ELSEVIER

Journal of Hydrology 169 (1995) 25–35

Journal of Hydrology

[1]

Optimal parameter estimation for Muskingum routing with ungauged lateral inflow

M.M. Kshirsagar[a], Balaji Rajagopalan[b], Upmanu Lall[b,*]

[a]Central Water and Power Research Station, Pune 411024, India

[b]Utah Water Research Laboratory, Utah State University, Logan, UT 84322-8200, USA

Received 17 June 1994; revision accepted 11 November 1994

Abstract

A Muskingum flood routing model coupled with an impulse response rainfall–runoff model for ungauged lateral inflows is presented. The model is designed to estimate the runoff hydrograph at a downstream point given runoff hydrographs at one or more upstream locations, and estimated rainfall hyetographs for ungauged catchments. A constrained, non-linear (successive quadratic) programming algorithm is used to estimate the parameters of this model from historical data. These parameters include the Muskingum K and x, the attenuation coefficient and the flow velocity of the impulse response function for the ungauged catchment. Tests of the model with synthetic situations and with a data set for the Godavari river reach in India are presented.

1. Introduction

A classical operational hydrologic problem is the estimation of runoff from a drainage consequent to rainfall and the routing of runoff downstream through a channel. Lateral inflow hydrographs into the main channel of interest can be estimated and added in dynamically as the routing proceeds. The Muskingum method continues to be popular for flood routing. Its parameters K ('travel time') and x ('discharge weighting factor') are usually estimated from historical data using a simple graphical or a least-squares procedure (see Singh 1992, p. 680) or linear programming methods (Gill, 1978; Stephenson, 1979). Linear impulse response models (Harpin and Cluckie, 1981) have been proposed for modelling ungauged runoff response to rainfall. Here we present a Muskingum flood routing model

* Corresponding author.

SSDI 0022-1694(94)02670-X

with multiple reaches coupled with an impulse response rainfall–runoff model for ungauged lateral inflows. The model is intended to estimate the runoff hydrograph at a downstream point given runoff hydrographs at one or more upstream locations, and estimated rainfall hyetographs for ungauged tributaries. A constrained, non-linear programming algorithm is used to estimate the parameters of the model.

The routing model is formally introduced through a schematic in the next section. The optimization problem for parameter estimation is then formulated. Tests with selected synthetic and real data sets that show the utility of the model conclude the presentation.

2. The streamflow model

The model is introduced through the hypothetical drainage system shown in Fig. 1. In this figure, we have two river reaches AC and CB. A and B are the upstream (u/s) and downstream (d/s) gauging sites where historic data on discharge are available. Subbasin c is a local ungauged catchment. Runoff from c is estimated using an impulse response rainfall–runoff model. The inflow ($I_{t,\ \mathrm{A}}$) from the upstream site A (see Fig. 1) is routed to the location C, where it is combined with the estimated lateral inflows (U_{tc}) from the ungauged catchment c and is then routed ($\hat{O}_{t,\ \mathrm{B}}$) to the downstream site B.

The Muskingum method for flood routing uses storage and continuity equations which are stated, respectively, as

$$S_m = K_m[I_m x_m + (1 - x_m)O_m] \tag{1}$$

and

$$\frac{\mathrm{d}S_m}{\mathrm{d}t} = I_m - O_m \tag{2}$$

where m is a reach index, S_m is the reach channel storage, I_m and O_m are the reach inflow and outflow, respectively, and K_m and x_m are the Muskingum parameters. Eqs. (1) and (2) when expressed in finite difference form and solved for the outflow at time

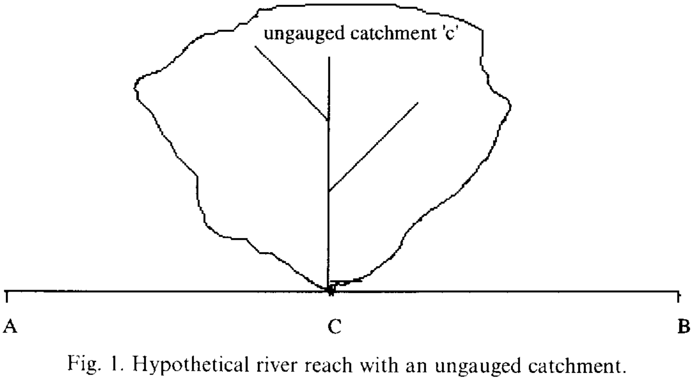

Fig. 1. Hypothetical river reach with an ungauged catchment.

