IDRiM (2019) 9 (2) **ISSN: 2185-8322**
DOI10.5595/idrim.2020.0357

Journal of Integrated Disaster Risk Management

Original paper

Improving Flood Forecasting in Karnali River Basin of Nepal Using Rainfall-Runoff Model and Complementary Error Model

Dilip Kumar Gautam [1] and Sumit Dugar [2]

Received: 19/05/2019 / Accepted: 26/11/2019 / Published online: 20/01/2020

Abstract Accuracy of flood forecast is important to take appropriate preparedness measures for saving lives and livelihoods of people residing in the floodplains. Predictions from flood forecasting models are usually uncertain which can be improved by complementing the hydrological model with an error model that can capture the information which the operational hydrological model lacks. This paper presents the application of this approach for improving daily flow forecasts for flood warning in Karnali River Basin of Nepal. A conceptual rainfall-runoff model, TUWmodel has been developed to model the rainfall-runoff processes and to predict the runoff at the outlet of the basin at Chisapani. The model has been calibrated for the period 2008-2011 with Nash-Sutcliffe Efficiency (NSE) 0.91 and percent bias (PBIAS) -0.7% and validated for the period 2012-2014 with NSE 0.88 and PBIAS -9.1% using observed temperature, precipitation and discharge data. A complementary ARIMA error model was developed from the error series for calibration set using automatic procedure and the predicted discharges were corrected using the error predictions from the error model. After error corrections, NSE and PBIAS were 0.95 and 0.1% respectively for calibration and 0.92 and 0.1% respectively for validation indicating significant improvements in the skill of forecasts.

Key words: Flood forecasting, error modeling, flood warning, hydrological model, ARIMA

1. INTRODUCTION

Hydrological models are important component of a flood forecasting system for estimating future flood discharge and assess the extent of inundation (Grange et al., 2015). Hydrologists have

[1] Practical Action Consulting, PO Box 15135, Kathmandu, Nepal, email: dilipgautam65@gmail.com
[2] Practical Action Consulting

used various types of hydrological models which can be classified based on process description, spatial representation or aspect of randomness (Todini, 2007). Based on process description, the models can be classified as physically based, conceptual and empirical data-driven models.

Lumped conceptual hydrologic models are the most common in research and used for operational flood forecasting (WMO, 2011). These models are based upon simple conceptual representation of the hydrologic cycle in the catchment with a structure of interconnected storages that represent overland flow, interflow and baseflow. Customization of such models involves estimating the model parameters by calibration with observed hydrologic data such as river level or discharge. The operational application of such model depends upon their capacity to reproduce observed hydrologic behavior of the catchment such as the rainfall–runoff processes. Examples of such models are NAM model (DHI, 2011), HBV model (Bergstörm, 1976), Soil Conservation Services (SCS) Curve Number method (Maidment, 1993), LCM model (Li and Liu, 2017), TANK model (Sugawara, 1995), SACRAMENTO model (Burnash, 1995), ARNO model (Todini, 1996) and the XINANJIANG model (Zhao et al,1980).

Amongst the various hydrologic processes, information on flood peak and time to the peaks, are important to take appropriate preparedness measures to save lives and livelihood assets. The flood peak information is used as input to the hydraulic model to estimate flood depth and discharge at various locations along the channel reach and floodplain (DHI, 2011, Pingping et al, 2018). Application of hydrologic models for flood forecasting is usually constrained by different sources of uncertainty such as inadequacy in model structure, incorrect model parameters, unreliable meteorological forecast or erroneous data (Grange et al., 2015). Consequently, the models fail to reproduce the observed hydrologic processes, hydrographs and flood peaks, and the inability to predict future floods accurately might have negative consequences in informing relevant stakeholders who make decisions based upon forecasts. Estimation of future flood peaks involves estimating the actual initial state of the basin, forecasting the inputs, and describing the different hydrologic processes that might provide an increased lead time. Ultimately, the quality of flood forecasts depends upon the accuracy and methodology applied whilst implementing each of these aforementioned aspects.

A data-driven time series model can be employed to enhance the prediction of floods by a conceptual model. Here, a calibrated conceptual model acts as the basic model that approximately captures the dominant hydrologic processes and forecasts the behavior of the catchment, results of which are deterministic in nature (Grange et al., 2015). A time series model can then be formulated on the errors. By analysing the error series, important information not captured by the conceptual model can be extracted which can be used for improving the prediction skill of a conceptual model. In this study, we used this technique for enhancing the performance of TUWmodel for flood forecasting in the Karnali River, Western Nepal.

Data-driven models based on the errors from a conceptual model can reveal whether the

conceptual model is adequate to identify essential relationships between the input–output data series (Kachroo, 1992). Data-driven error models can capture the persistency in the time series, which the conceptual model is unable to capture. Thus, the data-driven models can complement the conceptual model to improve the output (Serban and Askew, 1991).

Many researchers have demonstrated the application of data-driven model to improve the accuracy of conceptual models. Gragne et al. (2015) have applied complementary error modeling framework to improve real-time inflow forecasting into hydropower reservoirs in Norway by utilizing HBV model. Likewise, Morawietz et al. (2011) have evaluated different versions of autoregressive error models as post-processors for probabilistic streamflow forecasts. Similarly, Abebe and Price (2003) have applied artificial neural network model to improve the predictions of the conceptual rainfall-runoff model for the Sieve Basin in Italy. Liu et al. (2012) have also provided a comprehensive review of the data assimilation and error modeling in hydrologic forecasting.

The novelty of this research is that differs from the past work on the automatic estimation of parameters from the conceptual rainfall-runoff model, data-based ARIMA error model and Box-Cox transformation. Gragne et al. (2015) used already calibrated HBV model and the parameters of error model have been estimated by least squares method using an iterative algorithm. We applied Differential evolution (DE) global optimization algorithm available in R package **DEoptim** to automatically calibrate the model parameters. The parameters of error model and Box-Cox transformation have also been estimated automatically using R package '**forecast**'. The automatic procedure facilitates the application of this approach in real-time flood forecasting in an effective and efficient way.

The bias, persistence and heteroscedasticity present in the errors reflect structural inadequacy of the conceptual model to capture the complete catchment processes and, hence, are critical in defining the structure of the data-based error model. Here, we describe the errors in a transformed space with Box-Cox transformation (Box and Cox, 1964) and estimate the data-driven error model and the transformation parameters with an automatic procedure using R package '**forecast**' (See https://CRAN.R-project.org/package=forecast).

In the next section, we describe how the conceptual rainfall-runoff model and complementary error models are set-up. An example application in the Karnali River, Western Nepal is presented in the subsequent section which includes description of the study area, data used, findings from the evaluation of the complimentary error model and its application during calibration and validation, and results of forecasting capabilities.

2. STUDY AREA AND DATA

The Karnali is a perennial, trans-boundary river that originates from Tibet, crosses the dissected hill slopes of Western Nepal and flows into India. Tributaries of Karnali are the snow-fed rivers such as Mugu Karnali and Humla Karnali (Figure 1). The West Seti River (202 km) and the Bheri River (264 km) are the other major tributaries of the Karnali. It enters the lowlands of the Terai plains via a spectacular gorge at the town of Chisapani, where its catchment area is 45,583 km^2. The Karnali then remarkably divides into two main channels, Geruwa on the left and Kauriala on the right few kilometers downstream of Chisapani.

