JOURNAL OF
ENVIRONMENTAL HYDROLOGY

The Open Access Electronic Journal of the International Association for Environmental Hydrology
On the World Wide Web at http://www.hydroweb.com

VOLUME 22 2014

TOPMODEL PARAMETER OPTIMIZATION USING MONTE CARLO SIMULATION AND GENERALIZED LIKELIHOOD UNCERTAINTY ESTIMATION (GLUE) PACKAGE

Rajesh Vijaykumar Kherde [1]
Priyadarshi H. Sawant [2]

[1] **Symbiosis Institute of Technology, Lavale, Pune, India**
[2] **Sardar Patel College of Engineering, Andheri, Mumbai, India**

Topography plays an important role in the process of runoff generation dynamics. The TOPMODEL developed by Beven and Kirkby (1979) is a variable source area conceptual model that produces runoff hydrographs using topographic features of the basin. Although the TOPMODEL parameters are physically based and can be determined directly; physical measurement of TOPMODEL parameters may not be possible in data sparse regions. In such cases of ungauged catchments split record test model calibration process that uses the results of Monte Carlo simulation is usually adopted for parameter estimation. The calibration process reflects uncertainty. It is observed that there could be many different parameter sets that simulate the observed stream flow in terms of quantitative goodness of fit measure (Beven and Binley 1992). One set of parameter value is observed to be optimum, but many other from different parts of parameter range are acceptable simulators of runoff. Choosing the optimal parameter set remains a point blank question perhaps in the mind of modelers. This research paper demonstrates a method for arriving at the optimum parameter set whilst it evaluates the applicability of TOPMODEL in Wardha watershed of Maharashtra, India.

INTRODUCTION

In the year 1976, Beven & Kirkby proposed a topography based model for humid temperate areas that combines the advantages of lumped parametric as well as distributed model. The complex, dynamic and non linear rainfall runoff process is very difficult to understand. The soil moisture undergoes continuous changes with respect to space and time. In humid temperate areas hortons's infiltration excess overland flow model is not applicable as the infiltration capacities are generally higherthan normal rainfall intensities (Kirkby, 1969). The recent hydrological modeling research mostly focus on integration of hydrological models with GIS technology that uses digital terrain data.

The rainfall intensity exceeds the infiltration capacity only in a part of total catchment area (Beston 1964). The deficiency of soil moisture content is variable in nature, Quinn et al.(1991) studied subsurface flow and sensitivity of flow path direction. Zhang and Montgomery (1994) calculated topographic index for different grid sizes. Quinn concluded that to represent effect of topography more accurately the grid size should be less than 10 m. Pinol et al. while studying effective area of subsurface flow suggested modifications to topographic index approach. Brasington and Richards (1998) pointed out the break in model response and DEM resolution. TOP model when applied to Malprabha basin Venkatesh and Jain (2000) obtained acceptable simulation of flows. In upland areas the weighting of local storage deficit may be increased by introducing reference topographic index (Campling et al., 2002). The TOPMODEL is observed to be more suitable for catchments with shallow ground water table, hill slope topography having moist soil (Shufen & Huiping, 2004).

DESCRIPTION OF STUDY AREA

Wardha is one of the right bank tributaries of Pranhita River. The Wardha sub basin lies between latitude 19°18'N and 21°58'N and longitudes 77°20'E and 79°45'E. Wardha originates at an altitude of 777 m in the Betul district of Madhya Pradesh, India and enters Maharashtra about 32 km from its source. After traversing a distance of 528 km, it joins the Wainganga at an elevation of 146 m. The major left bank tributaries of the Wardha are the Kar, the Wena, The Jam and the Erai and the right bank tributaries are the Madu, the Bembla and the Penganga. The drainage area of the Wardha River is 47985 km^2 and throughout its course, the river flows through dense forests. The average annual rainfall for the entire sub-basin is 1,000 mm approximately. Wardha river basin frequently gets flooded on its banks and causes damage to nearby properties.

