Snowmelt Runoff Modelling of an Himalayan sub-tributary of the Ganges River in India: Comparison of Modelling Approaches

SARKAR A.[1,*], SONI A.[2], PECHLIVANIDIS I.[3],JAIN S.K.[1]and SINGH R.D.[1]

[1]Scientist, National Institute of Hydrology, Roorkee, India

[2]Deputy Manager, National Hydro Power Corporation, Faridabad, India

[3]Hydrology Research Unit, Swedish Meteorological and Hydrological Institute, Norrköping, Sweden

*corresponding author:

e-mail:archana_sarkar@yahoo.com

AbstractModelling of snowmelt runoff at the catchment scale is important for water resources management and flood protection. Mathematical representations of basin response to precipitation and snowmelt still remain a major challenge to hydrological research. Here, two models (SRM & ANN) are applied and compared for snowmelt runoff simulation of the Sharda river basin, which is a large (15280 km^2) trans-boundary central Himalayan River basin located within the Ganges River basin with 34% of its area being in Nepal. Observed data of daily precipitation, temperature and discharge as well as daily precipitation data from the WATCH dataset for the upper reaches of the catchment are used for model simulation. Snow cover data are derived from the MODIS dataset. Results show that the ANN technique outperforms conceptual modelling (SRM); however, due to the former's black-box nature, ANNs are useful only for short-term forecasting despite their high simulation accuracy. On the other hand, the conceptual models (SRM) even with lower accuracies than ANN are suitable for impact studies of environmental changes (landuse, climate etc.) as well reservoir operating policies.

Keywords: Hydrological modelling, Mountainous basins, Snow melt, MODIS, India

1. Introduction

Optimal utilization of available water resources depends on the accuracy of stream flow estimation, which in turn depends on the degree of understanding, simulation and forecasting of precipitationto runoff. Accurate values of runoff are necessary for economic planning of river basin projects for conservation and utilization of water for different purposes.

The transformation of precipitation to runoff is a complex, dynamic, and highly non-linear process, which is affected by many and often inter-related, physical factors. Besides precipitation, other factors that influence runoff are distribution & duration of precipitation, temperature, initial soil moisture, infiltration,land use, evapotranspiration, and catchment geomorphology. It is extremely difficult to clearly understand the complex physical process of runoff generation from a combination of these factors (Sarangi *et al.*, 2005). In case of Himalayan Rivers, there is an added complexity of snowmelt component in the runoff. Numerous modelling options are available for continuous-time modelling of precipitation-runoff relationships ranging from the empirical and physically based distributed models to black box lumped models. The empirical models (e.g. thewidely used SCS curve number method (McCuen,1982)) are generally unable to capture the inherent non-linear dynamics in the process of precipitation-runoff transformation. Conceptual models are way of simulating streamflow time-series as a function of climate inputs and represent the physical precipitation-runoff processes in a simplified manner. Conceptual models range from relatively complex models which attempt to explicitly represent all known components of the river system (catchment) including equations of water balance and conservation of energy (sometimes called 'physically-based' models), to simple models which lump many of the components into a small number of conceptual storages considering only water balance and may have as few as two parameters (McIntyre & Al-Qurashi, 2009). The conceptual models and the physically-based models that are in vogue for simulation of the precipitation-runoff process generally make use of a number of parameters, of which, many are either not yet readily available in literature or are difficult to ascertain for catchments from different geographical and climatic regions. Therefore, these models do not find successful application in data scarce scenarios. Data-driven modelling techniques, namely, Artificial Neural Networks (ANNs) have gained significant attention in recent years. ANN models do not presume a detailed understanding of the inherent physical processes, are able to produce results with limited data and do not require detailed basin descriptions or specific field measurements (Rajurkar *et al.*, 2004). Through comparative studies, it has been demonstrated that for hydrological problems that do not require explicit knowledge of the underlying hydrological processes, ANNs provide more efficient solutions than conventional approaches (Hsu *et al.*, 1995).

