Journal of Hydrology, 95 (1987) 323–339
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

[1]

COMBINED HYDROLOGIC SAMPLING CRITERIA FOR RAINFALL AND STREAMFLOW

DAVID G. TARBOTON, RAFAEL L. BRAS and CARLOS E. PUENTE*

Ralph M. Parsons Laboratory, Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139 (U.S.A.)

(Received August 23, 1986; revised and accepted April 14, 1987)

ABSTRACT

Tarboton, D.G., Bras, R.L. and Puente, C.E., 1987. Combined hydrologic sampling criteria for rainfall and streamflow. J. Hydrol., 95: 323–339.

This paper considers the joint sampling of the rainfall and streamflow processes. The sampling frequencies in time and space are obtained as a function of basin and rainfall characteristics. The effectiveness of different sampling strategies is measured by the variance of the error of estimated or predicted streamflow. This is related to the rainfall and basin rainfall-discharge properties through parameterizations of these processes. Rainfall is modelled as a stochastic process with covariance structure separable in time and space. Streamflow is parameterized in terms of the fluvial geomorphology of the basin. Linear systems theory is used to link precipitation to flow and to compute the variance of basin discharge. The variance of the error in prediction of streamflow is computed in terms of the following: (1) basin and rainfall model parameters; and (2) measurement strategy consisting of numbers of rain gages plus rainfall and flow measurement intervals. This error variance is used to assess the effectiveness of a measurement strategy. The results should be of use in the formulation of hydrologic sampling strategies.

INTRODUCTION

The inherent variability of hydrologic processes leads to questions like: "How much information is enough?" and "What kind of data are needed?" The answer to such questions depends on the particular objectives being pursued; this is why it is so difficult to provide general guidelines for the design of data-collection programs. This paper describes a study of the sampling of rainfall and streamflow in an interrelated fashion. Linear systems theory is used to link precipitation with streamflow. The variance of streamflow prediction error, as obtained from the linear representation, is used to assess the effectiveness of rainfall and streamflow sampling schemes.

The next section gives parameterizations of the rainfall process and of the basin response function. A state–space formulation, for providing the minimum variance linear estimate of flow, given rainfall and streamflow measurements,

* Present address: Department of Land, Air and Water Resources, University of California, Davis, CA 95616 (U.S.A.)

0022-1694/87/$03.50

is developed next. This state–space formulation is then exploited to solve the sampling problem under different rainfall and basin characteristics.

PARAMETERIZATION OF HYDROLOGIC PROCESSES

Rainfall

A particular model of rainfall is not assumed. Instead the only assumption is that the rainfall process has a covariance structure which is separable in time and space, with components stationary in time and homogeneous in space. Such a model has been used by Rodriguez-Iturbe and Mejia (1974), Bras and Rodriguez-Iturbe (1975) and Bras and Colon (1978). This assumption implies that the results will only be valid for areas small enough for spatial homogeneity to hold and at time scales shorter than where seasonality affects the stationarity, but longer than where non-stationarity within a single storm plays a role. We feel that this assumption is valid as a first approximation when designing sampling networks for completely ungaged catchments up to a size of about 100 × 100 km. Mathematically the separable covariance is:

$$\mathrm{Cov}[\zeta(t_1, z_1), \zeta(t_2, z_2)] = \sigma_p^2 C_T(|t_1 - t_2|) C_s(|z_1 - z_2|) \quad (1)$$

where $\zeta(t, z)$ is the precipitation rate at time t and spatial location z, with σ_p^2 the variance of precipitation at a point, and the functions C_T and C_s denoting respectively the covariance functions in time and space.

In particular, the following forms for such functions are assumed:

$$C_T(t) = \rho^t \quad (2)$$

and:

$$C_s(r) = \mathrm{e}^{-hr} \quad (3)$$

where ρ and h are the respective covariance parameters, and r denotes distance in space.

We are primarily interested in the area averaged precipitation as this is the input required for lumped rainfall–runoff process models. This is:

$$p(t) = \frac{1}{A} \int_A \zeta(t, z) \mathrm{d}z \quad (4)$$

where A is the area of the basin.

Using the covariance structure of the rainfall intensity process, eqn. (1), we can obtain the covariance of the area averaged rainfall:

$$\mathrm{Cov}[p(t_1), p(t_2)] = \sigma_p^2 C_T(|t_1 - t_2|) \frac{1}{A^2} \int_A \int C_s(|z_1 - z_2|) \mathrm{d}z_1 \mathrm{d}z_2 \quad (5)$$

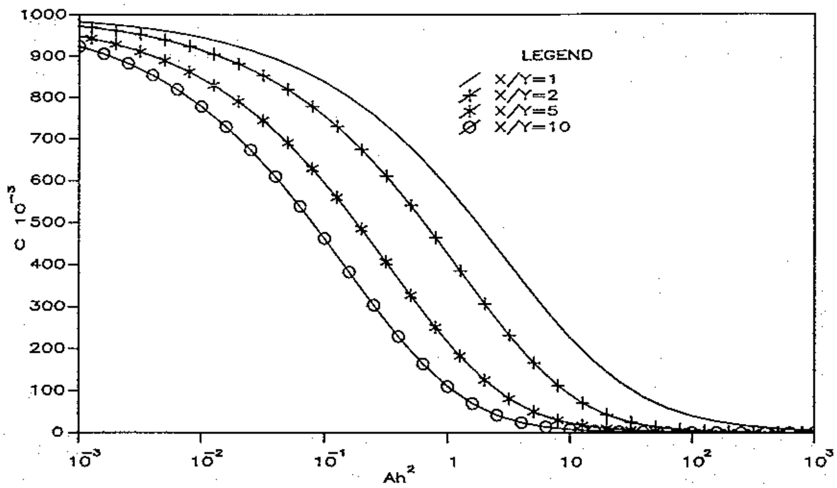

Fig. 1. Area reduction factor for a rectangular region with exponential correlation.