DELINEATION OF THE WARDHA BASIN

A high definition (30m) digital elevation model was obtained via ASTER GDEM. Every pixel of DEM represents the average elevation of 30m x 30m area. The GIS software interprets these elevations for producing various derivatives, one of which is to delineate watersheds. For delineation of watershed the GIS routine first search for any sinks. Sinks are low elevation points of DEM wherein the water is trapped. The flow of water would be disrupted had the sinks not been filled, resulting in number of ponds and lakes having no outlet. After raising the elevation of sinks the GIS software query the elevation of eight cells that surround any one cell in DEM to determine the largest change in elevation between two points – the direction of water flows. Based on the elevation of each cell and previously assigned flow direction the flow accumulation routine calculates the number cells that are contributing to the drainage of each cell. A unique value and flow direction is then assigned to every segment of the stream. The smallest watershed accumulates 500 cells while the largest watershed accumulates 1,00,000 cells. The watershed thus delineated (Figure 1) is found to be consistent with the manually delineated watershed by Ground water surveys and development Agency (GSDA), India.

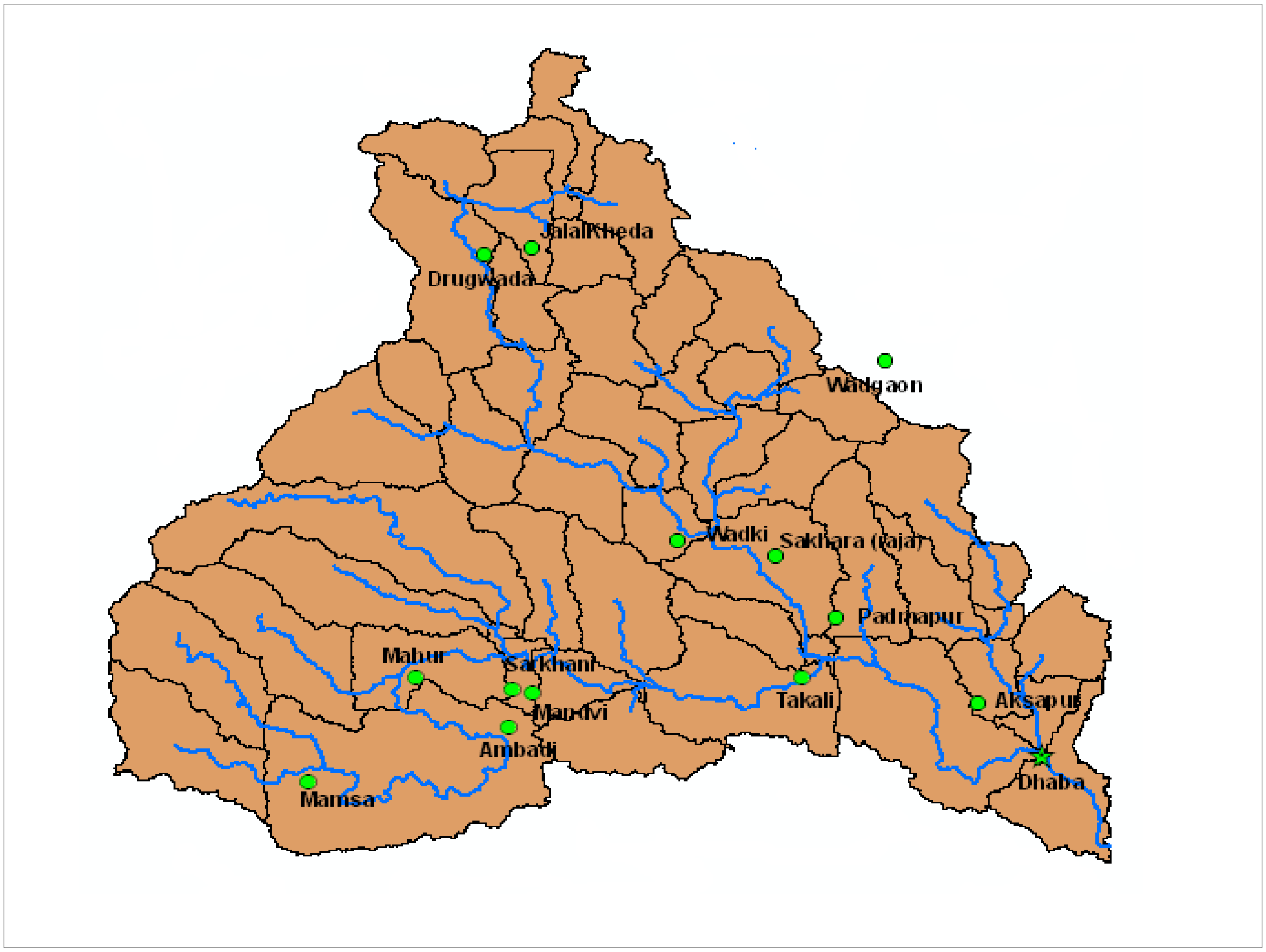

Figure 1. Delineated Wardha watershed with location of raingauges and gauge discharge station.