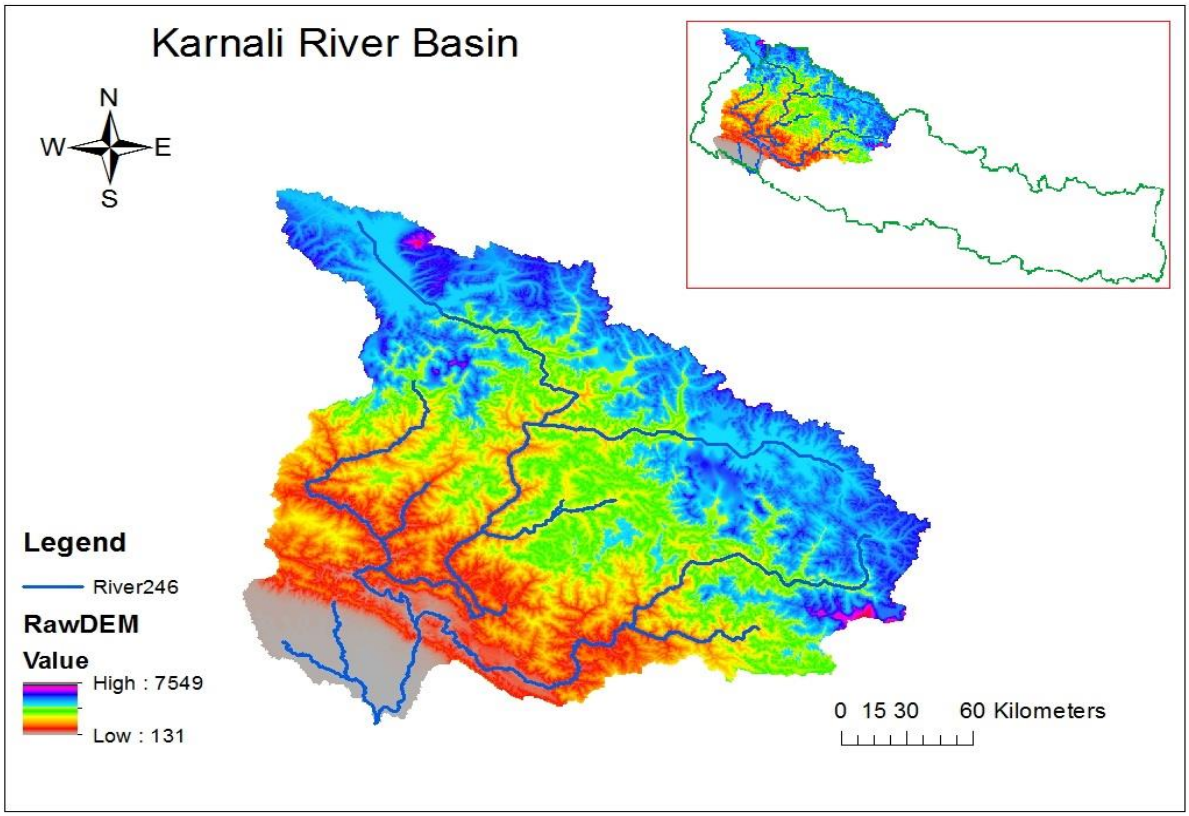

Figure 1. Karnali River Basin

The Terai plain downstream of Karnali is fertile and most of the alluvial plains are devoted to agriculture. In recent years, the development activities have increased with the construction of roads, irrigation, flood control embankments, electricity and telecommunication infrastructures downstream of Chisapani in the Terai belt. Many people reside alongside the floodplains because

of the economic opportunities which have inadvertently increased vulnerabilities due to flooding. Intense monsoon rainfall that triggers the floods turns into disaster because of high number of communities residing along the floodplains within the close proximity of the river. For example, the Karnali floods in 15 August, 2014 affected 173,800 people with 29,680 people displaced and 53 people killed. More than 1,240 houses were destroyed and 435 houses damaged.

.

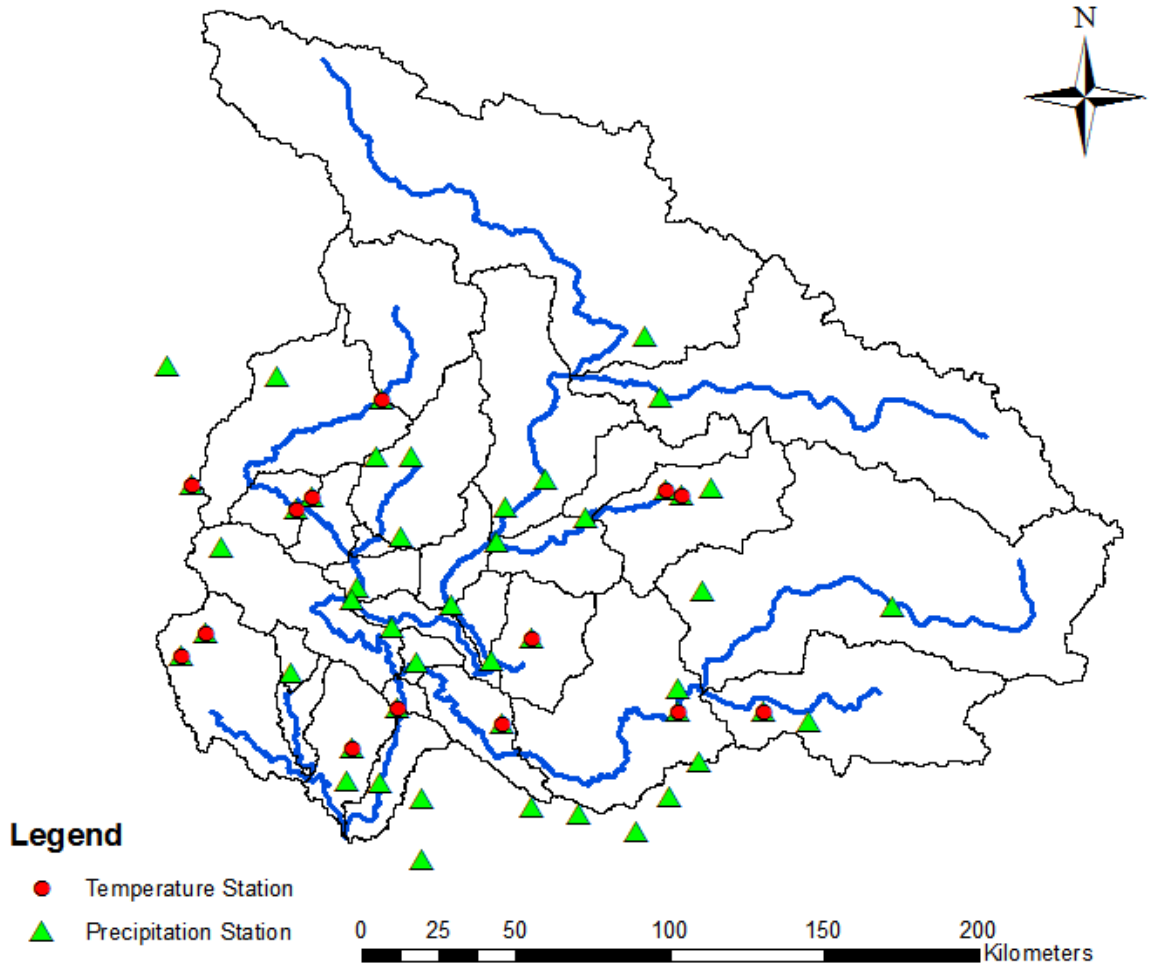

Figure 2. Precipitation and temperature stations

For this research, we used the observed discharge data of Chisapani station for 7 years (01/01/2008-31/12/2014) obtained from the Department of Hydrology and Meteorology of Nepal. Similarly, observed daily precipitation data of 47 stations and daily air temperature data of 14 stations have also been obtained for the same period. In the operational forecasting mode, the air temperature and precipitation input over the forecast lead time will be obtained from the Numerical Weather Prediction (NWP) model run by Meteorological Forecasting Division of the Department

of Hydrology and Meteorology of Nepal (http://mfd.gov.np/nwp/). As this study aims to improve hydrologic forecasts for flood forecasting by complementing the conceptual model by an error model, we assume that the predictions from the TUWmodel are made using the best possible input data. Hence, the observed air temperature and precipitation data are used for flood forecasts in hindcast mode. In operational setting with NWP forecast, the error model needs to be recalibrated. The basin average precipitation and temperature were computed by Thiessen polygon method and arithmetic average method respectively. Figure 2 depicts the location of precipitation and temperature stations whereas basin average precipitation, temperature and maximum daily discharge at Chisapani for the period 2008-01-01 to 2014-12-31 are presented in Figure 3.