step $t+1$ yield

$$\hat{O}_{t+1,\ m} = C_{1,\ m} I_{t+1,\ m} + C_{2,\ m} I_{t,\ m} + C_{3,\ m} \hat{O}_{t,\ m} \tag{3}$$

where $C_{1,\ m}$, $C_{2,\ m}$ and $C_{3,\ m}$ are expressed in terms of K_m and x_m and $\Delta t = t_2 - t_1$ the routing time step as

$$C_{1,\ m} = \frac{\Delta t - 2K_m x_m}{2K_m(1 - x_m) + \Delta t} \tag{4}$$

$$C_{2,\ m} = \frac{\Delta t + 2K_m x_m}{2K_m(1 - x_m) + \Delta t} \tag{5}$$

$$C_{3,\ m} = \frac{2K_m(1 - x_m) - \Delta t}{2K_m(1 - x_m) + \Delta t} \tag{6}$$

The parameters K_m and x_m are usually estimated using a graphical or least-squares procedure (Singh, 1992). Stephenson (1979) and Gill (1978) use linear programming techniques. These methods can work well in the absence of lateral inflows or when such flows are small. Extensions to consider ungauged lateral inflows are usually indicated.

The impulse response function approach suggested by Harpin and Cluckie (1981) is used to estimate the lateral inflows from any ungauged catchment c

$$U_{tc} = \mathrm{FAC}_c \sum_{j=1}^{n} h_{t_j c} P_{(t-t_j)c} + B_{0c} \tag{7}$$

where U_{tc} is the lateral inflow into the river at cross-section C at any time point t, P_{tc} is the rainfall over the ungauged catchment, FAC_c is a runoff coefficient (between 0 and 1) that accounts for abstractions and is used to estimate the rainfall excess from the storm rainfall and B_{0c} is the baseflow addition. Chow et al. (1989) (p. 139) mention that, due to the highly variable rainfall intensity, the runoff coefficient is difficult to determine from the observed data. Consequently we chose to determine FAC_c by the optimization as well.

It is generally known that FAC_c varies during a rainfall event, depending on rainfall intensity and duration. It can increase from near zero to above 0.5, as the storm event progresses. The FAC_c used here can be thought of as an average over the event and it is used to get an appropriate total runoff volume. A time-varying FAC_c could be accommodated if the investigator were willing to make additional assumptions regarding the time behaviour of FAC_c, e.g. one could parametrize it with a parametric function monotonically increasing with time, with an asymptote for large time and 0 for small time.

The variable h_{tc} is an impulse response function that gives the characteristic response of the basin to unit precipitation excess. It is given by Harpin and Cluckie (1981) as

$$h_{tc} = \int_{t-1}^{t} \frac{1}{2\sqrt[2]{\pi(\mathrm{ATK}_c)}} \frac{L_c}{t^{1.5}} \exp\left\{ -\left[\frac{(CV_c - L_c)^2}{4(\mathrm{ATK}_c)t} \right] \right\} \mathrm{d}t \tag{8}$$

where L_c is the length of the stream in the ungauged catchment (km), CV_c is the velocity of the flow (m s^{-1}) and ATK_c is the attenuation parameter. The extended trapezoidal rule (Section 4.1 in Press et al. (1989)) is used to evaluate the integral in Eq. (8) at each time point.

Any other model that allows computation of unit hydrograph for an ungauged basin with typically available information (e.g. rainfall over the ungauged basin, area of the basin, etc.) could be used. The approach used here is conceptually consistent with the level of complexity considered in the Muskingum routing. Harpin and Cluckie argue that this approach is likely to provide the longest forecast lead time, since the delay before runoff in the ungauged section may be of similar magnitude to the travel time of the main flood wave.