DATA COLLECTION AND COLLATION

The stream flow processes is simulated using following data:

Meteorological data: The precipitation and evaporation data (2002 to 2010) collected for the fifteen weather stations of Wardha stream were obtained from hydrological data user Group (HDUG) Nashik, Maharashtra, India. The data was then disposed with thiessen polygon method to calculate average precipitation and evaporation of the Wardha basin. The precipitation and evaporation is then converted to m/hr units. The time step of meteorological data used in the study is 24 hours.

Stream flow data: The stream flow data (from 2002 to 2010) for the pour point Dhaba hydrological station was obtained from the HDUG.

Topographic index: The topographical index the kernel of TOPMODEL is sensitive to the DEM resolution. Greater resolution of DEM results in higher values of topographic index. Topographic index reflects the spatial distribution of water storage capacity of grid. The ideal resolution is considered as 50m, in present study a DEM of 30m resolution is used. In Wardha basin the minimum value of topographic index is 0.615688 and the maximum is 17.85484.

TOPOGRAPHIC INDEX

The topographic slope tanβ represents the hydraulic gradient of saturated area. An exponential function of storage describes the conductivity (storage deficit) of soil profile, with the value of T_0 when the soil is saturated to the surface. A saturated zone is in equilibrium with a steady recharge rate over an upslope contributing area"a".

The soil water shortage "D" is taken as difference between soil moisture and saturated moisture. The saturation source area is the area where surface runoff generates when D ≥ 0.

The down slope saturated subsurface flow rate (*Qi*) per unit contour length at any point I on a hill slope is given by

$$Qi = T_0 \tan \quad \exp\left(\frac{-Di}{m}\right) \tag{1}$$

where D_i = local water storage deficit per unit plan area at the location of grid *i* (m), *m* = Coefficient of saturated transmissivity (m), T_0 = Efficient infiltration rate just to the extent of saturation (m^2/h), and tanβ = The hydraulic gradient on the basis that the slope is calculated as elevation change per unit distance in plan (rather than along the hillslope).

Assuming that for any time step quasi- steady-state flow exists throughout the soil and that the water enters the water table with spatially homogeneous recharge rate, the subsurface flow per unit contour length *Qi* is

$$Qi = r\,a \tag{2}$$

where *a* is the area of the hill slope per unit contour length that drains through point *i* .

Equations (1) and (2) combined together gives the water table depth for any point in terms of the topographic index [ln(a/tanβ)] , *m*, T_0 and *r*.

$$Di = -m \ln\left(\frac{r\,a}{T_0 \tan\beta}\right) \tag{3}$$

For fully saturated soil profile the local deficit is zero on the other hand when soil dries and water table drops the value of storage deficit increases. The mean storage deficit may be obtained by integrating Equation (3) over the entire catchment area *A*.

$$\overline{D} = \frac{1}{A} \sum A_i \left[-m \ln\left(\frac{r\,a}{T_0 \tan\beta}\right)\right] \tag{4}$$

where Ai is the area associated with point i. Assuming r to be spatially constant, $\ln(r)$ may be eliminated, this assumption results in

$$Di = \overline{D} + m\left[\lambda - \ln\left(\frac{a}{T_0 \tan\beta}\right)\right] \tag{5}$$

where $\ln\left(\frac{a}{T_0 \tan\beta}\right)$ is the soil-topographic index,

$$\lambda = \frac{1}{A}\sum A_i\left[\ln\left(\frac{a}{T_0 \tan\beta}\right)\right] \tag{6}$$

Areal average value of transmissivity is given by,

$$\ln(T_e) = \frac{1}{A}\sum A_i \ln(T_0) \tag{7}$$

Equation (5) may be rearranged to give

$$\frac{\left(\overline{D} - D_i\right)}{m} = -\left[\lambda - \ln\left(\frac{a}{\tan\beta}\right)\right] + \left[\ln(T_0) - \ln(T_e)\right] \tag{8}$$

where $\lambda = \frac{1}{A}\sum A_i\left[\ln\left(\frac{a}{T_0 \tan\beta}\right)\right]$ is topographic constant for the catchment.