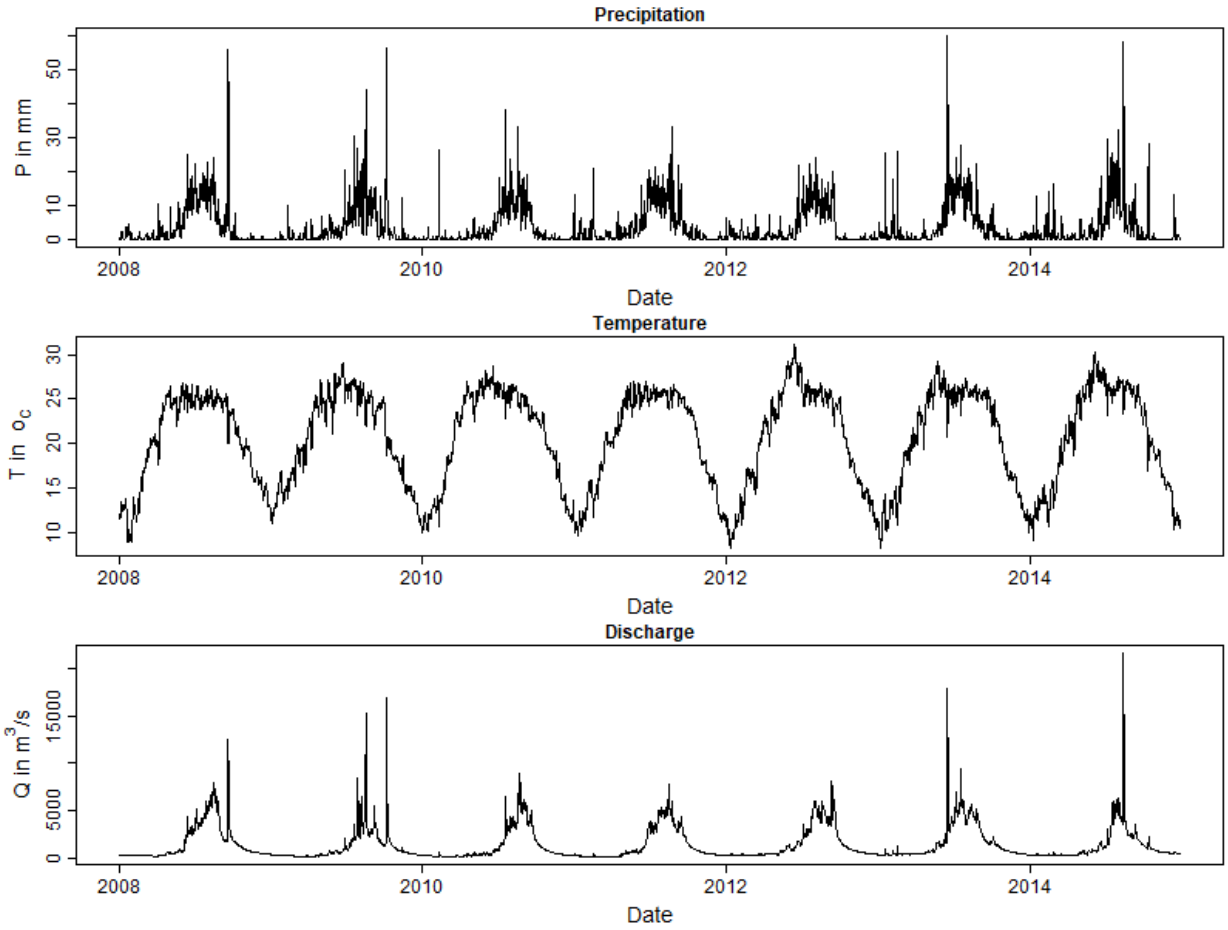

Figure 3. Basin average precipitation, temperature and maximum daily discharge at Chisapani

The modified Blaney-Criddle formula is one of the simplest methods for calculating reference evapotranspiration (Blaney and Criddle, 1950) which requires only mean daily temperature data. The original Blaney-Criddle formula is given by:

$$ET = p.(0.46.T + 8) \quad (1)$$

The modified Blaney-Criddle formula is as follows:

$$ET = -1.55 + 0.96\,p.(0.457.T + 8.128) \quad (2)$$

Where, ET = Reference evapotranspiration (mm/day)

T = mean daily temperature (°C)

p = mean daily percentage of annual daytime hours (it is function of latitude)

The mean daily percentage of annual daytime hours for 25-30 degree north latitude is given in Table 1. TUWmodel requires potential evapotranspiration as input which is calculated by modified Blaney-Criddle method using mean daily temperature data for the Karnali. R Codes have been written for this purpose.

Table 1. Mean daily percentage of annual daytime hours for 25-30 degree north latitude

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
p	0.24	0.255	0.27	0.29	0.305	0.315	0.31	0.295	0.28	0.26	0.245	0.235

The observed data were divided into calibration and validation set. Calibration and validation periods have been specified from 01/01/2008 to 31/12/2011 and 01/01/2012 to 31/12/2014 respectively. Both calibration and validation data set consisted of several flood peaks. There have been major flood events with the water level crossing the danger level in 2008, 2009, 2013 and 2014. The validation set consisted of the historically record flood peak on August 2014 with the recurrence interval of 1 in 100 years.

3. METHODOLOGY

3.1 Conceptual Rainfall-Runoff Model

Here, we have selected a lumped conceptual rainfall-runoff model **TUWmodel** for this research. **TUWmodel** follows the structure of the HBV (Hydrologiska Byråns Vattenbalansavdelning) model which runs on a daily or shorter time step and consists of a snow routine, a soil moisture routine and a flow routing routine (See Parajka, Merz and Bloeschl, 2007, Merz and Bloeschl, 2004). **Snow Routine**:

The snow routine represents snow accumulation and melt by a simple degree-day concept, using a degree-day factor DDF and a melt temperature parameter Tm. The catch deficit of precipitation

gauges during snowfall is corrected by a snow correction factor SCF. A threshold temperature interval Tr – Ts is used to distinguish between rainfall, snowfall and a mix of rain and snow. Mean daily precipitation P in an elevation zone is partitioned into rain P_r and snow P_s based on the mean daily air temperature T_a.

.

$P_r = P$ If $T_a \geq T_r$ (3)

$P_r = P\frac{T_a - T_s}{T_r - T_s}$ If $T_s < T_a < T_r$ (4)

$P_r = 0$ If $T_a < T_s$ (5)

$P_s = P - P_r$ (6)

Where T_a is air temperature, T_s is lower threshold temperature below which precipitation is snow and T_r is upper threshold temperature above which precipitation is rain. Melt starts at temperature above a threshold T_m.

$M = (T_a - T_m)\,DDF$ If $T_a > T_m$ and SWE > 0 (7)

Where, M is the amount of melt water per time step, DDF is the degree-day factor and SWE is the snow water equivalent. The catch deficit of the precipitation gauges during snowfall is corrected by a snow correction factor SCF. Changes in the snow water equivalent from days i -1 to i are accounted by Equation (8) below.

$$SWE_i = SWE_{i-1} + (SCF.P_s - M)\Delta t \quad (8)$$

Where SCF is snow correction factor and Δt is the time step.

Soil Moisture Routine:

The soil moisture routine represents runoff generation and changes in the soil moisture state of the catchment and involves three parameters: the maximum soil moisture storage FC, a parameter representing the soil moisture state above which evaporation is at its potential rate, termed the limit for potential evaporation LP, and a parameter in the non-linear function relating runoff generation to the soil moisture state, termed the non-linearity parameter β.

$S_{SM,\,i} = S_{SM,\,i-1} + P_r + M - E_a$ (9)

Where, S_{SM} is the soil moisture of a top soil layer controlling runoff generation and actual evaporation E_a. The contribution ΔS_{UZ} of rain and snowmelt to runoff is calculated by an explicit scheme as a function of the soil moisture of the top layer S_{SM} using a non-linear relationship with two free parameters, FC and β:

$$\Delta S_{UZ} = (\frac{S_{SM}}{FC})^{\beta} (P_r + M) \quad (10)$$

FC is the ma imum soil moisture storage. The parameter β controls the characteristics of runoff generation and is a non-linearity parameter. If the top soil layer is saturated, i.e. S_{SM} = FC, then all rainfall and snowmelt contribute to runoff. The actual evaporation E_a is calculated from potential evaporation E_p by a piecewise linear function of the soil moisture of the top layer:

$$E_a = E_p \frac{S_{SM}}{LP} \quad \text{If } S_{SM} < LP \quad (11)$$

$$E_a = E_p \quad \text{If } S_{SM} \geq LP \quad (12)$$

Where, LP is a parameter termed the limit for potential evapotranspiration E_p.

Flow Routing Routine:

Runoff routing on the hill slopes is represented by an upper and a lower soil reservoir S_{UZ} and S_{LZ} respectively. Excess rainfall ΔS_{UZ} enters the upper zone reservoir and leaves this reservoir through three paths: outflow from the reservoir based on a fast storage coefficient K_1; percolation to the lower zone with a constant percolation rate C_P; and, if a threshold of the storage state LS_{UZ} is exceeded, through an additional outlet based on a very fast storage coefficient K_0. Water leaves the lower zone based on a slow storage coefficient K_2. The outflow from both reservoirs Q_G is then routed by a triangular transfer function representing runoff routing in the streams.

$$B_Q = B_{MAX} - C_R Q_G \quad \text{If } (B_{MAX} - C_R Q_G) \geq 1 \quad (13)$$

$$B_Q = 1 \quad \text{Otherwise} \quad (14)$$

Where B_Q is the base of the transfer (triangular) function, B_{MAX} is the maximum base at low flows and C_R is a free scaling parameter.

The model's input data are precipitation, air temperature, potential evapotranspiration and catchment area. For this research, we used the model as a lumped model with one parameter set and input data set for the entire catchment. The model could also be used as a semi-distributed model by dividing the catchment into sub-basins. The 15 model parameters are described in Table 2 below.