If the spatially averaged covariance is denoted:

$$C_1 = \frac{1}{A^2}\int_A\int_A C_s(|z_1 - z_2|)\mathrm{d}z_1\,\mathrm{d}z_2 \tag{6}$$

then the covariance (5) becomes:

$$\mathrm{Cov}[p(t_1), p(t_2)] = C_1\sigma_p^2 C_T(|t_1 - t_2|) \tag{7}$$

C_1 is an area reduction factor since it relates the point variance to the area averaged autocovariance. Such reduction factor for a rectangular region and with $C_s(\cdot)$ given by eqn. (3), is given in Fig. 1. In this figure x/y is the aspect ratio, or length to width ratio of the region.

In practice it is not possible to directly measure the area averaged precipitation process (4). Rather, it is approximated from point rain gage measurements as follows:

$$\hat{p}(t) = \sum_{i=1}^{N} \gamma_i \zeta(t, z_i) \tag{8}$$

with the γ_i's representing weights corresponding to N gages with location vectors z_i. For simplicity it will be assumed that the gages are randomly located, so that $\gamma_i = 1/N$ is used for all the weights. Positioning of the rain gages and optimization on the weights so that $\hat{p}(t)$ better approximates $p(t)$ can be included in the assumed methodology but will not be considered here. The works of Rodriguez-Iturbe and Mejia (1974), Lenton and Rodriguez-Iturbe (1974) and Bras and Rodriguez-Iturbe (1975) should be used in such cases.

Using the assumed separable covariance structure of the rainfall intensity process leads to the autocovariance of mean area $\hat{p}(t)$ rainfall, as obtained from rain gages:

$$\mathrm{Cov}[\hat{p}(t_1), \hat{p}(t_2)] = \sigma_p^2 C_T(|t_1 - t_2|) \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} C_s(|z_i - z_j|) \tag{9}$$

Define:

$$C_2 = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} C_s(|z_i - z_j|) \tag{10}$$

then:

$$\mathrm{Cov}[\hat{p}(t_1), \hat{p}(t_2)] = C_2 \sigma_p^2 C_T(|t_1 - t_2|) \tag{11}$$

Notice that C_2 assumes specific locations for the rain gages, i.e., z_i. Evaluation of the effect of random placement of gages requires the elimination of such specific dependence on location. This is accomplished by computing the expected value of C_2 using the joint distribution of the location of any pair of rain gages, i.e., $f(z_i, z_j) = 1/A^2$.

The sum C_2 can be expanded as follows:

$$C_2 = \frac{1}{N^2} \left\{ N + 2 \sum_{i=2}^{N} \sum_{j=1}^{i-1} C_s(|z_i - z_j|) \right\} \tag{12}$$

and its expected value becomes:

$$E(C_2) = \frac{1}{N^2} \left\{ N + 2 \sum_{i=2}^{N} \sum_{j=1}^{i-1} \int_A \int_A \frac{1}{A^2} C_s(|z_i - z_j|) \mathrm{d}z_i \mathrm{d}z_j \right\} \tag{13}$$

which using eqn. (6) results in:

$$E(C_2) = \frac{1}{N^2} [N + N(N-1)C_1] = C_1 + \left[\frac{1 - C_1}{N}\right] \tag{14}$$

Now we are in a position to compute the covariance of the error in rainfall due to random discrete sampling in space. Define the error in rainfall as:

$$e(t) = \hat{p}(t) - p(t) \tag{15}$$

Using the definitions of the terms involved and the assumed covariance structure, eqns. (8), (4) and (1) respectively, we can obtain:

$$\mathrm{Cov}\{e(t_1), e(t_2)\} = \sigma_p^2 C_T(|t_1 - t_2|) \left[\sum_{i=1}^{N} \sum_{j=1}^{N} \frac{1}{N^2} C_s(|z_i - z_j|) - 2 \sum_{i=1}^{N} \frac{1}{AN} \int_A C_s(|z_i - z|) \mathrm{d}z + \frac{1}{A^2} \int_A \int_A C_s(|z_1 - z_2|) \mathrm{d}z_1 \mathrm{d}z_2 \right] \tag{16}$$

For randomly positioned gages, taking expected values of the terms inside square brackets we get:

$$\mathrm{Cov}\{e(t_1), e(t_2)\} = \sigma_p^2 C_T(|t_1 - t_2|)\left[\frac{1 - C_1}{N}\right] \tag{17}$$

It is worth emphasizing that if a particular gage configuration is known, the sums in eqns. (12) and (16) can be explicitly evaluated to give results of similar form to those obtained under the randomly positioned gages assumption, i.e., eqns. (14) and (17).

Basin response

The basin response is parameterized in terms of its instantaneous unit hydrograph. This gives runoff via the convolution:

$$q(t) = \int_0^t h(\tau)p(t - \tau)\mathrm{d}\tau \tag{18}$$

where $p(t)$ is the area averaged rainfall intensity, $h(t)$ the basin's transfer function or instantaneous unit hydrograph (IUH), and $q(t)$ the discharge at time t.