POTENTIAL EVAPOTRANSPIRATION (PET)

In data sparse region where only limited meteorological data such as air temperature is only available, the FAO-56 PM equation recommended as the standard for computing PET cannot be applied. In such situations, the air temperature based Hargreaves equation is recommended for estimating PET. In the present study Hargreaves Method was used to calculate the potential evapotranspiration, Equation (9). This PET is also converted from mm/day to m/h.

$$\text{PET} = 0.0023 * R_{ext} * \left(T_{avg} + 17.8\right) * \left(T_{max} - T_{min}\right)^{0.5} \tag{9}$$

Where, PET = potential evapotranspiration (mm/day), R_{ext}= daily extra terrestrial radiation (watts/m^2), T_{max}= daily maximum temperature (^oC), T_{min}= daily minimum temperature (^oC), T_{avg}= daily average temperature (^oC), Note: Radiation 1 MJ/m^2 = 0.408 mm/day.

HYDROGRAPH SIMULATION

For simulating stream flow process a 30 m resolution DEM is used to generate digital drainage network. The period from 2002 to 2004 was used for calibrating model parameters and the period from

2005 to 2008 was used for validation of daily stream flow simulation. For simulation of stream flow only the precipitation, evaporation and discharge data observed from June to October every year is used in the study, since other months are mostly dry. Nash-Sutcliffe efficiency is calculated to evaluate the performance of TOPMODEL, which express as

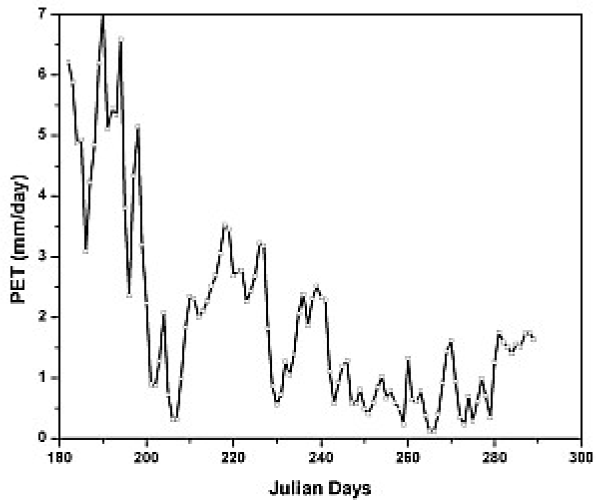

Figure. 2 Seasonal Variation of PET for the year 2002.

$$R^2 = 1 - \sum \left(Q_i - q_i \right)^2 / \sum \left(Q_i - Q \right)^2 \qquad (10)$$

where Q_i= measured stream flow, q_i = simulated stream flow, and Q = average measured stream flow.

For calibration the model parameter 'm' was varied keeping the values of other four parameters at initial value. The value of parameter 'm' is set at one which yields the highest efficiency. Next the value of $\ln(T_0)$ was varied to further enhance the efficiency. The same procedure was repeated for other parameters to arrive at set of parameters which gives highest efficiency. For best fit the other three criterions should move close to zero. These criteria are

Sum of squared errors,

$$SSE = \sum_{i=1}^{n} (Q_i - q)^2$$

Sum of squared log error,

$$SLE = \sum_{i=1}^{n} \left\{ \log(Q_i) - \log(q) \right\}^2$$

Sum of absolute error,

$$SAE = \sum_{i=1}^{n} \left| (Q_i - q) \right|$$

THE MODEL CALIBRATION PROCESS

An initial run of the model is made with the guessed values of the parameters based on the past research for Indian sub continent. To arrive at the best parameter set each chosen parameter is varied across its range, keeping the values of the other parameters constant. The range of parameter values for which model gives positive efficiency are only considered, parameter values giving negative efficiency are discarded.

The parameter 'm', controlling the decline of transmissivity with increasing storage deficit, shows the greatest control on performance. The reason behind the same is that, the decay parameter 'm' controls the effective depth of catchment soil profile. It does this in conjunction with parameter '$Ln(T_0)$' , which defines the transmissivity of the profile when saturated to the surface.