For our example, we have one ungauged catchment, and two subreaches (i.e. $c = 1$ and $m = 1, 2$). The estimated outflow at the downstream gauge is then determined by following the recursion given in Eq. (3), to the downstream gauge where

$$I_{t,\ m} = \hat{O}_{t,\ m-1} + U_{tc} \tag{9}$$

We use non-linear optimization for estimating the Muskingum routing parameters K_m and x_m, impulse response parameters ATK_c and CV_c, the runoff coefficients FAC_c and the uniform baseflow variables B_{0c}.

3. Parameter estimation

The parameter estimation problem can be formulated as a non-linear optimization problem with the objective function to be minimized as

$$Z = \sum_{t=1}^{n} (\hat{O}_{t,\ \mathrm{B}} - O_{t,\ \mathrm{B}})^2 \tag{10}$$

where $\hat{O}_{t,\ \mathrm{B}}$ and $O_{t,\ \mathrm{B}}$ are the estimated and observed outflows at the downstream gauge B. $\hat{O}_{t,\ \mathrm{B}}$ is estimated as described in the previous section. Note that $\hat{O}_{t,\ \mathrm{B}}$ is a non-linear function of the parameters defined through Eqs. (3)–(9).

Z is minimized subject to the following constraints.

(1) A bound on the fractional error in pointwise outflow

$$-\epsilon_t \leqslant (1 - \hat{O}_{t,\ \mathrm{B}}/O_{t,\ \mathrm{B}}) \leqslant \epsilon_t; \quad t = 1, \ldots, n \tag{11}$$

This constraint is desirable to limit the range of feasible solutions to ones that have desirable attributes, specifically, matching the time to peak and the outflow peak discharge. Tighter values of ϵ_t are recommended at points where the observed outflow is above the average outflow. One strategy for specifying the ϵ_t values is to prescribe $\epsilon_{t_{\mathrm{peak}}}$ (i.e. at time to peak t_{peak} where the outflow is O_p) and ϵ_{cr} (i.e. at the lowest outflow O_b) to a desired level. Now ϵ_t at any time point t is specified by logarithmic interpolation between $\epsilon_{t_{\mathrm{peak}}}$ and ϵ_{cr} as

$$\epsilon_t = \epsilon_{t_{\mathrm{peak}}} O_p^b (1/O_t)^b \tag{12}$$

where

$$b = \log(\epsilon_{t_{\mathrm{peak}}}/_{\mathrm{cr}})/\log(O_b/O_p) \tag{13}$$

Another strategy would be to adopt linear interpolation between $\epsilon_{t_{\mathrm{peak}}}$ and ϵ_{cr}. Typically, one should specify $\epsilon_{t_{\mathrm{peak}}} \ll \epsilon_{\mathrm{cr}}$. This scheme ensures that the peak flow characteristics are well preserved.

(2) A bound on the fractional error in outflow volume approximated as a discrete sum of the hydrograph ordinates

$$\epsilon_{\mathrm{vl}} \leqslant \frac{\sum_t O_{t,\ \mathrm{B}}}{\sum_t \hat{O}_{t,\ \mathrm{B}}} \leqslant \epsilon_{\mathrm{vu}} \tag{14}$$

We chose $\epsilon_{\mathrm{vl}} = 0.9$ and $\epsilon_{\mathrm{vu}} = 1.1$ for this study.

(3) Estimated baseflow is within some tolerance of a prior estimate BASE

$$(1-\tau)\mathrm{BASE} \leqslant \sum_c B_{0c} \leqslant (1+\tau)\mathrm{BASE} \tag{15}$$

for this study we chose $\tau = 0$ and $\mathrm{BASE} = O_{1,\ \mathrm{B}} - I_{1,\ \mathrm{A}}$.