The parameterization of the instantaneous unit hydrograph used here is that of Rodriguez-Iturbe and Valdes (1979). The IUH is interpreted as the probability density function of the travel time that a unit volume of effective rainfall, which lands randomly anywhere in the basin, takes to reach the basin outlet. The IUH was obtained by computing such probability density function, assuming the basin behaves as a continuous-time discrete-state Markov process. The states of the Markov chain are defined according to the order of the channel in which the unit volume is at a particular time in its journey to the basin outlet. The time which the unit volume spends in a particular channel was approximated by an exponential probability distribution, with a parameter inversely proportional to the typical length of streams of a given order. The IUH peak and time-to-peak given by Rodriguez-Iturbe and Valdes (1979) (rearranged for length and velocity expressed in compatible units) are:

$$q_p = 0.364\, R_L^{0.43}\left(\frac{L}{v}\right)^{-1} \tag{19}$$

$$t_p = 1.584\,(R_B/R_A)^{0.55}\, R_L^{-0.38}\left(\frac{L}{v}\right) \tag{20}$$

where R_B, R_A, and R_L are the Horton numbers: *bifurcation* ratio, *area* ratio, and *length* ratio, respectively; L is the length of the highest order stream; and v is the peak velocity on the streams, assumed constant over the basin.

A good approximation of this geomorphologic instantaneous unit hydrograph (GIUH) is given by Rosso (1984) in terms of the two parameter gamma probability density function. This is:

$$h(t) = [k\Gamma(m)]^{-1} (t/k)^{m-1} \exp(-t/k) \tag{21}$$

where k is a scale parameter, m is a shape parameter, and $\Gamma(\ )$ is the gamma function.

Such gamma model was introduced into hydrology by Nash (1957) as the result of a cascade of linear reservoirs, with m representing the number of reservoirs in cascade and k the discharge constant of each reservoir, i.e., storage equal k times discharge. Rosso (1984) gives the parameters of the gamma model in terms of geomorphologic parameters as:

$$m = 3.29 (R_B/R_A)^{0.78} R_L^{0.07} \tag{22}$$

$$k = 0.70 [R_A/(R_B R_L)]^{0.48} \left(\frac{L}{v}\right) \tag{23}$$

This representation is used here as a convenient, mathematically tractable form of the GIUH.

The precipitation term appearing in the convolution given by eqn. 18 should really be the runoff produced after infiltration and other reductions to the gross rainfall. There are many and varied runoff production mechanisms that can operate individually or simultaneously in a basin. In essence the runoff production represents another, probably nonlinear, transfer function altering the properties of the rainfall input. We have yet to find an acceptable way to parameterize what amounts to a spatially varying, random, transfer function. To the extent that this function alters the covariance structure of the rainfall our results will be affected. Although we acknowledge this serious limitation, and continue working on its resolution, we feel that ignoring those effects at this time still provides useful results. If the goal of the network is to gain knowledge on discharge the added uncertainty in the runoff producing mechanisms will favor increased flow sampling. The opposite would be true if interest is on increased knowledge of the rainfall process. Our results, we feel, still provide reasonable guidelines for sampling which are most probably underestimating some of the streamflow sampling needs.

STATE–SPACE FORMULATION

The covariance function of area averaged precipitation when the time component is assumed exponential, eqns. (7) and (2), is:

$$\mathrm{Cov}[p(t_1), p(t_2)] = C_1 \sigma_p^2 \rho^{|t_1 - t_2|} = C_1 \sigma_p^2 e^{\ln\rho |t_1 - t_2|} \tag{24}$$

This is the covariance function of a Markov process (Gelb, 1984) and can be represented by the stochastic differential equation:

$$\frac{\mathrm{d}}{\mathrm{d}t} p(t) = \ln\rho\, p(t) + w \tag{25}$$

where w is zero-mean white noise with spectral density $2C_1\sigma_p^2 \ln(1/\rho)$. Similarly the rainfall error process due to discrete rain gages, (eqn. 15), can be represented as:

$$\frac{d}{dt}e(t) = \ln\rho\, e(t) + w' \tag{26}$$

where w' is zero-mean white noise with spectral density $2[(1 - C_1)/N]\sigma_p^2\ln(1/\rho)$.

Rainfall observations are generally cumulative over the observation interval. This is represented mathematically as an integral of the observed process, over the interval Δt:

$$\int_{t-\Delta t}^{t} \hat{p}(t)\mathrm{d}t = \int_{t-\Delta t}^{t} [p(t) + e(t)]\mathrm{d}t \tag{27}$$

Define:

$$l(t) = \int_{0}^{t} [p(t) + e(t)]\mathrm{d}t \tag{28}$$

Then rainfall observations can be represented as:

$$l(t) - l(t - \Delta t) \tag{29}$$

The differential equivalent of (28) is:

$$\frac{d}{dt}l(t) = e(t) + p(t) \tag{30}$$

The Nash model parameterization of the basin is convenient in that it allows descriptions of the basin response through a finite number of differential equations:

$$\left.\begin{aligned}
\frac{dq_1}{dt} &= -\frac{1}{k}q_1 + \frac{1}{k}q_0\\
\frac{dq_2}{dt} &= \frac{1}{k}q_1 - \frac{1}{k}q_2\\
&\vdots\\
\frac{dq_m}{dt} &= \frac{1}{k}q_{m-1} - \frac{1}{k}q_m
\end{aligned}\right| \tag{31}$$

where q_i is flow from the ith Nash reservoir, and q_0 are the inflows (effective rainfall) to the basin model.

The state–space model of combined rainfall and runoff is formed by combining eqns. (25), (26), (30), and (31) into a single vector stochastic differential equation in which the inflow to the basin component is taken as the true area averaged precipitation $p(t)$:

$$\frac{d}{dt}\boldsymbol{x} = \boldsymbol{F}\boldsymbol{x} + \boldsymbol{w} \tag{32}$$

where $\boldsymbol{x}$ is the state vector:

$$\boldsymbol{x} = [e(t)l(t)p(t)q_1(t)...q_m(t)] \tag{33}$$

$$\boldsymbol{F} = \begin{bmatrix} \ln\rho & 0 & 0 & \dots & \dots & & \dots & 0 \\ 1 & 0 & 1 & 0 & \dots & & \dots & 0 \\ 0 & 0 & \ln\rho & 0 & \dots & & \dots & 0 \\ 0 & 0 & \frac{1}{k} & -\frac{1}{k} & 0 & & \dots & 0 \\ 0 & 0 & 0 & \frac{1}{k} & -\frac{1}{k} & 0 & \dots & 0 \\ \vdots & & & & & & & \vdots \\ 0 & \dots & \dots & \dots & 0 & & \frac{1}{k} & -\frac{1}{k} \end{bmatrix} \tag{34}$$