SENSITIVITY ANALYSIS & UNCERTAINTY ESTIMATION

The Generalised Likelihood Uncertainty Estimation (GLUE) package (Beven and Binley 1992) is used for sensitivity analysis and uncertainty estimation. The GLUE concept rejects the idea of optimum parameter set. In fact the parameter set that predict required variable is dominant in fitting available observations. The results of Monte carlo simulations are used to determine prediction limits of required variables. The parameter set that produce reasonably good results in evaluation are given greatest weight in prediction. Prediction limits are based on available sample of prediction and likelihood weighted distribution of variable may vary greatly. The value of likelihood measure should increase monotonically with increasing performance of the model. Applying the lower threshold serve this purpose by considering zero values for likelihood value of parameter sets below threshold. Thus analysis focuses on parameter sets rather than their behaviour. Likelihood measure thus reflects ability of particular model to predict particular series of observations. GLUE methodology is an exercise in model rejection. The acceptable parameter sets are often found to come from the wide ranges of parameter space. This pose a difficulty in interpretation of feasible parameter sets.

Following methodology was adopted to arrive at optimum parameter set.

1. For Monte Carlo simulations optimum parameter set obtained from the results of manual trials are used as current values. The range of parameter values giving efficiency above 70 percent is displayed in Table 1.

Table 1. Parameter values defined with TOPMODEL

Parameter	Value	Unit	Manually evaluated range for sensitivity analysis
m	0.017	M	0.017 – 0.024
$Ln(T_0)$	0.023	m^2/hr	0.01– 0.3
SR_{max}	0.20	M	0.001– 0.3
SR_{init}	0.0020	M	0.001 – 0.03
ChVel	3600	m/hr	1000 – 5000

2. List of best simulators is obtained for the first year of calibration period using GLUE package. GLUE package lists twenty such record sets.

3. For the best twenty parameter sets efficiencies are obtained for the subsequent years for calibration period as tabulated in Table 2.

Table 2. Ranked parameter sets obtained from sensitivity analysis.

Rank	m	Ln(To)	SRmax	SRinit	ChVel	Efficiency for the Year of Calibration period		
						2002	2003	2004
1	0.0177	0.0620	0.1370	0.0131	4145.8500	0.6993	0.7214	0.7378
2	0.0181	0.0323	0.0875	0.0286	3040.4990	0.6991	0.7213	0.7368
3	0.0190	0.2899	0.0559	0.0200	1199.0120	0.6822	0.7096	0.7363
4	0.0200	0.2224	0.0497	0.0027	1973.2430	0.6670	0.7131	0.7354
5	0.0217	0.0998	0.0270	0.0194	3194.6990	0.6324	0.7191	0.7352
6	0.0176	0.1355	0.1090	0.0193	4377.5370	0.7013	0.7093	0.7349
7	0.0212	0.1282	0.0162	0.0056	1638.5580	0.6530	0.0721	0.7360
8	0.0222	0.0822	0.0175	0.0190	1662.3440	0.7033	0.7119	0.7325
9	0.0215	0.0819	0.0088	0.0244	2221.6320	0.6539	0.7155	0.7320
10	0.0216	0.2028	0.0047	0.0086	2322.6820	0.6500	0.7159	0.7308
11	0.0176	0.0729	0.2272	0.0014	3210.6900	0.6973	0.7158	0.7302
12	0.0181	0.0316	0.1464	0.0076	4051.4210	0.672	0.7069	0.7301
13	0.0174	0.1146	0.2542	0.0044	2254.5460	0.679	0.7152	0.7300
14	0.0222	0.0983	0.0073	0.0177	4947.8980	0.700	0.7091	0.7295
15	0.0228	0.2612	0.0027	0.0079	2097.7020	0.6276	0.0228	0.7293
16	0.0172	0.0613	0.2687	0.0076	4840.2430	0.622	0.7101	0.7291
17	0.0228	0.0593	0.0114	0.0214	1390.7880	0.6125	0.7041	0.7288
18	0.0173	0.2681	0.1486	0.0128	4724.8820	0.676	0.7136	0.7288
19	0.0184	0.1046	0.1246	0.0054	3479.3260	0.619	0.7132	0.7278
20	0.0216	0.2028	0.0047	0.0086	2322.6820	0.708	0.7071	0.7259

4. Introduction of new data in validation phase leads to rejection of many more parameter sets. A poor performance in the validation phase of split-record test is perhaps a reason enough for rejection.