(4) Bounds on decision variables

$$K_{m,\ \mathrm{l}} \leqslant K_m \leqslant K_{m,\ \mathrm{u}} \tag{16}$$

$$0.0 \leqslant x_m \leqslant 0.5 \tag{17}$$

$$CV_{c,\ \mathrm{l}} \leqslant CV_c \leqslant CV_{c,\ \mathrm{u}} \tag{18}$$

$$\mathrm{ATK}_{c,\ \mathrm{l}} \leqslant \mathrm{ATK}_c \leqslant \mathrm{ATK}_{c,\ \mathrm{u}} \tag{19}$$

$$0.0 \leqslant \mathrm{FAC}_c \leqslant 1.0 \tag{20}$$

$$B\{0c,\ \mathrm{l}\} \leqslant B_{0c} \leqslant B_{0c,\ \mathrm{u}} \tag{21}$$

These bounds condition the solution to lie in a suitable feasible region, where $K_{m,\ \mathrm{l}}$, $CV_{c,\ \mathrm{l}}$, $\mathrm{ATK}_{c,\mathrm{l}}$, $B_{0c,\ \mathrm{l}}$, $K_{m,\ \mathrm{u}}$, $CV_{c,\ \mathrm{u}}$, $\mathrm{ATK}_{c,\mathrm{u}}$ and $B_{0c,\ \mathrm{u}}$ are the lower and upper bounds of the parameters K_m, CV_c, ATK_c and B_{0c}, respectively.

The parameter estimation scheme described here is easily generalized to situations with multiple outflow records (i.e. multiple reaches). In such cases one could move reach by reach solving independent optimization problems for each one. We feel that this is preferable to solving a single large optimization problem for all reaches simultaneously. The computational effort involved in solving the single larger optimization problem is significantly greater than that for solving the sequence of individual problems. The optimization process can be readily automated to solve the reach by reach sequence of optimization problems. Since, the accuracy constraints are specified for outflow at each reach, identifiability of the overall model is maintained.

A number of options are available for treating multiple storms. One could solve the optimization problem once for multiple storms and determine 'average' optimal values for the parameters K_m, x_m, CV_c, ATK_c, FAC_c and B_{0c}. Alternatively, one could solve for these parameters storm by storm, and then try to relate their variation to storm and antecedent condition attributes. The choice between these two strategies may depend on the purpose of the application and the data and resources available.

Sometimes it may be desirable to consider a 'non-linear' flow generation and routing model, where the parameters depend on the magnitude of flow or rainfall, and may even be related to known drainage basin characteristics such as slope, drainage area, surface soils and elevation. If parametric functions that can describe such relations are available or can be assumed, the associated parameters can also be solved for in the same manner as described here.

4. The solution algorithm

It can be seen from the model described above, that the objective function and the constraints are non-linear in the decision variables. This problem is solved using a Feasible Sequential Quadratic Programming (FSQP) algorithm developed and implemented by Zhou and Tits (1993). This algorithm solves the minimization of a set of smooth non-linear objective functions subject to general smooth non-linear constraints. Zhou and Tits (1993) argue that this algorithm is globally convergent and locally superlinear convergent.

A succession of quadratic programs is solved to determine the optimal solution formed by Taylor series approximations of the functions at each solution point (see Luenberger, 1973). The reader is referred to Zhou and Tits (1993) for details of the algorithm. The FSQP routines can be obtained by contacting Professor Andre'L. Tits, Electrical Engineering Department and Institute for Systems Research, University of Maryland, College Park, MD 20742 (e-mail: andre@eng.umd.edu).

5. Applications

The model was applied to three cases, two synthetic data sets, with and without lateral inflows, and one real data set from Dhalegaon–Gangakhed reach of the Godavari river in India. The initial values for the parameters K and x were given based on the observed discharge hydrographs, and for other parameters (CV_c, ATK_c and FAC_c) were based on past experience.

5.1. Case 1 (synthetic data set without lateral inflows)

Consider a reach AB (Fig. 1) with a triangular inflow hydrograph at A as shown in Fig. 2. The observed outflow hydrograph is obtained by routing this hydrograph to B using $K = 4$ h and $x = 0.2$. The number of time periods considered was 25 and the number of constraints was 52. The estimated K and x from our model are 4.0 h and 0.2, respectively. Here we used $\epsilon_t = \epsilon = 0.01$. The hydrographs are shown in Fig. 2.