$\boldsymbol{w}$ is a white noise vector with constant in time spectral density matrix, $\boldsymbol{Q}$, as follows:

$$\boldsymbol{Q} = 2\sigma_p^2 \ln\frac{1}{\rho} \begin{bmatrix} \frac{1-C_1}{N} & 0 & \dots & & 0 \\ 0 & 0 & & & \\ \vdots & & C_1 & & \vdots \\ & & & 0 & \\ & & & & \ddots \\ 0 & & \dots & & 0 \end{bmatrix} \tag{35}$$

This completes the development of the continuous time state-space formulation. To practically implement such formulation for discrete measurements, the equivalent discrete form is used:

$$\boldsymbol{x}(t) = \boldsymbol{\phi}(t, t_0)\boldsymbol{x}(t_0) + \boldsymbol{w}(t, t_0) \tag{36}$$

where $\boldsymbol{\phi}(t, t_0)$ is the transition matrix, from time t_0 to time t, which satisfies the differential equation:

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\phi}(t, t_0) = \boldsymbol{F}\boldsymbol{\phi}(t, t_0) \tag{37}$$

with initial conditions $\phi(t_0, t_0) = \boldsymbol{I}$. $\boldsymbol{w}(t, t_0)$ is zero-mean white noise with covariance related to the continuous spectral density by:

$$\boldsymbol{Q}(t, t_0) = \int_{t_0}^{t} \boldsymbol{\phi}(t, \tau)\boldsymbol{Q}\boldsymbol{\phi}(t, \tau)^T \mathrm{d}\tau \tag{38}$$

Equations (37) and (38) are solved numerically.

The discrete formulation is completed by giving the discrete measurements. Rainfall measurements, at intervals Δt, are modelled as:

$$z_1(t) = H_1 \boldsymbol{x}(t) + J\boldsymbol{x}(t - \Delta t) + v_1(\Delta t) \tag{39}$$

where $H_1 = [0\ 1\ 0 \ldots 0]$, $J = [0\ -1\ 0 \ldots 0]$, and $v_1(\Delta t)$ is a zero-mean measurement error with variance $R_1(\Delta t)$. R_1 depends on Δt is because with greater accumulation of rainfall a greater error is expected.

Flow measurements are modelled as:

$$z_2(t) = H_2 \boldsymbol{x}(t) + v_2 \tag{40}$$

where $H_2 = [0\ 0 \ldots 1]$, and v_2 is the zero-mean flow measurement error with variance R_2.

Linear filtering techniques can be applied to find each states minimum variance estimate and the corresponding error covariance matrix. Such error covariance matrix can be computed without knowledge of the actual observations. The variance of the error on runoff (the last diagonal element of the error covariance matrix) will be used to measure the effectiveness of a sampling strategy. The procedure employed for the propagation and updating of state and error covariance estimates is described next. It is based on the Kalman filter solution of a system with integrated measurements as given by Brown (1983). The solution presented here generalizes that of Brown (1983), in that an instantaneous runoff measurement can occur anywhere independently on rainfall sampling. Details of the derivation can be obtained in Tarboton (1987).

Suppose that at time t_0 we have an estimate of the state $\hat{\boldsymbol{x}}(t_0)$ which has estimation error covariance $\Sigma(t_0)$. Then using the discrete model described above we get an estimate of state at time t:

$$\hat{\boldsymbol{x}}(t) = \boldsymbol{\phi}(t, t_0)\hat{\boldsymbol{x}}(t_0) \tag{41}$$

which has error covariance:

$$\Sigma(t) = \boldsymbol{\phi}(t, t_0)\Sigma(t_0)\boldsymbol{\phi}(t, t_0)^T + \boldsymbol{Q}(t, t_0) \tag{42}$$

Equation (42) is used to propagate the covariance of estimation error at times between rainfall or runoff measurements.

The estimate may be updated, due to a flow measurement, eqn. (40), as follows:

$$\hat{\boldsymbol{x}}(t+) = \hat{\boldsymbol{x}}(t-) + \boldsymbol{K}[z(t) - H_2\hat{\boldsymbol{x}}(t-)] \tag{43}$$

where $t-$ and $t+$ indicate respectively the state estimates just before and after the update for the measurement at time t. $\boldsymbol{K}$ is the Kalman gain matrix which provides the minimum variance linear estimate of the states, it is:

$$\boldsymbol{K} = \Sigma(t-)H_2^T[H_2\Sigma(t-)H_2^T + R_2]^{-1} \tag{44}$$

The covariance of estimation error after the updating of runoff becomes:

$$\Sigma(t+) = (\boldsymbol{I} - \boldsymbol{K}H_2)\Sigma(t-) \tag{45}$$

In the problem at hand, rainfall measurements are cumulative values, not instantaneous measurements. Several streamflow measurements may also exist during the time rainfall is being accumulated. The traditional Kalman filter cannot handle this situation in the updating step when a rainfall total becomes available. Tarboton (1987) derives modified filter expressions to permit the handling of cumulative rainfall measurements. The state update is:

$$\hat{x}(t+) = \hat{x}(t-) + M[z_1(t) - H_1\hat{x}(t-) - J\hat{x}_s(t-\Delta t)] \tag{46}$$

where $x_s(t-\Delta t)$ is the smoothed estimate of the state at time $t-\Delta t$ due to all measurements up to, but not including, those at t. The gain matrix M is given by:

$$M = (\Sigma(t-)H_1^T + \phi'(t, t-\Delta t)\Sigma(t-\Delta t)J^T)L^{-1} \tag{47}$$

where:

$$\begin{aligned} L = {} & H_1\Sigma(t-)H_1^T + R_1 + J\Sigma_s(t-\Delta t)J^T + H_1^T\phi'(t, t-\Delta t)\Sigma(t-\Delta t)J^T \\ & + J\Sigma(t-\Delta t)\phi'(t, t-\Delta t)^T H_1^T \end{aligned} \tag{48}$$

with $\Sigma_s(t-\Delta t)$ denoting the error covariance matrix of the smoothed state estimate at time $t-\Delta t$; and $\phi'(t, t-\Delta t)$ representing the effective generalized transition matrix of the system when flow updates (at times $t_1, \ldots, t_{n-1}$) are made during the interval of accumulation of rainfall:

$$\begin{aligned} \phi'(t, t-\Delta t) = {} & \phi(t, t_{n-1})(I - K_{n-1}H_2)\phi(t_{n-1}, t_{n-2})(I - K_{n-2}H_2)\ldots \\ & (I - K_1H_2)\phi(t_1, t-\Delta t) \end{aligned} \tag{49}$$

The error covariance update after rainfall measurement is:

$$\Sigma(t+) = \Sigma(t-) - MLM^T \tag{50}$$

Now the theoretical development is complete. State predictions are given by eqn. (41) and the prediction error covariance by eqn. (42). Updates of these estimates due to rainfall or runoff measurements are achieved using eqns. (43)–(50). Before giving results it is instructive to see how the basin parameters affect the streamflow variance. With the rainfall covariance, eqn. (7), and basin transfer function, eqn. (21), taking the variance of eqn. (18) gives the streamflow variance. This is most effectively achieved in the frequency domain (Bras and Rodriguez-Iturbe, 1985). Here we numerically integrated the flow spectral density to obtain the results presented in Fig. 2.

As a bound on how good we can possibly expect our predictions to be, we can assume our initial error variance matrix $\Sigma(0) = 0$. This implies we know all states perfectly at time $t = 0$. Equation (42) is used to propagate the error variance from this starting condition. Figure 3 shows how the flow [state $q_m(t)$] error variance propagates with time for different basin parameters k and m. The flow error variance has been normalized by the flow process variance

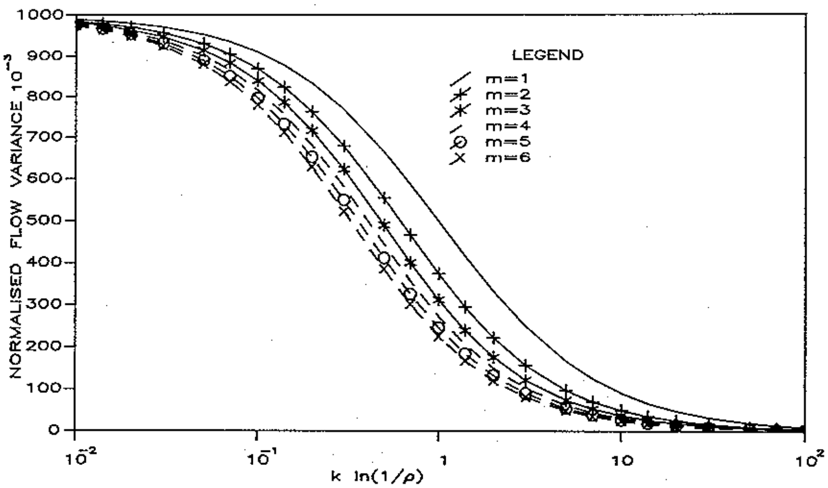

Fig. 2. Flow variance from basin parameters.

which is known (Fig. 2). Time has been normalized by the factor ln $(1/\rho)$. Notice that the flow error variance quickly reaches an asymptotic steady state, which is the flow process variance.

If sampling of rainfall and discharge in time is done periodically (i.e., at regular intervals), a similar asymptotic state will be reached but not monotonically as in Fig. 3. The observed cyclic behavior, due to periodic sampling is illustrated in Fig. 4. The peaks in Fig. 4 give the maximum uncertainty in our estimate with sampling and occur just prior to a flow measurement. The predicted flow variance in our estimate of flow for a given lead time is obtained

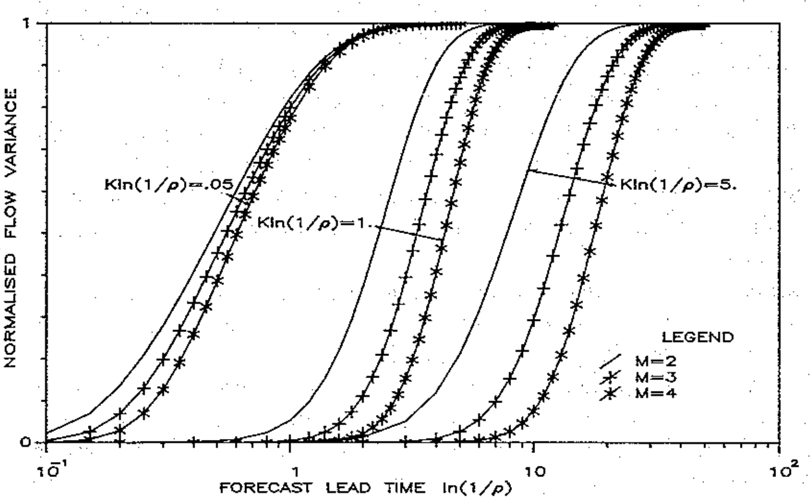

Fig. 3. Stationary rainfall model. Normalized error variance based on perfect knowledge.