5. Out of the best simulators, the parameter sets those, which resulted in maximum efficiency for all the years of calibration period are selected for analysis in validation period.

6. The efficiencies calculated for validation period are tabulated in the Table 3.

7. The record set (14) is observed to give maximum efficiencies for all the years of validation period and hence taken as optimum parameter set.

8. Optimum parameter set thus obtained with efficiencies for calibration and validation period are tabulated in Table 4.

Table 3. Model efficiency for top four parameter sets during validation period.

Record Set of Table 2	Efficiency for the year of validation period			
	2005	2006	2007	2008
6	0.721	0.723	0.713	0.724
8	0.701	0.666	0.653	0.691
14	**0.746**	**0.713**	**0.726**	**0.735**
20	0.723	0.707	0.711	0.729

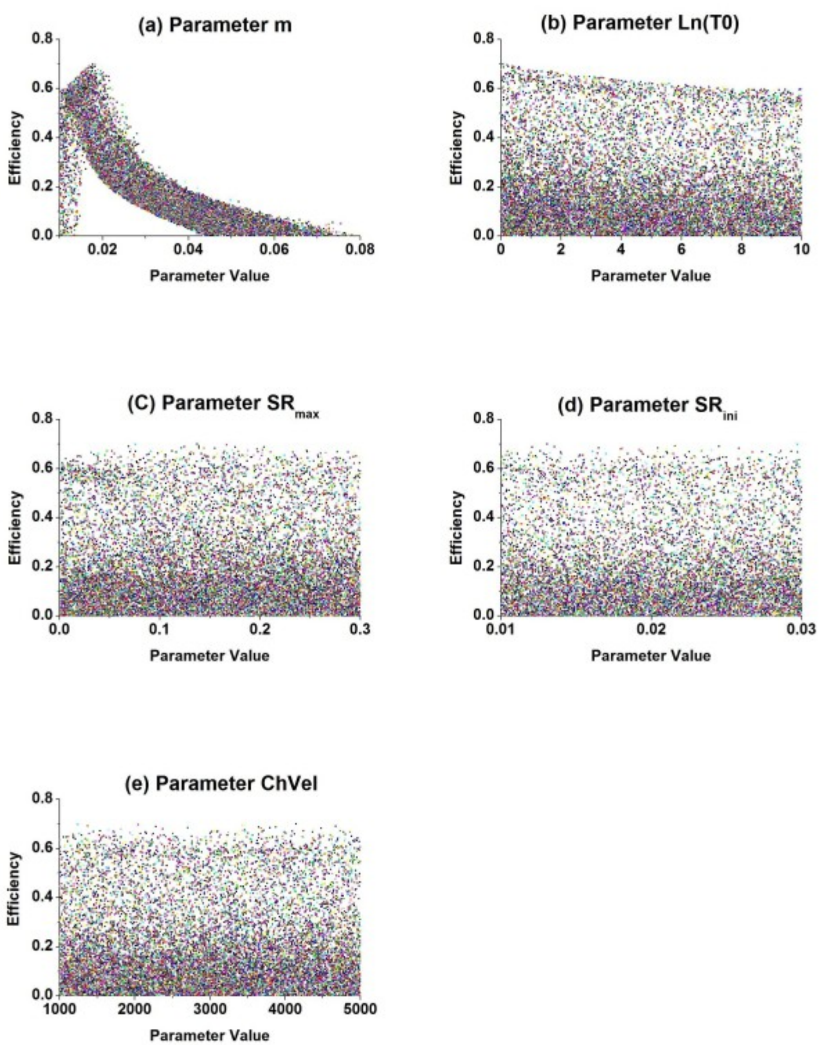

Figure 3. Scatter plots of goodness of fit expressed as a modeling efficiency.