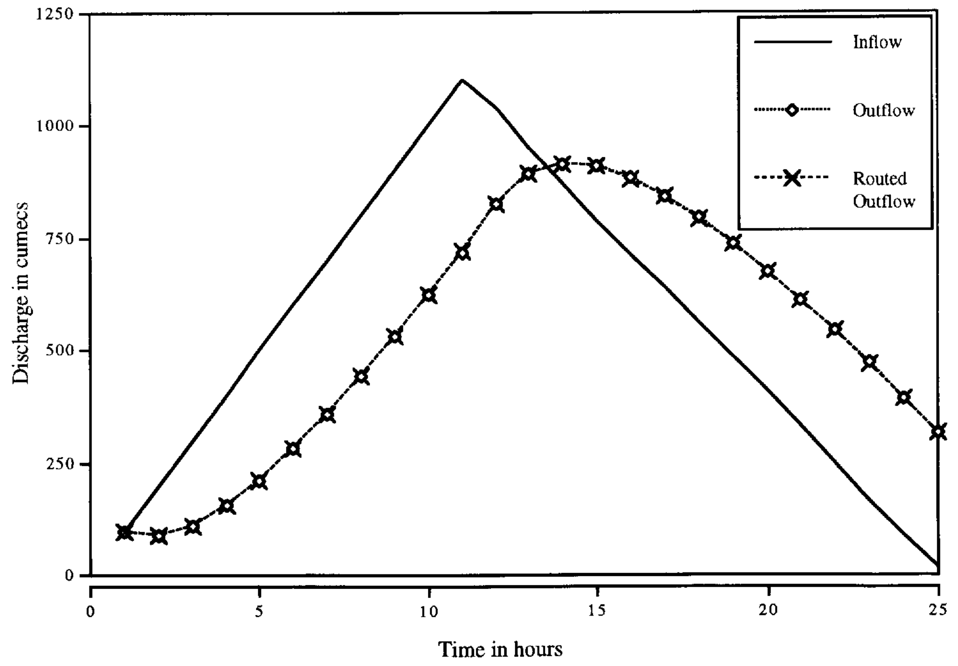

Fig. 2. Flood hydrographs — synthetic data set (case 1) without lateral inflow.

5.2. Case 2 (synthetic data set with lateral inflows)

Reach AB is considered with a inflow hydrograph at A as shown in Fig. 3. The rainfall considered over the single ungauged catchment is shown in Fig. 3. Lateral inflows were derived from this rainfall using Eqs. (7) and (8) with $CV_c = 1.80$ m s^{-1}, $ATK_c = 400$ and $FAC_c = 1.00$. The baseflow in the ungauged catchment was assumed to be zero. The observed outflow hydrograph is obtained by routing this to B using $K_1 = 5.5$ h, $x_1 = 0.2$, $K_2 = 4.0$ h and $x_2 = 0.2$ for the two reaches, respectively. The number of time periods considered were 48 and the number of constraints was 92. The estimated parameters from the model were $K_1 = 5.5$ h, $x_1 = 0.2$, $K_2 = 4.0$ h, $x_2 = 0.2$, $CV_c = 1.80$ m s^{-1}, $ATK_c = 400$ and $FAC_c = 1.00$. Here $\epsilon_t = \epsilon = 0.01$ was used. The hydrographs are shown in Fig. 3.

5.3. Case 3 (real data set with lateral inflows)

The Godavari is one of the major rivers in India, flowing through the states of Maharashtra and Andhra-Pradesh. The reach under consideration was Dhalegaon–Gangakhed which is about 500 km. from the source. Flow observations are made at Dhalegaon and Gangakhed and the length of the river between this reach is 93 km. The ungauged catchment between these two gauging sites is 3035 km^2. The map of the river system is shown in Fig. 4. Most of the floods in this region are due to rainfall