This study compares two different approaches to continuous-time modeling (semi-distributed conceptual modelling and black box ANN modeling) of daily precipitation-runoff response for the Sharda river basin, and compares the two techniques in terms of their statistical performance and applicability.

2. The Study Area and Data Used

In the present study, the Sharda River basin upto NHPC's Tanakpur barrage (Figure 1) has been taken up as the study area. It is a trans-boundary major sub- tributary of the river Ganges called Mahakali in Nepal. The ShardaRiver or Mahakali river demarcates Nepal's western border with India. The area of the Sharda basin upto Tanakpur barrage is 15280 km^2 (34% in Nepal), with elevation ranging from 250 to 7000m.Within the study area, the river flows mostly in the hilly area and emerges into plains at Bramhadeo, 5 km upstream of the basin's Tanakpurthe outlet.

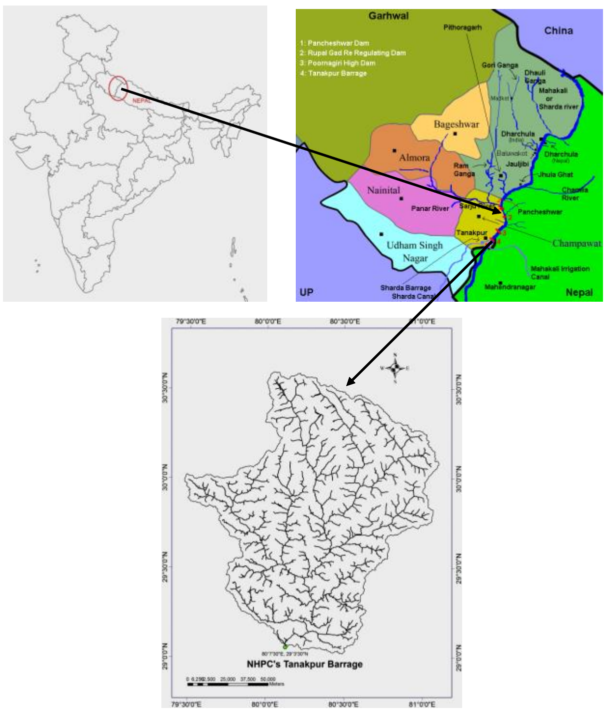

Figure 1. Study area

The Sharda river basin is characterized predominantly by hilly terrain, deep gorges and river valleys. The region broadly falls under four major divisions: (i) the Great Himalayan Ranges (snow-covered regions), (ii) Alpine and pasture land (covered by snow during four months of the winter season), (iii) Middle Himalaya (characterized by highest population) and (iv) river valleys (characterized by service centres and institutions). Permanent snow occurs above 5000m elevation. The basin has a diverse climate;the Southern Terai plain is sub-tropical while the Siwaliks are sub-tropical to warm temperate. In the Mid-hills, climate varies between warm temperate and cool temperate, whereas high Himalaya experiences a temperate to alpine or arctic climate. The basin experiences two rainy seasons, first in the summer (June –September), when the southwest monsoon brings about 75% of its total annual precipitation; and in the winter, accounting for the remaining precipitation. The average annual precipitation in the lower and middle part of the basin is about 2000mm. The average annual discharge of the basin upto Tanakpur barrage is about 715 m^3/sec.The Sharda Valley in Uttrakhand has a vast potential for Water Resources Development.

Observed daily discharge data at the Tanakpur Barrage on the Sharda River, dailyprecipitation data for the Indian part of the Sharda river for two NHPC's meteorological stations (i.e. at Tanakpur barrage and Dhauliganga dam), anddaily observed maximum and minimum temperature data at NHPC's Dhauliganga Dam have been procured for 5 years (2006 – 2010) from the Design & Engineering Division (D&E Div.), NHPC Corporate Office, Faridabad, Haryana, India and used in the present study. For the Nepal portion of the basin, no observed precipitationgauge data could be available, and therefore, gridded data fromthe WATCH Project (Weedon *et al.* 2011) has also been used for the upper reaches of the basin.