Table 2. Parameters of TUWmodel

S No.	Parameter	Description	Range	Unit
1	SCF	snow correction factor	0.9-1.5	
2	DDF	degree day factor	0.0-10.0	mm/degC/timestep
3	Tr	threshold temperature above which precipitation is rain	1.0-3.0	degC
4	Ts	threshold temperature below which precipitation is snow	-3.0-1.0	degC
5	Tm	threshold temperature above which melt starts	-2.0-2.0	degC
6	LP	parameter related to the limit for potential evaporation	0.0-1.0	
7	FC	field capacity, i.e., max soil moisture storage	0-600	mm
8	BETA	the nonlinear parameter for runoff production	0.0-20.0	
9	K0	storage coefficient for very fast response	0.0-2.0	timestep
10	K1	storage coefficient for fast response	2.0-30.0	timestep
11	K2	storage coefficient for slow response	30.0-250.0	timestep
12	LSUZ	threshold storage state, i.e., the very fast response start if exceeded	1.0-100.0	mm
13	CP	constant percolation rate	0.0-8.0	mm/timestep
14	BMAX	maximum base at low flows	0.0-30.0	timestep
15	CR	free scaling parameter	0.0-50.0	$timestep^2/mm$

Table 3 presents the model output vector. The model output consists of a vector of simulated runoff as q (mm/timestep), and 11 additional vectors.

Table 3. Model output vector

SN	Output	Description	Unit
1	q	simulated runoff	mm/timestep
2	qzones	simulated runoff for each zone	mm/timestep
3	q0	surface runoff	mm/timestep
4	q1	subsurface runoff	mm/timestep
5	q2	baseflow	mm/timestep
6	rain	liquid precipitation	mm/timestep
7	snow	solid precipitation	mm/timestep
8	melt	snowmelt	mm/timestep
9	moist	soil moisture	mm
10	swe	snow water equivalent	mm
11	suz	upper storage zone	mm
12	slz	lower storage zone	mm

The model is available freely as R library (See https://cran.r-project.org/package=TUWmodel) . Further coding is needed for processing input data, calibration and validation, forecasting and o utput generation. Codes have been written in R to customize the TUWmodel for Karnali River.

3.2 Calibration of Model Parameters

The performance of a hydrological model depends on how well its parameters are calibrated. Hence, the calibration process should be chosen carefully to maximize the performance of the model. Calibration of a hydrological model is a tedious process. A hydrological model can be calibrated using manual trial and error method, global optimization method or combination of both. For a hydrological model having many parameters, manual method can be extremely time consuming and uncertain to identify optimum parameter values. Alternatively, a global optimization algorithm can be effective and efficient. Differential evolution (DE) is a stochastic, population-based global optimization algorithm using crossover, mutation and selection operators

that is effective on many problems of interest in science and technology. DE is particularly well-suited to find the global optimum of a real-valued function of real-valued parameters, and does not require that the function be either continuous or differentiable. DE has been successfully applied in a wide variety of fields, from computational physics to operations research. Detailed description and R code of differential evolution can be found on the website: http://www1.icsi.berkeley.edu/~storn/code.html.

Optimum parameters of TUWmodel were found using DE optimization algorithm (Mullen et al., 2011) which allows box constraints, that is each variable can be given a lower and/or upper bound (Byrd et al., 1995). The initial value must satisfy the constraints. We used R package DEoptim to automatically calibrate the model parameters. The DEoptim package is available at https://cran.r-project.org/web/packages/DEoptim/index.html.

3.3 The Complementary Error Model

The error model aims at exploiting the bias, persistence and heteroscedasticity in the errors and estimating the errors is likely to occur in the forecast lead time. Forecasting the error in the lead time is regarded as a two-step process that comprises of offline identification and estimation of the error model, and error predictions based on most recent information.

In this study, we aim to fit Auto Regressive Integrated Moving Average (ARIMA) model for the error series (Brockwell and Davis, 1996). First, we compute error series as the difference between observed and predicted discharge at time t expressed as $Y_t = Q_t - \widehat{Q}_t$. Then the errors are analyzed to see whether these are random and stationary or show some trend and heteroscedasticity. If the series is non-stationary, then it is stationarized by differencing. This is then followed by assessment of the pattern of autocorrelation function (ACF) and partial autocorrelation function (PACF) to determine if lags of the stationarized series and/or lags of the forecast errors should be included in the forecasting equation. The autocorrelation function (ACF) plot shows the correlation of the series with itself at different lags. The partial autocorrelation function (PACF) plot shows the amount of autocorrelation at lag k that is not adequately explained by lower-order autocorrelations. The suggested error model is fitted and the error diagnostics are checked, particularly for the error ACF and PACF plots, to comprehend if all coefficients are significant and the entire pattern have been explained. Patterns that remain in the ACF and PACF may suggest the need for additional autoregressive (AR) or moving average (MA) terms.

Let *Y* denote the *original* series and *y* denote the *differenced* (stationarized) series.

No difference (d=0): $y_t = Y_t$ (15)

First difference (d=1): $y_t = Y_t - Y_{t-1}$ (16)

Second difference (d=2): $y_t = (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}) = Y_t - 2Y_{t-1} + Y_{t-2}$ (17)

The AutoRegressive Integrated Moving Average (ARIMA) error forecasting equation is given by $\hat{y}_t = \mu + \emptyset_1 y_{t-1} + \cdots + \emptyset_p y_{t-p} - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q}$ (18)

Where, μ is constant term, φ and θ are parameters of autoregressive (AR) and moving average (MA) terms and p and q are order of autoregressive and moving average terms (Nau, 2014).

Positive ACF at lag 1 indicates AR series and negative ACF at lag 1 indicates MA series. For the AR series, ACF dies out gradually and PACF cuts off sharply after a few lags. For the MA series, ACF cuts off sharply after a few lags and PACF dies out more gradually (Nau, 2014).

We used R package 'forecast' to estimate the orders and parameters of ARIMA model automatically (See http://github.com/robjhyndman/forecast). Version 7.2 of the package was used for this study (Hyndman and Khandakar, 2008). The auto.arima function in 'forecast' package estimates the best fit ARIMA model to univariate time series according to either Akaike's Information Criterion (AIC) or Bayesian Information Criteria (BIC) value.

We applied Box-Cox transformation to the error series (y) before ARIMA model is estimated (Box and Cox, 1964). A Box-Cox transformation is a way to transform non-normal dependent variables into a normal shape. The Box-Cox transformation is given by:

$$y(\lambda) = \begin{cases} \frac{y^{\lambda}-1}{\lambda}, & \lambda \neq 0 \\ logy, & \lambda = 0 \end{cases} \quad (19)$$

At the core of the Box-Cox transformation is an exponent, lambda (λ), which varies from -5 to 5. We used BoxCox.lambda function of R package 'forecast' to estimate the Box-Cox transformation parameter (λ) automatically by minimizing the coefficient of variation of the error series (Guerrero,1993).

3.4 Goodness-of-fit Measures

For evaluating the goodness-of-fit (GOF) measures of the model, the following 20 numerical measures are defined.

Table 4. Goodness-of-fit measures

GOF measures	Description
me	Mean Error
mae	Mean Absolute Error
mse	Mean Squared Error
rmse	Root Mean Square Error
nrmse	Normalized Root Mean Square Error (-100% <= nrms <= 100%)
PBIAS	Percent Bias
RSR	Ratio of RMSE to the Standard Deviation of the Observations, RSR = rms / sd(obs). (0 <= RSR <= +Inf)
rSD	Ratio of Standard Deviations, rSD = sd(sim) / sd(obs)
NSE	Nash-Sutcliffe Efficiency (-Inf <= NSE <= 1)
mNSE	Modified Nash-Sutcliffe Efficiency
rNSE	Relative Nash-Sutcliffe Efficiency
d	Index of Agreement (0 <= d <= 1)
md	Modified Index of Agreement
rd	Relative Index of Agreement
cp	Persistence Index (0 <= PI <= 1)
r	Pearson Correlation coefficient (-1 <= r <= 1)
R2	Coefficient of Determination (0 <= R2 <= 1). Gives the proportion of the variance of one variable that is predictable from the other variable
bR2	R2 multiplied by the coefficient of the regression line between sim and obs (0 <= bR2 <= 1)
KGE	Kling-Gupta efficiency between sim and obs (0 <= KGE <= 1)
VE	Volumetric efficiency between sim and obs (-Inf <= VE <= 1)

We used the R package hydroGOF to compute the goodness-of-fit measures (https://cran.r-project.org/web/packages/hydroGOF/index.html).

The Nash-Sutcliffe Efficiency (NSE) coefficient and Percent Bias (PBIAS) are the most important performance measures widely used to evaluate the model performance. The NSE and PBIAS can be defined respectively as follows:

$$NSE = 1 - \frac{\sum_{t=1}^{N}(Q_t^{pred} - Q_t^{obs})^2}{\sum_{t=1}^{N}(Q_t^{obs} - Q^{mean})^2} \quad (20)$$

$$PBIAS = \frac{\sum_{t=1}^{N}(Q_t^{pred} - Q_t^{obs})}{\sum_{t=1}^{N} Q_t^{obs}} \times 100 \quad (21)$$

Where Q_t^{pred} is predicted flow, Q_t^{obs} is observed flow and Q^{mean} is the mean observed flow.

PBIAS measures the tendency of the predicted flows to be larger or smaller than the observed flows. Hence, it gives a measure of mass conservation. The optimal value is 0.0, whereas positive value indicates a tendency of model overestimation and negative value indicates a tendency of model underestimation.

