}, x_{k,l}, \ldots, x_{n,l})$

where

$P(x_{j,l}, x_{k,l}, \ldots, x_{n,l})$ is the joint probability of l_{th} event of the variable set $(X_j, X_k, \ldots, X_n)$

The objective function is defined as:

$$\text{Max } T(X_1, X_2 \ldots, X_m; X_j, X_k, \ldots, X_n) = \max_{j,k,\ldots,n} H(X_j, X_k, \ldots, X_n) \quad (14)$$

ESTIMATION PERFORMANCE CRITERION

The performance of networks designed by Shannon's and Fisher's methodologies described in earlier section were compared on the basis of their estimation accuracy. Multiple linear regression analysis was used for this purpose. The potential stations were identified by both methodologies on the basis of an initial set of observations at all of the grid point locations in a basin. The values associated with the events at the optimal stations act as the independent variables. The values associated with the events at common grid point locations not appearing in the optimal networks specified by either method act as common dependent variables. Regression coefficients were computed using the first N_1 observations. Estimated values were obtained for the dependent variables using the N_2 subsequent observations of the independent variables and compared with the true values. The error variance was then determined for both methods.

ANALYSIS

The time series data for eight point locations were simulated using the multivariate correlation approach. The input correlation matrix of the simulated data, together with the expected values at each point location, are defined as in Table 1. Out of 300 values simulated at each point location, 150 values were used in optimum station selection by the methodologies based on Fisher's and on Shannon's information. The remaining 150 values were used to study the predictive performance of the optimum selected stations at the common ungauged point locations.

Using the formulation as described in previous section, the optimum station sets and the corresponding information transmission were determined using Shannon's discrete entropy concept at the bivariate levels. The best single, two, and three station sets were (#5); (#3, and #5); and (#3, #4, and #5) respectively as shown in Table 2.

Similarly, the information content of all possible station pairs were obtained using Fisher's information criterion of Eq.(6). The optimum stations were identified by maximizing the information content as summarized in Eq.(7). The optimum station(s) and their corresponding information content are listed in Table 3. The optimum single, two, and three station sets were found to be (#3); (#3, #6); and (#3, #6, and #8) respectively.

At the multivariate level of information transmission, Eq.(14) for Shannon's methodology, and Eq.(12) for the Fisher's information criterion, were applied. The optimum stations in Shannon's case were (#4, and #5); and (#3, #4, and #5). Fisher's information criterion identified point location (#3) as the optimum single station, point locations (#4, #6) as the optimum two station set, and point locations (#3, #4, #6) as the optimum three station combinations. These results are listed in Table 3.

The parameters of the estimation models were computed from the initial 150 simulated observations. Values at the common point locations not included in the list of optimal design by either method (#1, #2 and #8) were then estimated by the optimum station sets obtained by the different criteria. The error variances thus obtained are listed in Tables 4 and 5.

Table 1 Correlation Matrix of the Simulated Data

i \ j	1	2	3	4	5	6	7	8
1	1.0	.35	.34	.31	.30	.13	.10	.05
2		1.0	.30	.25	.22	.20	.16	.14
3			1.0	.50	.42	.36	.33	.21
4				1.0	.24	.18	.13	.21
5					1.0	.51	.35	.25
6						1.0	.45	.25
7							1.0	1.30
8								1.0

TABLE 2 Optimum Retained Stations and the Optimum Shannon's Information Transmission

Bivariate		Multivariate	
Optimum Retained Stations	Optimum Information	Optimum Retained Stations	Optimum Information
5	3.536		
3,5	5.312	4,5	3.629
3,4,5	6.972	4,5,8	4.525

TABLE 3 Optimum Retained Stations and the Optimum Fisher's Information

Bivariate		Multivariate	
Optimum Retained Stations	Optimum Information	Optimum Retained Stations	Optimum Information
3	5.366		
3,6	5.813	4,6	5.853
3,6,8	6.222	3,6,8	6.310

TABLE 4 Comparison of Shannon and Fisher Methodologies using Estimation Error Method: Single Station Case

Grid	Station	Error Variance	Methodologies
1	5	0.885	Shannon (Biv.)
1	3	0.919	Fisher (Biv.)
2	5	0.835	Shannon (Biv.)
2	3	0.907	Fisher (Biv.)
8	5	0.887	Shannon (Biv.)
8	3	0.906	Fisher (Biv.)

TABLE 5 Comparison of Shannon and Fisher Methodologies using Estimation Error Method: Two Station Case

Grid	Station	Error Variance	Methodologies
1	4,5	0.895	Shannon (Mult.)
1	3,5	0.908	Shannon (Biv.)
1	3,6	0.917	Fisher (Biv.)
1	4,6	0.895	Fisher (Mult.)
2	4,5	0.814	Shannon (Mult.)
2	3,5	0.906	Shannon (Biv.)
2	3,6	0.909	Fisher (Biv.)
2	4,6	0.819	Fisher (Mult.)
8	4,5	0.888	Shannon (Mult.)
8	3,5	0.898	Shannon (Biv.)
8	3,6	0.901	Fisher (Biv.)
8	4,6	0.870	Fisher (Mult.)

CONCLUSION

Although the hydrologic network system is a fundamental tool in the design, development, and operation of water resources projects but no universal method of its design has yet emerged. The existing methods are applicable to very specific problems and depend upon very restrictive assumptions.