The elevation data used in this study is the 90 m resolution (3-arc SRTM).The data are freely available at: http://seamless.usgs.gov/Website/ Seamless/.

For snow cover mapping in Sharda basin, the Moderate Resolution Imaging Spectroradiometer (MODIS) snow cover products (Riggs *et al.* 2007) from 2006 to 2010 covering the Himalayan range in and around the basin have been used. The snow product used in this study is MODIS/Terra Snow Cover 8-Day L3 Global 500 m Grid (MOD10A2), Version 5 (Riggs *et al.*, 2007), which is available from 24 February, 2000 onwards. Spatial resolution of MOD10A2 is 500×500 m^2 in every 8-day period that begins on the first day of each year and continues to first few days of the next year. This product represents the maximum snow extent in the given 8-day period.

3. Brief Description of Modelling Techniques and Application

The following sections provide a brief description of the two modelling techniques used for runoff simulation in the present study.

3.1. Semi-distributed Conceptual Model: Snowmelt-Runoff Model (SRM)

Snowmelt-Runoff Model (SRM) is a degree-day based deterministic conceptual model developed to simulate and forecast daily streamflow in mountainous basins where snowmelt is a major runoff component (Martinec *et al.*, 2008; Wang and Li, 2006). In the present study SRM's windows version which is known as WinSRM, has been used to simulate runoff in the Sharda River at NHPC's Tanakpur barrage. SRM uses snow cover information and meteorological data (daily precipitation and daily maximumand minimum or average temperature) as input. Details of WinSRM (modelling concept, governing equations and application) are available in Martinec *et al.*(2008).

In the present study, SRM has been calibrated for the Sharda basin up to the Tanakpur site. The basin has been divided into 10 elevation zones using SRTM data (Figure 2). Daily snow cover area (SCA) in each elevation zone has been estimated using the MODIS data. The spatial variation of SCA in the Sharda basin for the period 2006-2007 is presented in Figure 3. It can be seen that the maximum snow in the basin is observed around the international boundary between India and Tibet. The elevation in this zone varies from 4500 to 7000m. The lower portion of the basin falling in India is almost snow free. Maximum snow coverage is in the month of March.

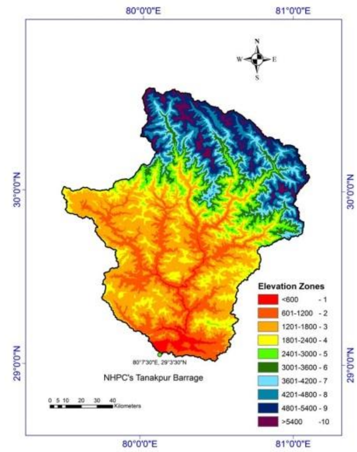

Figure 2.DEM of Sharda basin upto Tanakpur Barrage divided into 10 elevation zones

Observed daily precipitation data from two stations in the Indian part (Tanakpur and Dhauliganga) have been used. For the Nepal part and the higher elevations, high resolution gridded precipitation data from the WATCHdatasethas been used. Observed daily maximum and minimum temperature data in the Indian part at Dhauliganga have been used to calculate the daily mean temperature. For the upper parts, lapse rate has been used. In this way, each elevation zone has been allotted precipitation and mean temperature data on a daily basis. Discharge data available at the outlet of the basin (Tanakpur) have been used. The entire data base has been prepared for two time periods (calibration and validation), i.e., 2006-2009 and 2009-2010 respectively with October to September as the water year. Various model parameters have been optimised during model calibration. SRM gives the total snowmelt input to the basin for each day as an output table. From this table, the net snowmelt input (applying the snow runoff coefficient) from each zone has been calculated. As the SRM does not separate snowmelt contribution from the total runoff, we can consider that the net snowmelt input volume to the basin will be available at the basin outlet with some delay. Considering this, the average snowmelt contribution in the streamflow of the Sharda basin at Tanakpur barrage has been calculated by taking the ratio of average annual snowmelt runoff volume input to the basin and average annual total runoff volume computed at the basin outlet.