The Nash-Sutcliffe Efficiency coefficient (NSE) measures the fraction of the variance of the observed flows explained by the model in terms of the relative magnitude of the residual variance (noise) to the variance of the flows (information). It can range from $-\infty$ to 1. An efficiency of 1 (NSE = 1) corresponds to a perfect match of modeled discharge to the observed data. An efficiency of 0 (NSE = 0) indicates that the model predictions are as accurate as the mean of the observed data, whereas an efficiency less than zero (NSE < 0) occurs when the observed mean is a better predictor than the model or, in other words, when the residual variance is larger than the data variance. Essentially, the closer the model efficiency is to 1, the more accurate the model is. In general, the model simulation can be judged as satisfactory if NSE > 0.5 and PBIAS ±25% for streamflow (Moriasi et al, 2007).

Additionally, the time series plots of observed and simulated hydrograph and scatterplots are also compared. Visual inspection of simple hydrograph plots that compare the predictions to actual measurements in calibration and validation set can provide significant information about how much the predictions are close to the observations for different flow regimes.

4. RESULTS

The differential evolution optimization algorithm as described above was used for automatic calibration of 15 parameters of TUWmodel. The mean square error was defined as the objective function to minimize. The objective function converged after 356 iterations. Figure 4 below shows the objective function plot.

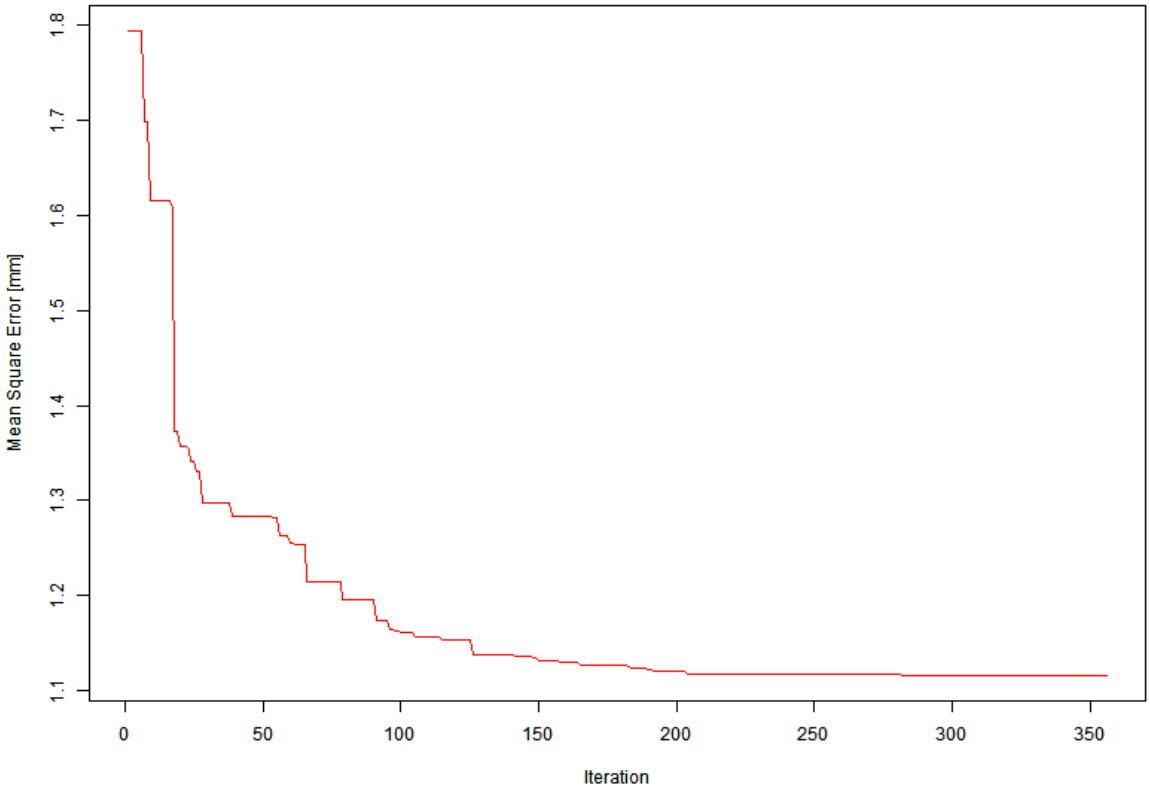

Figure 4. Objective function plot

The set of optimal parameter values obtained by automatic calibration are presented in Table 5 below.

Table 5. Optimal parameter values for calibration set

S No.	Parameter	Description	Range	Optimal value	Unit
1	SCF	snow correction factor	0.9-1.5	1.19	
2	DDF	degree day factor	0.0-10.0	3.35	mm/degC/ timestep
3	Tr	threshold temperature above which precipitation is rain	1.0-3.0	2.94	degC
4	Ts	threshold temperature below which precipitation is snow	-3.0-1.0	-2.49	degC
5	Tm	threshold temperature above which melt starts	-2.0-2.0	1.08	degC
6	LP	parameter related to the limit for potential evaporation	0.0-1.0	1	
7	FC	field capacity, i.e., max soil moisture storage	0-600	288.34	mm
8	BETA	the nonlinear parameter for runoff production	0.0-20.0	0.38	
9	K0	storage coefficient for very fast response	0.0-2.0	1	timestep
10	K1	storage coefficient for fast response	2.0-30.0	2.72	timestep
11	K2	storage coefficient for slow response	30.0-250.0	30	timestep
12	LSUZ	threshold storage state, i.e., the very fast response start if exceeded	1.0-100.0	32.83	mm
13	CP	constant percolation rate	0.0-8.0	5.01	mm/timestep
14	BMAX	maximum base at low flows	0.0-30.0	4.88	timestep
15	CR	free scaling parameter	0.0-50.0	32.81	$timestep^2/mm$

Figures 5 and 6 present the time series plot and scatterplot of observed and simulated flows at Chisapani station for calibration periods. The joint hydrograph of observed and simulated flows shows that the model is able to produce flow pattern nicely but it underestimates the peak flows. Figures 7 and 8 present the time series plot and scatterplot of observed and simulated flows at Chisapani station for validation periods. The model is able to produce flow pattern nicely but it underestimates the peak flows for the validation data set also. The performance of TUWmodel model was similar for calibration and validation period.