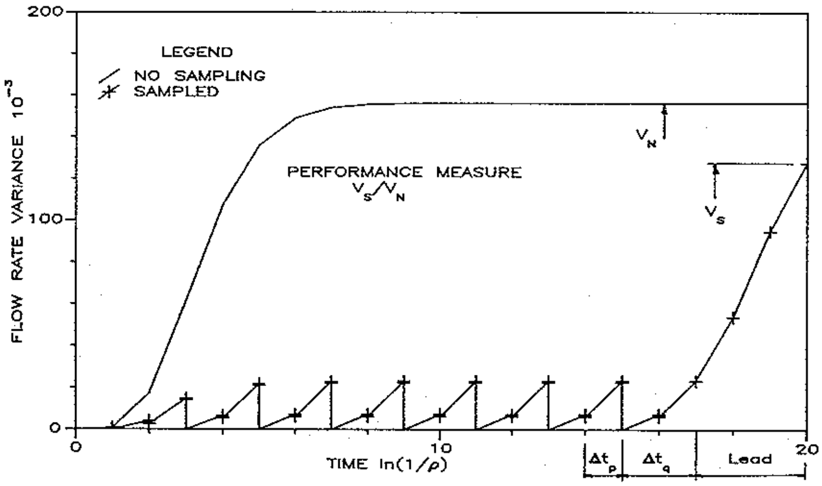

Fig. 4. Definition of flow rate variance criterion.

by propagating that obtained from updates prior to the time from which forecasts are made, as shown on the right of Fig. 4. The maximum predicted error variance V_s, compared to the process variance V_n is used as a measure of the effectiveness of the sampling strategy. In Fig. 4 and in the remainder of this paper, the measurement noise variances [R_1 and R_2, eqns. (39) and (40)] were taken as zero. This was done for simplicity and so as not to confound effects of error due to insufficient measurements and errors in the measurements.

SAMPLING DESIGN STRATEGY

Thus far many different parameters and variables have been introduced. All have an effect on the error variance of our estimate of discharge. Here we simplify the complex interrelations between such variables by normalizing them with respect to rainfall parameters. In this way, quasi-general results will be presented which show the trade-offs between rainfall and discharge sampling to obtain a desired degree of accuracy. The results are quasi-general in that through normalization the number of parameters has been reduced from five to two, and full results are given only for typical sets of the two remaining normalized parameters.

The five parameters which describe basin response and climate (rainfall) defined previously are k, m, ρ, Ah^2, σ_p^2. First, time is normalized by measuring it in terms of the rainfall correlation time $\ln 1/\rho$. Then discharge error variance is normalized by the known process variance. This has already been done in Fig. 3 and is the ratio V_s/V_N in Fig. 4. The choice of a sampling strategy involves defining the rainfall and runoff sampling intervals Δt_p and Δt_q and the number of gages N. Note that the number of gages only affects the normalized error variance of estimation by the relative magnitude of terms in the noise spectral density matrix eqn. (35). Define this effect through the ratio:

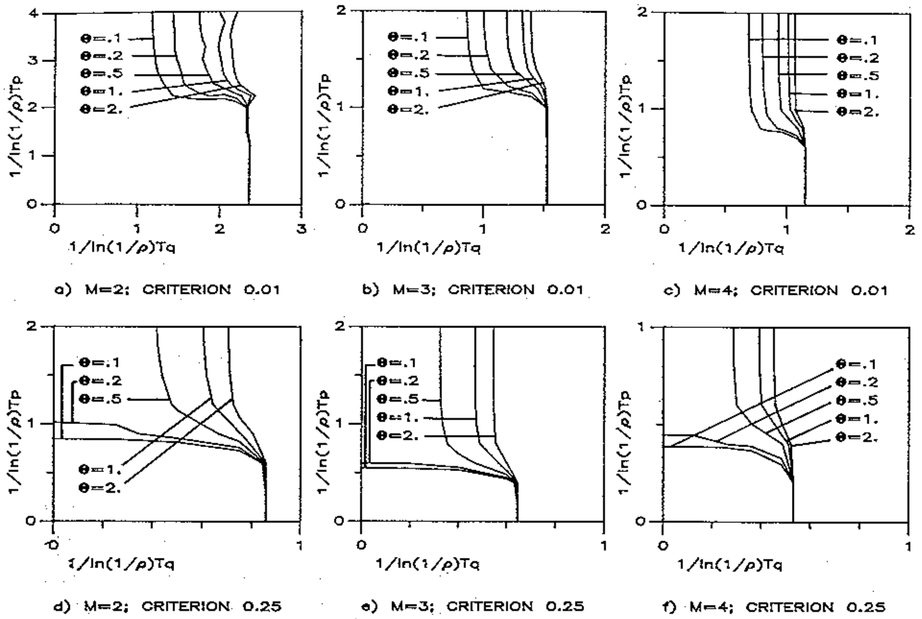

Fig. 5. Sampling strategy selection with $k \ln(1/\rho) = 1$ and 0 forecast lead time.