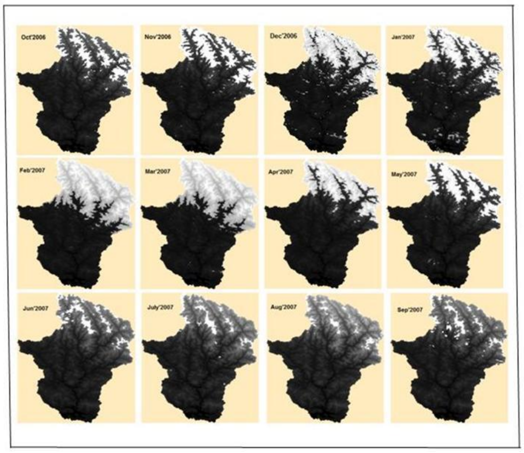

Figure 3. Sequential Snow cover in the Sharda basin as seen in MODIS images for the period Oct. 2006–Sep. 2007

3.2Black Box Model: Artificial Neural Network (ANN)

An ANN is a computing system made up of a highly interconnected set of simple information processing elements, analogous to a neuron, called units. The neuron collects inputs from both a single and multiple sources and produces output in accordance with a predetermined non-linear function. An ANN model is created by interconnection of many of the neurons in a known configuration. The primary elements characterizing the neural network are the distributed representation of information, local operations and non-linear processing.The main principle of neural computing is the decomposition of the input-output relationship into series of linearly separable steps using hidden layers. The theory of ANN has not been described here and can be found in many books such as Haykin (1994).

In the present study, a back propagation ANN with the generalized delta rule as the training algorithm has been employed. The structure of the ANN models was three layer back propagation ANN developed with non-linear sigmoid as activation function uniformly between the layers. Nodes in the input layer were equal to number of input variables, nodes in hidden layer were varied from the number of input nodes to approximately double of input nodes (Zhu, *et al.*, 1994) and the nodes in the output layer was one as the models provide single output. In this way, various ANN models were trained considering different hidden node numbers on a trial and error fashion and the best performing model has been reported in the results. Input output data used for ANN models is exactly similar to the SRM models for calibration as well as validation. The training of various ANN models has been accomplished through the ANN software, namely, Neural Power (NPP 2.5, 2004).

The training initiated with the normalization (re-scaling) of all input and output data with the maximum value of respective variable thus reducing the data domain in the range 0 to 1 and 0.1 to 0.9 respectively. This was accomplished through the software. All interconnecting links between nodes of successive layers were assigned random values called weight between +0.5 to –0.5 and a constant value of 0.15 and 0.8 was considered for learning rate and momentum respectively. The batch back propagation (BBP) learning algorithm has been adopted for training of all the precipitation-runoff ANN models because it produced the highest simulation accuracy. The network weights were updated after presenting each pattern from the learning data set, rather than once per iteration. At a point when average error of the network started to rise, the training was stopped. The performance of the model was tested through the statistical criterion discussed in the following section. Based on the best test results, ANN model was identified as the best performing model and used for the comparative study.

4. Performance Evaluation of Models

For the present study, the model performance on a daily basis has been evaluated using the non-dimensional coefficient of determination 'DC', also commonly known as Nash-Sutcliffe efficiency (Nash and Sutcliffe, 1970) as given by the equation:

$$DC = 1 - \frac{\left\{\sum_{t=1}^{n}(Q_0 - Q_e)^2\right\}}{\left\{\sum_{t=1}^{n}(Q_0 - \bar{Q}_0)^2\right\}} \qquad (1)$$

Where,

DC = Nash-Sutcliffe coefficient of goodness of fit

Q_o= daily observed discharge at basin outlet

Q_e= daily estimated discharge at basin outlet

$\bar{Q}_o$= mean of observed discharge

n = number of days of discharge simulation

The value of the coefficient of determination is analogous to the Nash-Sutcliffe efficiency and is a direct measure of the proportion of the variance of the recorded flows explained by the model.