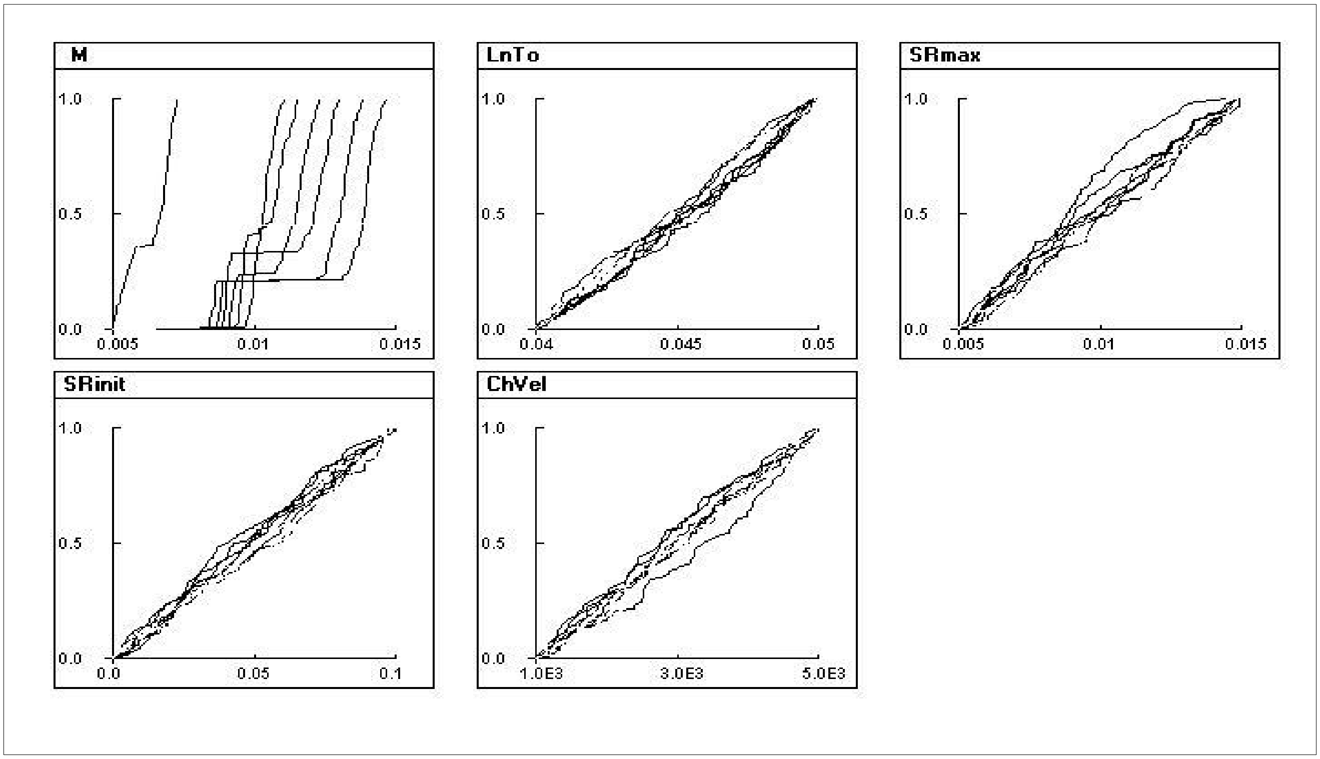

Figure 4. Parameter sensitivity plots.

Table 4. Efficiencies for the optimum parameter set obtained from GLUE analysis**.**

Optimum parameter set						
Efficiency for the calibration year			Efficiency for the validation year			
2002	2003	2004	2005	2006	2007	2008
0.700	0.7091	0.7295	0.746	0.713	0.726	0.735

CONCLUSIONS

The structural simplicity of TOPMODEL and comparatively less number of parameter estimation has made it most preferred hydrological model by researchers. The applicability of TOPMODEL is tested for the Wardha basin. The available data of rainfall, evaporation and discharge is split into two groups. The first group (2002-2004) data was used for calibration of the model and the validation is performed using the second group (2005-2008) data.

Manual trials resulted in one optimum parameter set with the range of parameter values for which model gives acceptable efficiency. A threshold of 0.7 is considered acceptable. Then each chosen parameter is varied across its range, keeping the values of the other parameters constant. It is observed that there could be many different parameter sets that will simulate the observed stream flow in terms of efficiency. For uncertainty estimation GLUE package is used to obtain top twenty best simulators. For each of these parameter set, efficiency is calculated for subsequent calibration years (Table 3). It is observed that for four out of twenty best simulators TOPMODEL gives maximum efficiency. These four parameter sets are then tested for validation period (Table 4) and the one giving maximum efficiency for calibration as well as validation period is chosen as best simulator (Table 5).