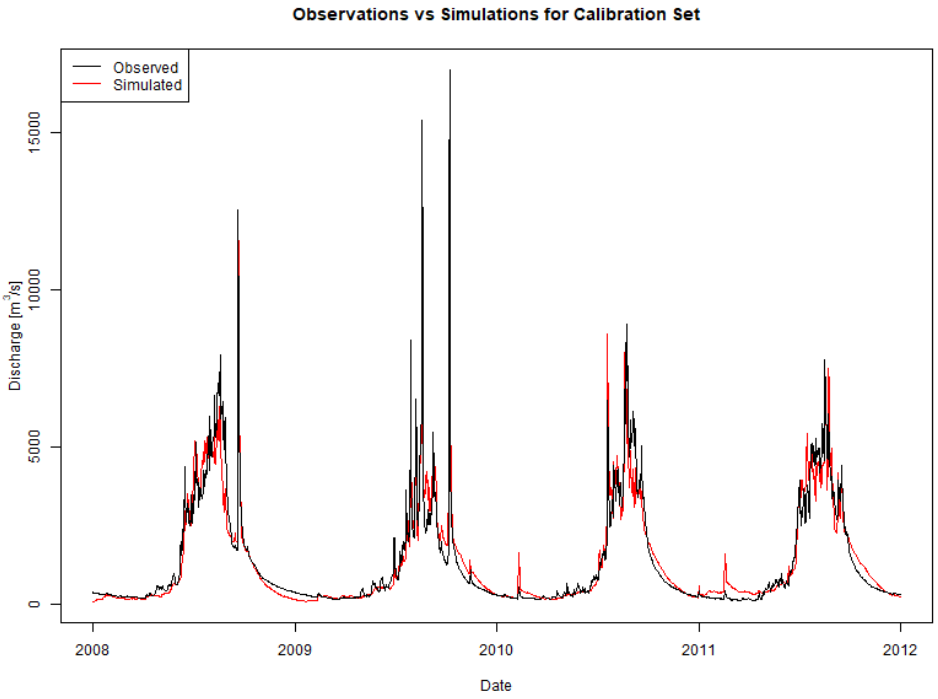

Figure 5. Time series plot for calibration set

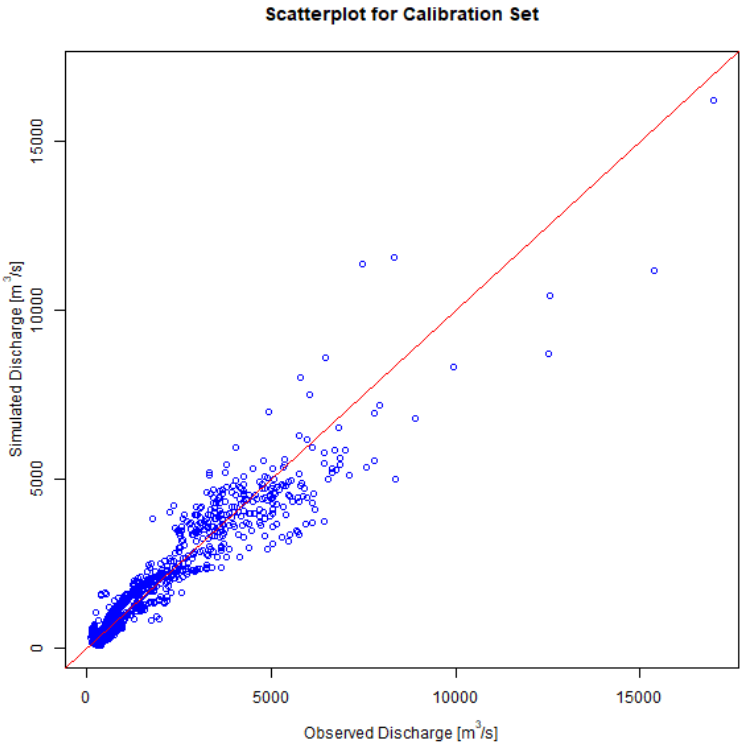

Figure 6. Scatterplot for calibration set

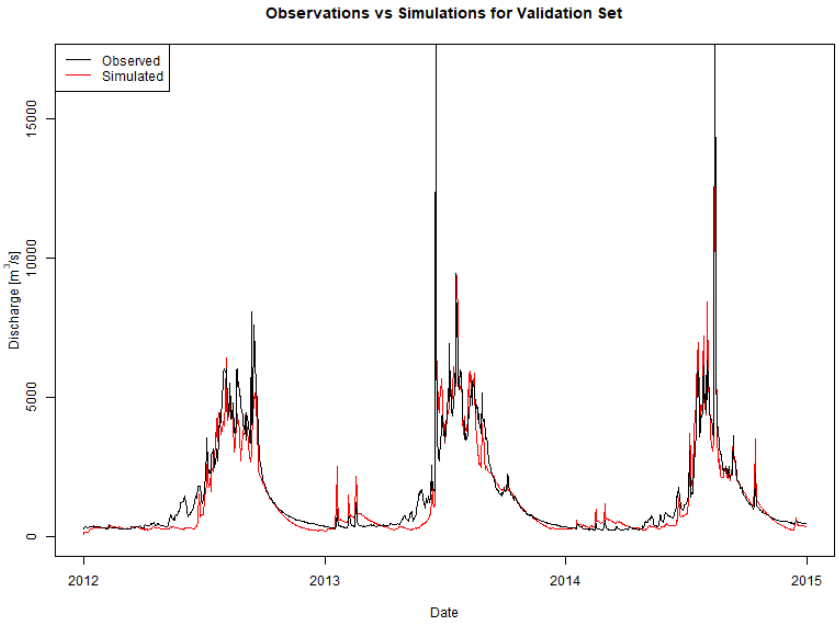

Figure 7. Time series plot for validation set

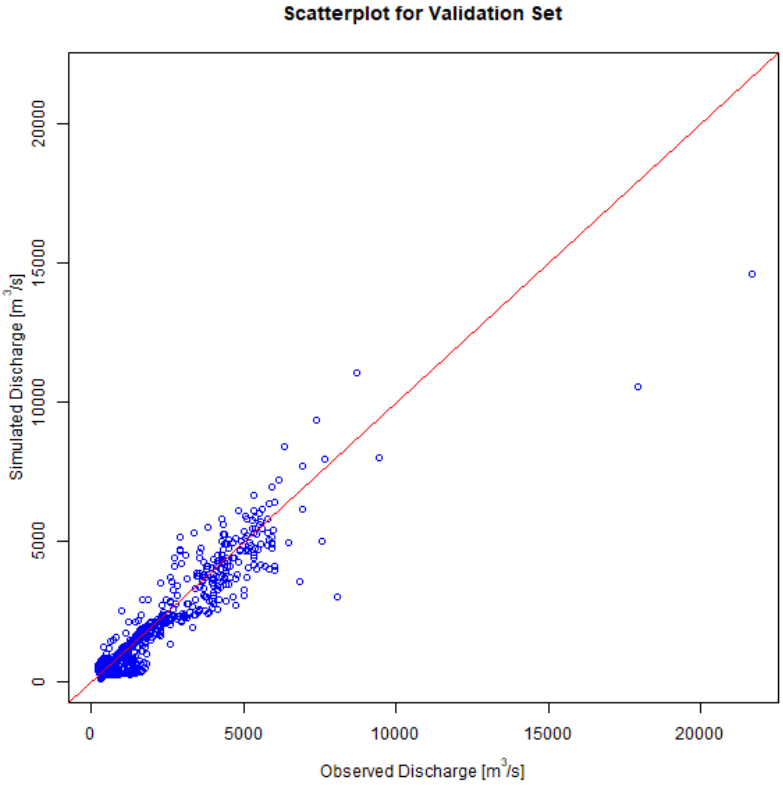

Figure 8. Scatterplot for validation set

The autocorrelation function (ACF) and partial autocorrelation function (PACF) of the error series have been analyzed. The error plot in Figure 9 shows the variability of errors. The error variations are high for high flows. The ACF and PACF plots in Figure 9 indicate the high degree of persistency and presence of both AR and MA terms.

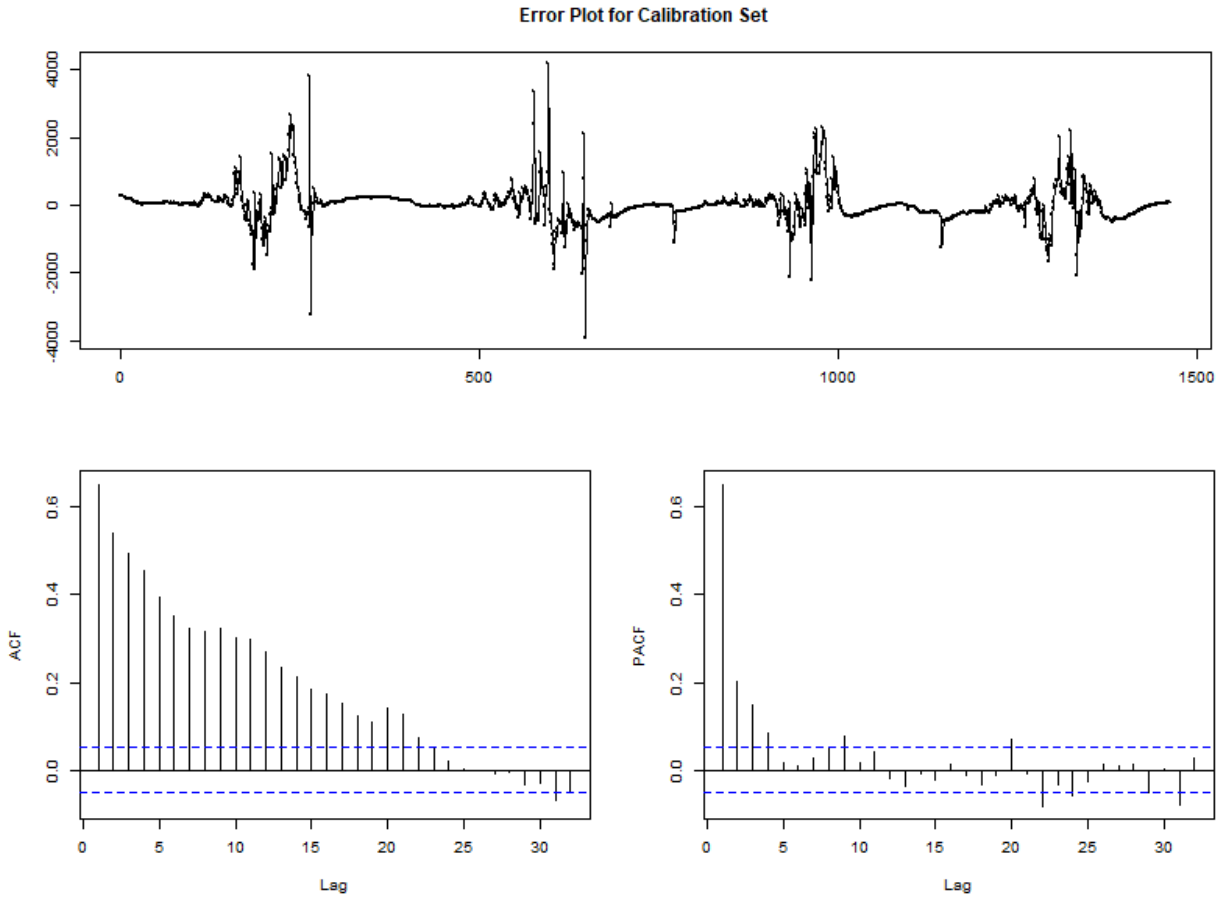

Figure 9. Error, ACF and PACF plot

The Box-Cox transformation parameter (λ) value was found to be 1.0463. The ARIMA model structure was identified as ARIMA (2,1,1) with coefficient $\varphi_1 = 0.3337$, $\varphi_2 = 0.0624$ and $\theta_1 = -0.8513$.