The model performance has also be determined by computing the percentage volume difference between the measured and computed runoff as:

$$D_v = \frac{\left\{\sum_{t=1}^{n} Q_0 - \sum_{t=1}^{n} Q_e\right\}100}{\left\{\sum_{t=1}^{n} Q_0\right\}} \qquad (2)$$

Where,

D_v is the percentage volume difference between the observed and estimated discharge at basin outlet.

5. Results and Discussion

The results of daily runoff simulation for the calibration and validation period for both the models are shown in Table 1.

Table 1. Performance of Models

Model	Training (Calibration)			Testing (Validation)		
	DC	Dv (%)	Snowmelt (%)	DC	Dv (%)	Snowmelt (%)
SRM	0.79	-5.37	16.4	0.77	9.77	16.2
ANN	0.83	-16.8	N.A.	0.79	4.41	N.A.

It can be seen from Table 1 that the DC value is higher for the ANN model in the calibration phase and lower in the validation phase as compared to theSRM performance. However, the performance based on the Dv criterion is reversed, with ANN showing higher overestimation during calibration and lower underestimation in the validation phase. Nevertheless, the SRM model is capable of producing an quantitative estimate of the snowmelt input of the basin due to its conceptual model representation(note that is not possible with ANN models).

Further, a graphical comparison between the two modelsis shown in Figure 4. This figure presents theobserved runoff series and model simulated runoff series based on the SRM and ANNmodels during the calibration and validation periods. Figure 5 shows the evaluation (scatter) plots of the two techniques for the calibration and validation phases.

From Figure 4 and 5, it is clear that the ANN model generates more accuratepredictions of the observed runoff compared to SRM model, especially for the peak runoff. However, in the lower range of runoff, SRM simulation is comparatively closer to the observed values whilst the ANN simulations are overestimated. The reason could be the sudden high discharge in the river duringfew days in 2008, 2009 and 2010 due to heavy precipitationevents. SRM is unable to model these high values but ANN gives closer simulation mainly due to the nonlinear transformation process involved.