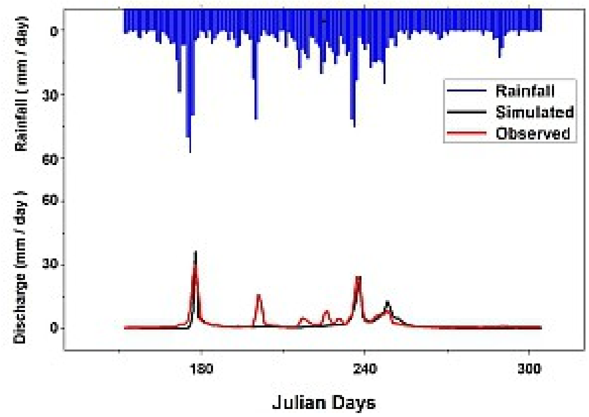

Figure 5. Observed and simulated hydrographs of Wardha watershed during calibration (2002).

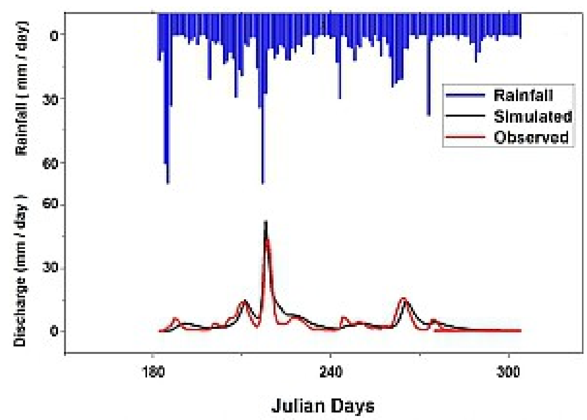

Figure 6. Observed and simulated hydrographs of Wardha watershed during validation (2006).

Hydrographs using the best simulator parameter set are displayed in Figures 5 and 6. The simulation has given better insight into the response of the catchment at different periods of the season. The TOPMODEL performed reasonably well as a continuous hydrograph simulator in the Wardha basin.

The sensitivity analysis showed greater control on performance of the model by the parameters ‘m’, ‘Ln(T0)’ and SRmax, while other two parameters SRini and CHVEl does not significantly affect the performance of the model.

ACKNOWLEDGMENT

We are very grateful to Dr K.C. Khare of the Department of Civil Engineering, Symbiosis International University, Pune, India and Dr S. Hariharan of the Department of Civil Engineering, Thadomal Shahani Engineering College, Bandra, Mumbai, India, for thoroughly reviewing the manuscript.

REFERENCES

Betson, R. P. (1964), What is watershed runoff?, J. Geophys. Res., 69(8), 1541–1552.

Beven, K. and Kirkby, M.J., 1979 a. A physically-based variable contributing area modelof basin hydrology. Hydrological Sciences Bulletin. 24, 43-69.

Beven, K.J. and Binley, A.M., 1992. The future of distributed models: model calibration and uncertainty prediction, Hydrological Processes, 6, p.279–298.

Brasington, J. and Richards, K., 1998. Interactions between Model predictions. Parameters and DTM Scales for Topmodel. Computers & Geosciences, 24(4), pp. 299-314.

Campling, P., Gobin, A., Beven, K. and Feyen, J. (2002), Rainfall-runoff modeling of a humid tropical catchment: the TOPMODEL approach. Hydrol. Process., 16: 231–253.

Kirkby, M.J., 1969. Infiltration, through flow and overland flow. Water, Earth and Man, Methuen, pp. 215-228.

Quinn, P. F. B. J., Beven, K., Chevallier, P., & Planchon, O. (1991). The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models. Hydrological processes, 5(1), 59-79.

Sun Shufen, Deng Huiping, 2004: A Study of Rainfall-Runoff Response in a Catchment Using TOPMODEL . Adv. Atmos. Sci., 21(1), 87-95.

Venkatesh B. and Jain, M.K., 2000. Simulation of daily flows using topography based rainfall-runoff model. J. Institution of engineers (India) Vol. 81, pp 127-132.

Zhang W. and Montgomery D.R., 1994. Digital elevation model grid size, landscape representation, and hydrological simulations. Water resources Research. 30(4): 1019-1028.

ADDRESS FOR CORRESPONDENCE
RajeshVijaykumarKherde
Symbiosis Institute of Technology
Lavale, Pune
India
Email: rkherde@gmail.com