Figures 10 and 11 present the time series plot and scatterplot of observed and simulated flows at Chisapani station for calibration period after error correction. The figure clearly shows the improvement in predicting peak flows after error correction for calibration set. Figures 12 and 13 present the time series plot and scatterplot of observed and simulated flows at Chisapani station for validation period after error correction. It also shows the improvement in predicting peak flows after error correction for validation set.

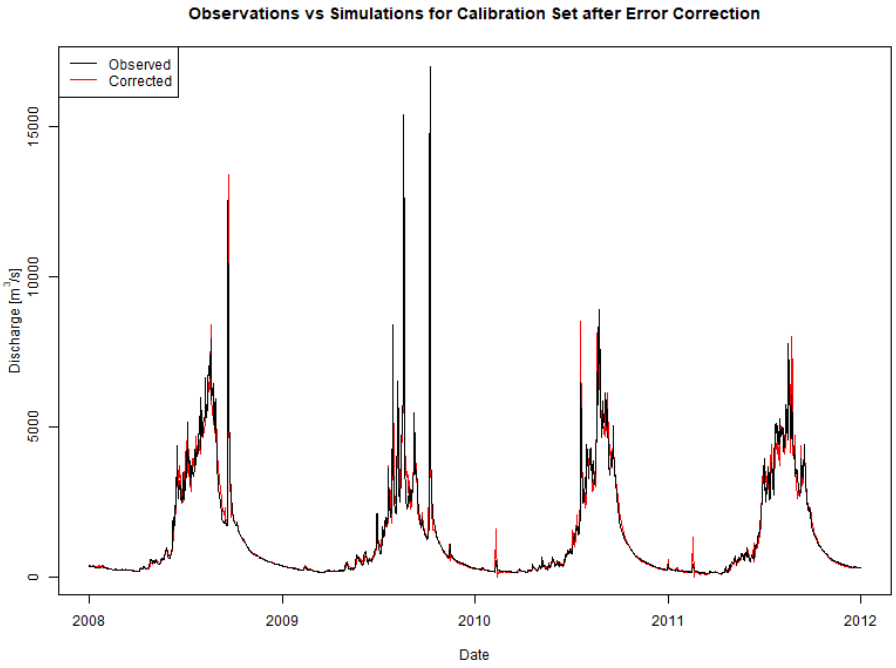

Figure 10. Time series plot for calibration set after error correction

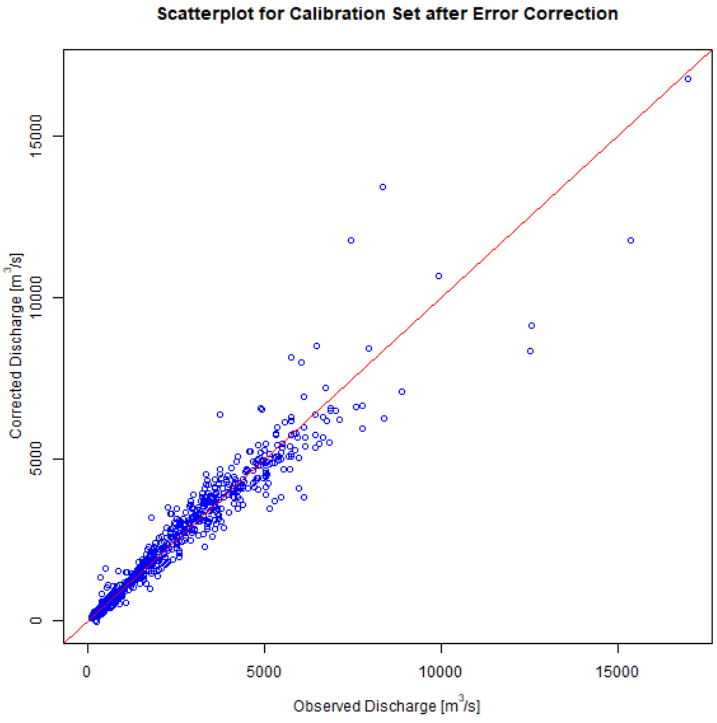

Figure 11. Scatterplot for calibration set after error correction

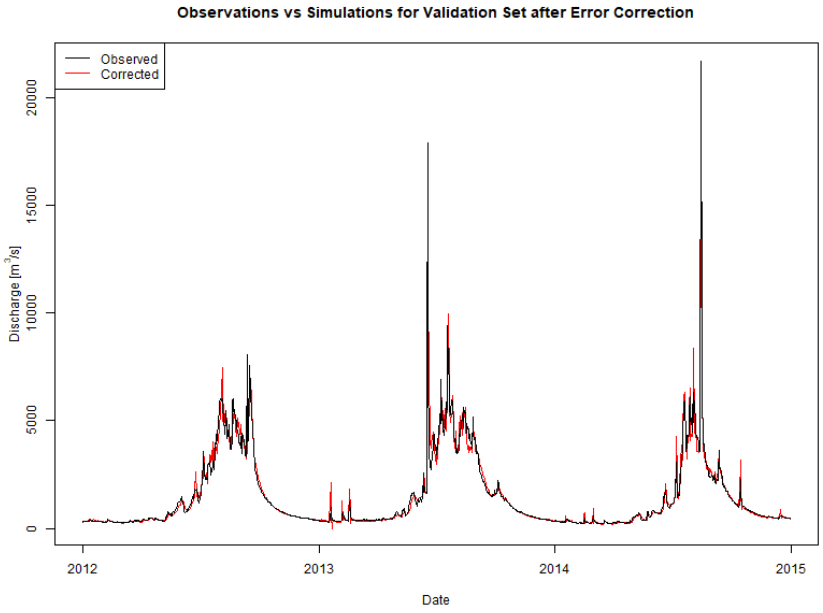

Figure 12. Time series plot for validation set after error correction

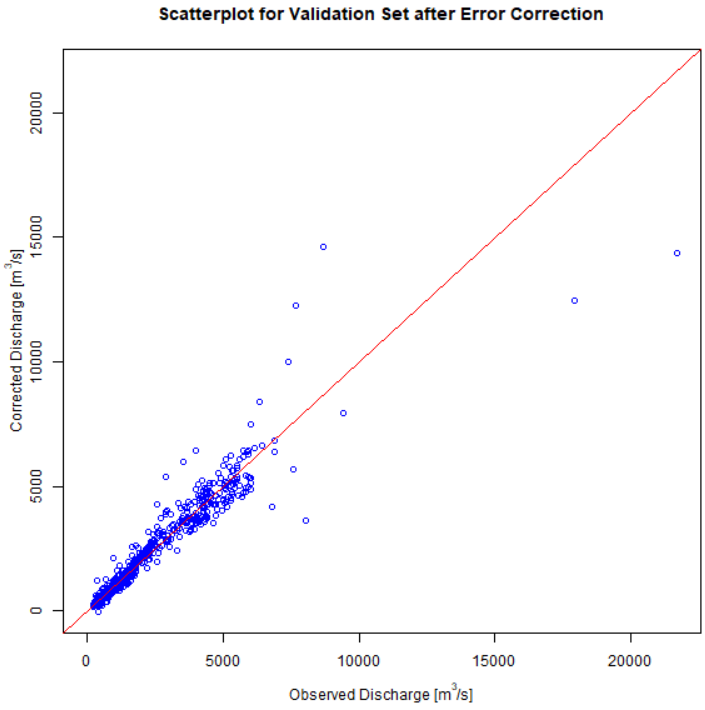

Figure 13. Scatterplot for validation set after error correction

Table 6 below summarizes the model performance before and after employing complementary error model.