A methodology based on Shannon's information theory is proposed which does provide a universal dimensionless network performance measure. It provides a simple criterion which treats the network as an integrated system. Unlike Fisher's information measure to hydrologic network design, the proposed methodology is not restricted to linearity and normality assumptions.

Based on the simulated example and the tabulated results, the following conclusions are drawn:

1. When estimating on the basis of single station data, estimation error variance using Shannon's bivariate information criterion, was less than that obtained with the comparable Fisher's criterion (Table 4).

2. When estimating on the basis of data from two stations, Shannon's bivariate and multivariate information transmission criteria in most cases outperformed Fisher's criteria (Table 5).

REFERENCES

Benson, M.A. (1972). Use of multiple regression analysis in the design of a stream gaging networks - Practice in the U.S.A. Casebook on Hydrometeorological Network Design Practice: WMO No. 324, Geneva, Switzerland, III-3.2.1 - 3.2.4.

Benson, M.A. (1965). Allocation of stream-gaging stations within a country. WMO/IASH Symp. on the Design of Hydrologic Network, Quebec, IASH Pub. No. 67, 222-228.

Brass, R.L., and Rodriguez Iturbe, I. (1976). Rainfall network design for runoff prediction. Water Resources Research, 12, 6, 1197-1208.

Carrigan, P.H., and Golden, H.G. (1975). Optimizing information transfer in a stream gaging network. Water Resources Investigation 30-75, U.S. Geological Survey, 25 p.

Davis, D.R., Kisiel, C.C., and Duckstein, L. (1972). Optimum design of mountainous raingage networks using Bayesian decision theory. Int. Symp. on Distribution of Precipitation in Mountainous Area, Norway.

Dawdy, D.R., Kubik, H.E., and Close, E.R. (1970). Value of streamflow data for project design - A pilot study. Water Resources Research, 6, 4, 1045-1050.

Desi, F., Czelani, R., and Rackoczi, F. (1965). On determining the rational density of precipitation measuring networks. Proc. WMO/IASH Symp. on the Design of Hydrological Networks, Quebec, IASH Pub. No. 67, 127-129.

Drozdov, O.A. (1936). A method for setting up a network of meteorological station for a level region. Trudy GGO, 12, 3.

Fiering, M.B. (1965a). An optimization scheme for gaging. Water Resources Research, 1, 4, 463-470.

Fiering, M.B. (1965b). Use of correlation to improve estimates of the mean and variance. U.S. Geological Survey, Prof. Paper 434-C, 9 pp.

Ganguli, M.K., Rangarajan, R., and Panchang, G.M. (1951). Accuracy of rainfall estimates - Data of Damodar catchment. Irrigation and Power Journal, 8, 278-284.

Hershfield, D.M. (1965). On the spacing of raingages. Proc. WMO/IASH Symp. on the Design of Hydrological Networks, Quebec, IASH Pub. No. 67, 72-79.

Horton, R.E. (1923). The accuracy of areal rainfall estimates. Monthly Weather Review, 348-353.

Husain, T. (1979). Shannon's information theory in hydrologic network design and estimation. Ph.D. Thesis, Department of Civil Engineering, Univ. of B.C., Canada.

Jacobi, S. (1975). Economic optimum record length. Nordic Hydrology, 6, 28-42.

Langbein, W.B. (1954). Streamgaging network. IASH Pub. No. 38, 293-303.

Maddock, T. (1977). An optimum reduction of gauges to meet data program constraints. Bull. Int. Assoc. for Hydrologic Sciences, 19, 3, 337-345.

Matalas, N.C. (1973). Optimum gaging station locations. IBM Scientific Computing Symp. on Water and Air Resource Management, Yorktown Heights, N.Y., 85-94.

Matalas, N.C., and Gilory, E.J. (1968). Some comments on the regionalization in hydrologic series. Water Resources Research, 4, 6, 1361-1369.

Moss, M.E. (1972). Expected economic record length as basis for hydrologic network design. Proc. Int. Symp. on Water Resources Planning, Mexico City.

Moss, M.E., and Karlinger, M.R. (1974). Surface water network design by regression analysis and simulation. Water Resources Research, 10, 3, 427-433.

Moss, M.E., and Dawdy, D.R. (1973). The worth of data in hydrologic design. Highway Research Record, 479.

Shannon, C.E., and Weaver, W. (1949). The mathematical theory of communication. The University of Illinois Press, Urban, Illinois.

Solomon, S.I. (1972). Multi-regionalization and network stategy. Casebook on Hydrometeorological Network Design Practice: WMO No. 324, pp. III-3.3-1 to 3.3-11.

Solomon, S.I., Denouvilliez, J.P., Chart, E.J., Woolley, J.A., and Cadou, C. (1968). The use of square grid system for computer estimation of precipitation, temperature, and runoff. Water Resources Research, 4, 5, 919-929.

Stol, Ph. Th. (1972). The relative efficiency of the density of raingage networks. J. Hydrology, 15, 193-208.

741

Uryvaeu, V.A. (1965). Basic principles governing the design of a hydrological network. Proc. WMO/IASH Symp. on the Design of Hydrological Networks, Quebec, IASH Pub. No. 67, 199-206.

Wallis, J.R., and Matalas, N.C. (1972). Information transfer via regression in markovian world. RC 4207, IBM Research Centre, Yorktown Heights, N.Y.