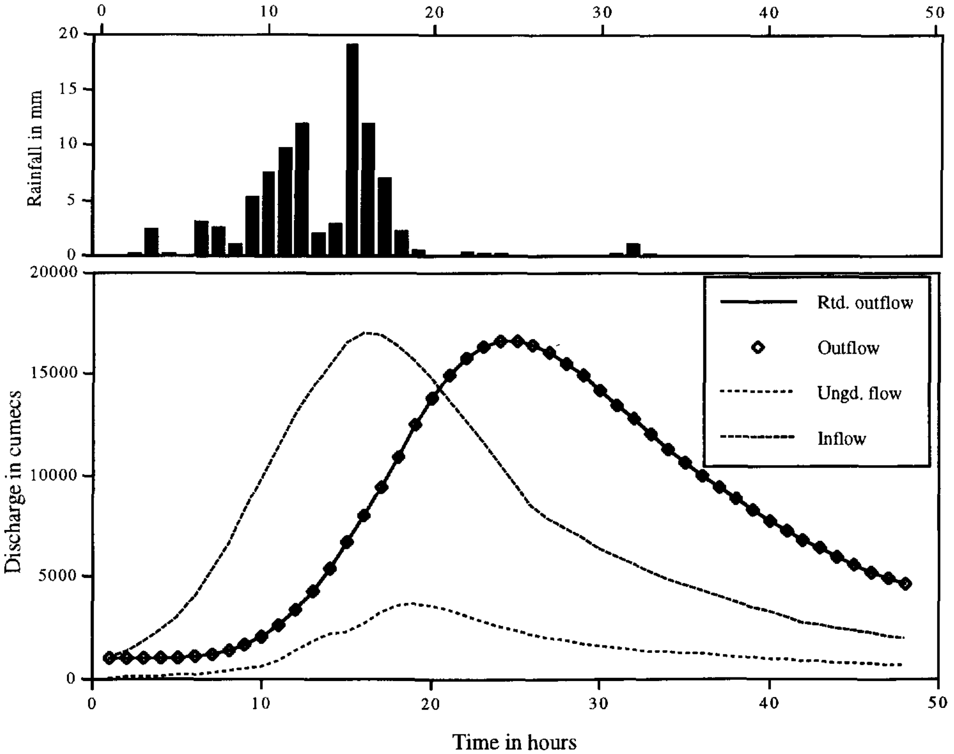

Fig. 3. Rainfall over single ungauged catchment and flood hydrographs — synthetic data set (case 2) with lateral inflow.

from the southwest monsoon, which is active from June–October. Observed discharge at Dhalegaon is the inflow (I_t) and the observed discharge at Gangakhed is the outflow (O_t). Daily weighted rainfall over the catchment was obtained using Thiessen polygon method with daily rainfall data from seven stations situated in and around the catchment. The daily weighted rainfall thus obtained is distributed over a 24 h period using the hourly rainfall pattern from the hourly rain gauge station, Gangakhed (see Fig. 4). The weighted hourly rainfall pattern thus obtained, for the ungauged catchment is shown in Fig. 5. If data from more than one hourly rain gauge station are available then a weighted hourly pattern is obtained for distribution of the daily weighted rainfall. The flood case studied here was during the period 31 August 1977 (10:00 h) to 5 September 1977 (24:00 h). The values obtained for the parameters were $K_1 = 2.00$ h, $x_1 = 0.01$, $K_2 = 8.42$ h, $x_2 = 0.01$ for the two subreaches with $CV_c = 4.50$ m s^{-1}, $\text{ATK}_c = 20.22$ and $\text{FAC}_c = 0.16$ for the ungauged lateral inflows. Following the prescription given earlier B_{0c} is taken to be 26.0 cumecs (i.e. $O_{1,\text{B}} - I_{1,\text{A}} = 36.0 - 10.0$). We chose ϵ_t using Eqs. (12) and (13) with $\epsilon_{t_{\text{peak}}} = 0.1$ and $\epsilon_{\text{cr}} = 0.8$. The number of time periods for this storm was 135 and the number of constraints was 84. These estimates of K are consistent with the field estimates of travel time in these reaches and also agree with the values obtained from other