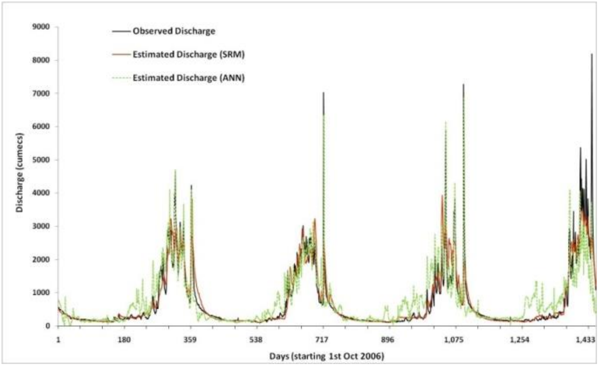

Figure 4. Comparison of observed and simulated hydrographs of SRM and ANN

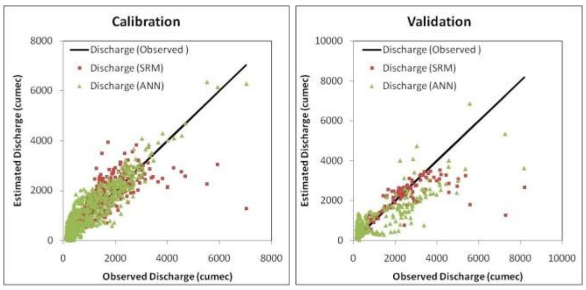

Figure 5. Evaluation(scatter) plots for SRM and ANN

6. Conclusions

This paper investigates the snowmelt runoff processesin a trans-boundary Himalayan sub-tributary of the Ganges River in India, namely Sharda River basin using two different modelling techniques namely semi-distributed conceptual modelling in the form of SRM model and Black-box modelling in the form of ANN model. Runoff simulation has been carried out on a daily time scale with similar input-output data for both models. A comparison has been carried out between the modelling techniques. The comparative analysis of the two techniques clearly demonstrates the superiority of the ANN technique over conceptual modelling (SRM). Although ANN models produce higher accuracy of runoff estimation compared to SRM, the advantage of the latter model is in the estimation of the snowmelt component of runoff. The advantage of such runoff segregation is manifolds in water resources sector. Therefore, the two techniques, can be used in a complementary manner for modelling precipitation-runoff response in large Himalayan River basins; for example, ANN models could be very useful for real time flood forecasting whereas the SRM model would be ideal for use in developing operating policies for a hydropower project.

Acknowledgements

The authors gratefully acknowledge NHPC, India for providing the hydro-meteorological data for the present study. The work presented in this paper is part of a collaborative international research project and the authors are also thankful to the Swedish Research Council (VR) for funding this investigation through the WaterRain- Him project (348-2014-33).

References

Haykin S. (1994), Neural Networks - a Comprehensive Foundation, Macmillan, New York.

Hsu K-L., Gupta H.V. and Sorooshian, S. (1995), Artificial Neural Network Modelling of the Rainfall-Runoff Process, *Water Resources Research*, **31(10),** 2517-2530.

Martinec, J., Rango, A., Roberts, R. (2008). Snowmelt Runoff Model (SRM), User's Manual, (Updated Edition 2008, Windows Version 1.11*),* USDA Jornada Experimental Range, New Mexico State University, Las Cruces, NM 88003, U.S.A.

McCuen R. H. (1982), A Guide to Hydrologic Analysis using SCS Methods, Prentice-Hall, Englewood Cliffs, NJ.

McIntyre N. and Al Qurashi A. (2009), Performance of Ten Rainfall-Runoff Models Applied to an Arid Catchment in Oman, *Environmental Modelling and Software*, **24,** 726-738.

Nash J.E. and Sutcliffe J.V. (1970), River Flow Forecasting through Conceptual Models, *Journal of Hydrology*, **10**, 282-290.

Neural Power (2004), Neural Networks Professional Version 2.5, CPC-X Software, Copyright: 1997-2004.

Rajurkar M.P., Kothyari U.C. and Chaube U.C. (2004), Modelling of the Daily Rainfall-Runoff Relationship with Artificial Neural Networks, *Journal of Hydrology*, **285**, 96-113.

Riggs G.A., Hall D.K. and Salomonson, V.V. (2007), MODIS snow products user guide to collection 5, doi:modis-snow-ice.gsfc.nasa.gov/sug_c5.pdf.

Sarangi A., Madramootoo P.E., Prasher S.O. and Patel R.M. (2005), Performance Evaluation of ANN and Geomorphology-based Models for Runoff and Sediment Yield Prediction for a Canadian Watershed, *Current science*, **89**(12).

Wang, J., and Li, S.(2006) Effect of climatic change on snowmelt runoffs in mountainous regions of inland rivers in Northwestern China. *Science in China-Series D: Earth Sciences*, **49**, 881-888.

Weedon GP, Gomes S, Viterbo P, *et al*. (2011) Creation of theWATCH forcing data and its use to assess global and regional reference crop evaporation over land during the twentieth century. *J Hydrometeorol*, **12**, 823– 848. doi:10.1175/2011JHM1369.1

Zhu, M., Fujita, M., and Hashimoto, N. (1994) Application of neural networks to runoff prediction. Stochastic and Statistical Method in Hydrology and Environmental Engineering, vol. 3, K.W. Hipel *et al*., eds., Kluwer, Dordrecht, The Netherlands, 205-216.