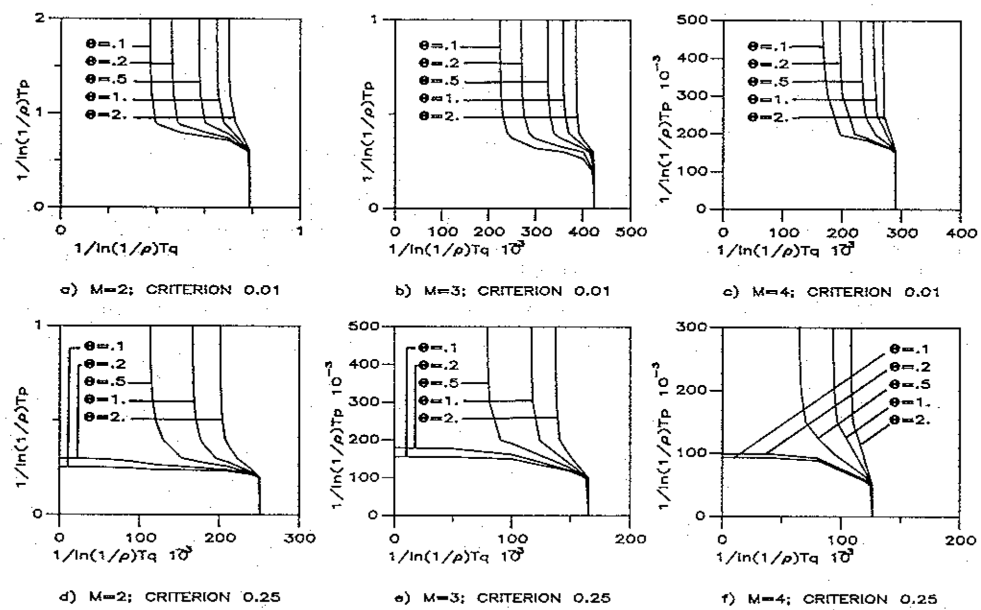

Fig. 6. Sampling strategy selection with $k \ln(1/\rho) = 5$ and 0 forecast lead time.

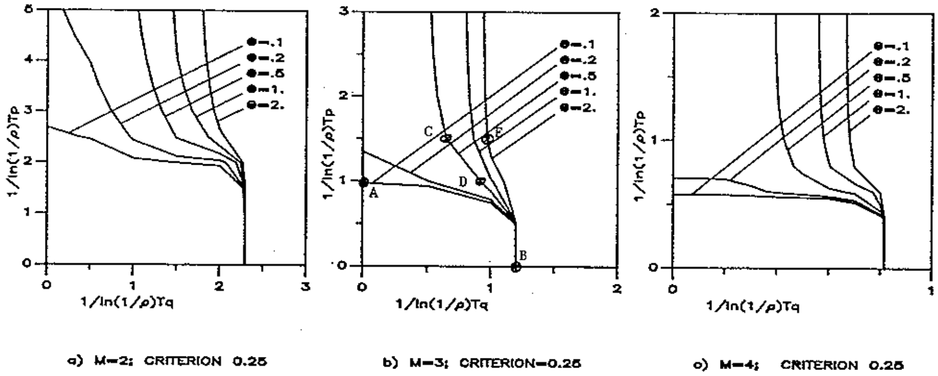

Fig. 7. Sampling strategy selection with $k \ln(1/\rho) = 5$ and $\ln(1/\varrho)$ forecast lead time.

$$\theta = \frac{1 - C_1}{C_1 N} \tag{51}$$

Since the effect of the exponential correlation structure in time, Ah^2, is also included in this ratio, we have reduced the sampling selection to three dimensionless design variables $\Delta t_p \ \ln 1/\rho$, $\Delta t_q \ln 1/\rho$, and θ, which will be obtained based on the two dimensionless parameters $k \ln 1/\rho$ and m.

A possible design criterion using this procedure may be the specification of acceptable normalized error variance at a specific forecast lead time. Figure 3

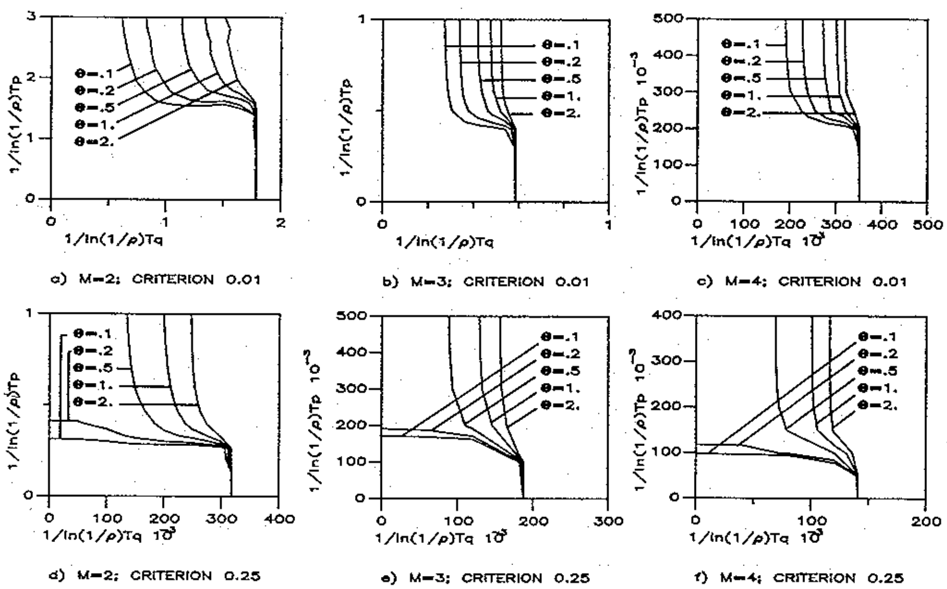

Fig. 8. Sampling strategy selection with $k \ln(1/\rho) = 5$ and $\ln(1/\rho)$ forecast lead time.

TABLE 1

Example of selection of sampling strategy for a 1.2 h forecast

Point	θ	Rainfall sampling frequency $\frac{1}{\ln\frac{1}{\rho}\Delta t_p}$	Discharge sampling frequency $\frac{1}{\ln\frac{1}{\rho}\Delta t_q}$	Rainfall sampling interval (h) Δt_p	Discharge sampling interval (h) Δt_q	Number of rain gages N
A	0.1	0.97	0	0.86	∞, none	$12.2 \simeq 12$
B	–	0	1.2	∞, none	0.70	–
C	0.5	1.45	0.65	0.57	1.3	$2.4 \simeq 2$
D	0.5	1	0.92	0.83	0.90	$2.4 \simeq 2$
E	2	1.45	0.97	0.57	0.86	$0.6 \simeq 1$

which gives the variance of predictions based on perfect knowledge should be considered when doing this so as not to form an unduly stringent criterion. Here criteria of 0.01 and 0.25 are used for normalized discharge error variance. These correspond to discharge estimates having standard deviations that are 10 and 50% of the process standard deviation. Figures 5–8 give results for typical values of the dimensionless parameters $k \ln 1/\rho$ and m. These show the sampling strategy required to meet the given design criterion at the specified forecast lead times.