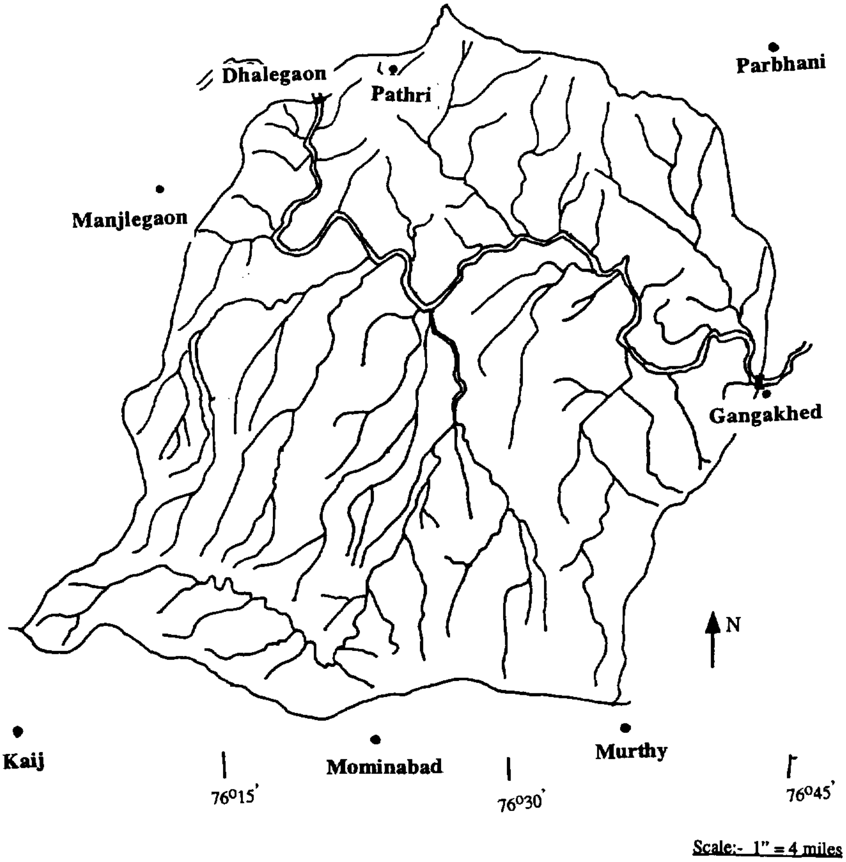

Fig. 4. Map of ungauged catchment between the reach Dhalegaon and Gangakhed — the Godavari river, India. Dots show the location of raingauges.

methods. The routed and the observed hydrographs along with the estimated ungauged lateral inflows are shown in Fig. 5. Acceptable agreement between the routed and the observed outflow hydrograph is observed.

6. Summary and conclusions

A straightforward application of non-linear optimization to the determination of parameters for a simple, linear model of flood routing and rainfall runoff generation is presented. The presentation is kept concise and the examples simple for the sake of clarity.

The two synthetic examples showed that the optimization scheme is capable of recovering the right parameters at least for simple and identifiable situations, where we know that the underlying model is correct. However, there is no assurance that the

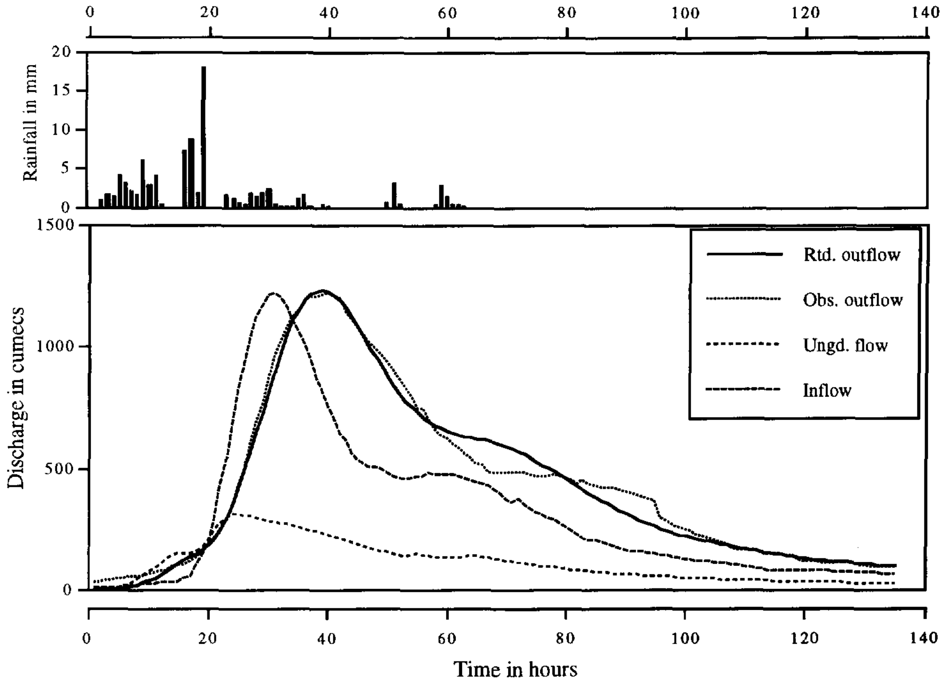

Fig. 5. Rainfall over ungauged catchment and flood hydrographs — Dhalegaon–Gangakhed reach of the Godavari river, India.

simple, linear flow routing and runoff generation models used here are correct in a 'real' situation. In such a case, the parameters cannot be determined uniquely.

It is unlikely that except in a contrived example the routed and observed outflow hydrographs would or should be identical. The representation of the physical processes is far from complete or unique in such models. Moreover, model parameters that lead to a certain minimum of the fitting criteria (least sum of squares of errors) need not be uniquely specified. The same total squared error may be obtained by parameter choices that (1) undershoot the peak outflow dramatically, match a large part of the outflow hydrograph and then decay slowly relative to the observed outflow and (2) match the outflow hydrograph with a small pointwise error throughout, rather than a large under or overshoot at a point. Hence the need for an optimization solution that is constrained to have bounded pointwise error as well, as is done here.

The optimization algorithm merely steers us towards one of possibly many acceptable solutions. Part of this is due to possible model misspecification, and part is due to the lack of information on causative processes and their operative values during the events of interest.

Increased sophistication in process representation calls for substantially greater amount of information. The associated increase in the number of parameters to be solved reduces the degrees of freedom, but may not improve the identifiability of the model. The simplicity of the model presented here is attractive for exploratory practical applications with the amount of data typically available in developing countries.

Acknowledgements

Partial support of this work by the United Nations Development Program for their funding of the first author's fellowship at UWRL, Logan is acknowledged. We thank P.P. Vaidyaraman, V.M. Bendre, CWPRS, Pune, India and D.S. Bowles, UWRL, Logan, for their encouragement in this work. We are very grateful to Andre'L. Tits for provision of the optimization subroutine (FSQP Version 3.3b), the manual and his help with installation of the same.

References

Chow, V.T., Maidment, D.R. and Mays, L.W., 1989. Applied Hydrology. McGraw-Hill, New York.

Gill, M.A., 1978. Flood routing by the Muskingum method. J. Hydrol., 36: 353–363.

Harpin, R. and Cluckie, I.D., 1981. Real-time flood routing in the presence of high magnitude ungauged lateral inflows. In: V.P. Singh (Editor), Rainfall–Runoff Relationships. Water Resources, Littleton, CO.

Luenberger, D.G., 1973. Introduction to Linear and Nonlinear Programming. Addison-Wesley, Reading, MA.

Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T., 1990. Numerical Recipes. Cambridge University, New York.

Singh, V.P., 1992. Elementary Hydrology. Prentice Hall, Englewood Cliffs, NJ.

Stephenson, D., 1979. Direct optimization of Muskingum routing coefficients — an extension to the paper by M.A. Gill (1978): 'Flood routing by the Muskingum method'. J. Hydrol., 41: 161–165.

Zhou, J.L. and Tits, A.L., 1993. User's Guide for FSQP Version 3.3b: A Fortran Code for Solving Constrained Nonlinear (Minimax) Optimization Problems, Generating Iterates Satisfying All Inequality and Linear Constraints. TR-92-107r3, Electrical Engineering Department and Institute for Systems Research, University of Maryland, College Park, MD.