Notice that with low θ ratio increasing rainfall sampling frequency markedly reduces the required flow sampling that is necessary to achieve the same accuracy. A tradeoff between rainfall and streamflow sampling is clearly present. For a high θ ratio ($\theta > 2$), which occurs with low spatial density of rain gages, the rainfall sampling has little effect as evidenced by the almost vertical line indicating that the same streamflow sampling is required to achieve the given accuracy no matter how rainfall is sampled in time. Observe also that there is a rainfall sampling frequency below which all the lines join into one vertical line. This means sampling rainfall below such merging point is ineffective and does not provide any information additional to that provided by the streamflow sampling.

Example

Consider designing a sampling network for a river basin with Horton numbers $R_B = 3.5$, $R_A = 4$ and $R_L = 2.5$, length of highest order stream $L = 7\,\text{km}$ and stream peak flow velocity estimated to be $v = 4\,\text{km}\,\text{h}^{-1}$. The basin can be approximated by a rectangle $15 \times 10\,\text{km}$. Using eqns. (22) and (23) we compute $m = 3.16$ and $k = 0.84\,\text{h}$. Assume rainfall parameters $\varrho = 0.3$ (at a 1 h lag) and $h = 0.1\,\text{km}^{-1}$. Based on $Ah^2 = 1.5$, Fig. 1 gives $C_1 = 0.45$. The normalized basin parameter is $k \ln 1/\rho = 1.01$, so Figs. 5 and 7 apply for strategy selection. We wish to use the $\ln 1/\rho = 1.2$ h forecast lead criterion so points A to E are obtained from Fig. 7(b). The strategies for these are listed in Table 1. These strategies all provide information with the same error variance, namely that the variance of a 1.2 h forecast is 0.25 that of process variance. Tradeoffs between the number of gages and frequencies of measurement are evident. The cheapest or most convenient strategy from a comparison like this should be implemented.

CONCLUSIONS

Parameterizations of the rainfall and of the basin response which are simple and allow the use of linear systems theory for analyzing the problem of combined rainfall and streamflow measurement have been given. To provide a minimum variance linear estimate of discharge from rainfall and streamflow measurements combined, a state space approach was developed. This approach has the advantages that it can be applied when rainfall is modelled by nonstationary process and can represent rainfall measurements realistically as integrals over time.

The results presented here were obtained using the rainfall-discharge state–space formulation relating the variance of estimation error of flow to the measurement of discharge (in time) and of rainfall (in time and space). The findings are parameterized in terms of the rainfall and basin characteristics. These results could be useful in the design of hydrologic measurement networks.

This work has limitations in that it assumes linearity in the basin response and, it is therefore dependent on the accuracy of the model and parameters. While the basin response parameters may be obtained from the basin geomorphology, the choice of climate (rainfall) parameters values is not easy. Another deficiency is that the effect of uncertainty in the infiltration has not been accounted for. We are investigating ways of doing this through incorporation of the contributing area concept into the state–space formulation.

ACKNOWLEDGEMENTS

This work is part of a continuing effort sponsored by the U.S. Geological Survey (Grant No. 14-08-0001-G1143) as authorized by Public Law 98-242. The authors want to acknowledge the help of Mrs. Elaine Healy and Ms. Carole Solomon in preparing the final document.

REFERENCES

Bras, R.L. and Colon, R.V. 1978. Time averaged areal mean of precipitation: estimation and network design. Water Resour. Res., 14(5):878–888.

Bras, R.L. and Rodriguez-Iturbe, I., 1975. Rainfall runoff as spatial stochastic processes: data collection and synthesis. Dep. Civ. Eng., MIT, Cambridge, Mass., Ralph M. Parsons Lab. Tech. Rep. No. 196, 382 pp.

Bras, R.L. and Rodriguez-Iturbe, I., 1985. *Random Functions and Hydrology*. Addison Wesley, Reading, Mass., 559 pp.

Brown, R.G., 1983. *Introduction to Random Signal Analysis and Kalman Filtering*. Wiley, New York, N.Y., 347 pp.

Gelb, A., 1984. *Applied Optimal Estimation*, MIT Press, Cambridge, Mass., 374 pp.

Lenton, R.L. and Rodriguez-Iturbe, I., 1974. On the collection, the analysis, and the synthesis of spatial rainfall data. Dep. Civ. Eng., MIT, Cambridge, Mass., Ralph M. Parsons Lab. Tech. Rep. No. 194, 218 pp.

Nash, J.E., 1957. The form of the instantaneous unit hydrograph. *IAHS AISH* Publ. 42: 114–118.

Rodriguez-Iturbe, I. and Mejia, J.M., 1974. The design of rainfall networks in time and space. Water Resour. Res., 10(4):713–728.

Rodriguez-Iturbe, I. and Valdes, J.B., 1979. The geomorphologic structure of hydrologic response. Water Resour. Res., 15(5):1409–1420.

Rosso, R., 1984. Nash model relation to Horton order ratios. Water Resour. Res., 20(7):914–920.

Tarboton, D.G., 1987. Hydrologic sampling: a Characterization in terms of rainfall and basin properties. M.S. Thesis Massachusetts Institute of Technology, Cambridge, Mass.