Table 6. Summary of model performance

GOF measures	Calibration Set		Validation Set	
	TUWmodel only	TUWmodel and Error model	TUWmodel only	TUWmodel and Error model
me	-9.64	1.19	-134.14	1.44
mae	296.87	154.87	334.36	175.9
mse	275056	152559.4	372484.9	249920.7
rmse	524.46	390.59	610.32	499.92
nrmse %	29.5	22	34	27.8
PBIAS %	-0.7	0.1	-9.1	0.1
RSR	0.3	0.22	0.34	0.28
rSD	0.94	0.99	0.93	0.97
NSE	0.91	0.95	0.88	0.92
mNSE	0.77	0.88	0.74	0.86
rNSE	0.8	0.98	0.88	0.98
d	0.98	0.99	0.97	0.98
md	0.88	0.94	0.87	0.93
rd	0.94	1	0.97	1
cp	0.44	0.69	0.53	0.69
r	0.96	0.98	0.94	0.96
R2	0.91	0.95	0.89	0.92
bR2	0.85	0.93	0.79	0.89
KGE	0.92	0.97	0.87	0.95
VE	0.78	0.89	0.77	0.88

The results before error correction show that TUWmodel is good in conserving runoff volume with percent bias (PBIAS) -0.7% for the calibration period. The Nash-Sutcliffe efficiency (NSE)

for calibration set was 0.91. The PBIAS for the validation period was -9.1%. The NSE for validation set was 0.88. After employing the error model, the calibration efficiencies calculated using the PBIAS and NSE metrics improved to 0.1% and 0.95 respectively. Corresponding values for the validation period also improved to 0.1% and 0.92 respectively. The relative index of agreement (rd) also improved from 0.94 to 1 for the calibration set and 0.97 to 1 for the validation set. This indicates that the time to the peak flow is exactly matching after employing error correction. The performance indices for both calibration and validation sets clearly show the significant improvement of the performance of TUWmodel after employing complementary error model.

5. DISCUSSION AND CONCLUSIONS

In this study, we presented an application of a complementary data-based error correction model to enhance the performance of a conceptual rainfall-runoff model for flood forecasting. A conceptual rainfall-runoff model TUWmodel was developed for Karnali basin in Western Nepal. Four years of daily data from 01/01/2008 to 31/12/2011 were used to calibrate the model. The calibration data set consisted of major flood events in 2008 and 2009. The parameters of TUWmodel were calibrated automatically using differential evolution optimization algorithm. The model was then validated using three years of daily data from 01/01/2012 to 31/12/2014. The validation data set consisted of historically high flood event of August 2014. Although, only seven years of data were used to calibrate and validate the model, the data set consisted of major flood events crossing the danger level. In recent years, such flood events haven't been observed.

The errors of the calibration set from the conceptual rainfall-runoff model were then analyzed to identify the bias, persistency and heteroscedastic behavior. We outlined a procedure for extracting useful information from the bias, persistency and heteroscedasticity exhibited by the error series. We also presented an automatic procedure to identify the model structure and the parameters of the complementary error correction model using a freely available R package 'forecast'.

A data-driven ARIMA model was developed from the error series of the calibration set and employed to correct the predictions made by the conceptual model on both calibration and validation set. Application of the error correction model to both calibration and validation data set revealed that this procedure could effectively improve forecast accuracy of the conceptual rainfall-runoff model. This shows that the accuracy of a flood forecasting system can be significantly improved by setting up a data-driven error correction model to complement a conceptual rainfall-runoff model operating in the simulation mode.

In operational flood forecasting, the precipitation and temperature forecast will be obtained from NWP model in gridded form. Basin average precipitation and temperature forecast need to be computed from gridded data. Both the rainfall-runoff model and the error correction model

should be recalibrated for NWP inputs. Parameters of the conceptual rainfall-runoff model and error correction model will be calibrated automatically and kept unchanged during operational application. The automatic optimization module DEoptim provides an efficient method to update parameters of conceptual rainfall-runoff model whereas the auto.arima function provides an efficient method for automatically estimating parameters of error correction model. The whole system can be deployed in automatic operational forecasting mode by updating input time series in real-time.

The major limitations of this study are that daily data are used for calibration and validation of the model and the performance of the model isn't tested using NWP inputs. Sub-daily or hourly data aren't available for the study. Application of sub-daily or hourly data is required to capture the flood peak precisely. Although the data used in the model have covered historical peak events, longer length of data will help for better estimate of model parameters. We haven't tested the performance of the model in operational flood forecasting using NWP inputs and kept that research for future.

ACKNOWLEDGEMENT

This study has been financially supported by Practical Action Consulting, Nepal. We gratefully acknowledge the hydrological and meteorological data provided by the Department of Hydrology and Meteorology, Nepal

REFERENCES

Abebe, A. J. and Price, R. K., 2003. Managing uncertainty in hydrological models using complementary models, *Hydrolog. Sci. J.*, 48, 679–692.

Bergstörm, S., 1976. Development and application of a conceptual runoff model for Scandinavian catchments. Department of Water Resources Engineering, Lund Institute of Technology/University of Lund, Bulletin Series A, no. 52.

Blaney, H. F., Criddle, W. D., 1950. Determining water requirements in irrigated areas from climatological and irrigation data. In *ISDA Soil Conserv. Serv.*, SCS-TP-96.

Box, G. E. P. and Cox, D. R., 1964. An analysis of transformations. *JRSS B* 26 211–246.

Brockwell, P. J. and Davis, R. A., 1996. Introduction to Time Series and Forecasting, *Springer*, New York. Sections 3.3 and 8.3.

Burnash, R. J. C., 1995. The NWS River Forecast System – Catchment Modeling. In Computer models in watershed hydrology, V. P. Singh (Ed.), Water Resources Publications.

Byrd, R. H., Lu P., Nocedal, J. and Zhu, C., 1995. A limited memory algorithm for bound constrained optimization. *SIAM J. Scientific Computing*, 16(5), 1190–1208.

DHI, 2011. MIKE 11: A Modeling System for Rivers and Channels, Reference Manual, DHI Water & Environment, Denmark.

Gragne, A. S., Sharma, A., Mehrotra, R., and Alfredsen, K., 2015. Improving real-time inflow forecasting into hydropower reservoirs through a complementary modelling framework, *Hydrol. Earth Syst. Sci.*, 19, 3695-3714, https://doi.org/10.5194/hess-19-3695-2015.

Guerrero, V.M., 1993. Time-series analysis supported by power transformations. *Journal of Forecasting*, 12, 37–48.

Hyndman, R.J. and Khandakar, Y., 2008. Automatic time series forecasting: The forecast package for R, *Journal of Statistical Software*, 26(3).

Kachroo, R. K., 1992. River flow forecasting: Part 1 – A discussion of the principles, *J. Hydrol.*, 133, 1–15.

Li, J. and Liu, C., 2017. Improvement of LCM model and determination of model parameters at watershed scale for flood events in Hongde basin of Chaina. *Water Science and Engineering*, 10(1), 36-42, https://doi.org/10.1016/j.wse.2017.03.006.

Liu, Y., Weerts, A. H., Clark, M., Hendricks Franssen, H.-J., Kumar, S., Moradkhani, H., Seo, D.-J., Schwanenberg, D., Smith, P., van Dijk, A. I. J. M., van Velzen, N., He, M., Lee, H., Noh, S. J., Rakovec, O., and Restrepo, P., 2012. Advancing data assimilation in operational hydrologic forecasting: progresses, challenges, and emerging opportunities, *Hydrol. Earth Syst. Sci.*, 16, 3863-3887, https://doi.org/10.5194/hess-16-3863-2012.

Maidment, D. R., 1993. Handbook of Hydrology, 1st Edn.NewYork: McGraw Hill Publication.

Merz R and Blöschl G, 2004. Regionalisation of catchment model parameters. *Journal of Hydrology* 287: 95–123. DOI: 10.1016/j.jhydrol.2003.09.028.

Morawietz, M., Xu, C. Y., Gottschalk, L., Tallaksen, L. M., 2011. Systematic evaluation of autoregressive error models as post-processors for a probabilistic streamflow forecast system, *J. Hydrol.*, 407(1–4), 58–72, doi:10.1016/j.jhydrol.2011.07.007.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., Veith, T. L., 2007. Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. *Transactions of the ASABE*. 50 (3): 885–900

Mullen, K.M, Ardia, D., Gil, D. L., Windover, D., Cline, J., 2011. An R Package for Global Optimization by Differential Evolution. *Journal of Statistical Software*, 40(6), 1-26, 2011. URL: http://www.jstatsoft.org/v40/i06/.

Nau, R, 2014. Lecture notes on forecasting. Fuqua School of Business, Duke University. URL: http://people.duke.edu/~rnau/Slides_on_ARIMA_models--Robert_Nau.pdf.

Parajka, J., Merz, R., Bloeschl, G., 2007. Uncertainty and multiple objective calibration in regional water balance modelling: case study in 320 Austrian catchments, *Hydrological Processes*, 21, 435-446, doi:10.1002/hyp.6253.

Pingping Luo, Dengrui Mu, Han Xue, Thanh Ngo-Duc, Kha Dang-Dinh, Kaoru Takara, Daniel Nover, Geoffrey Schladow, 2018. Flood inundation assessment for the Hanoi Central Area, Vietnam under historical and extreme rainfall conditions, *Scientific Reports (Nature)*, 2018, 8:12623, DOI:10.1038/s41598-018-30024-5

Serban, P. and Askew, A. J., 1991. Hydrological forecasting and updating procedures, in: *Hydrology for the Water Management of Large River Basins*, Proceedings of the Vienna symposium, Vienna, 357–369.

Sugawara, M., 1995. Tank model. In Computer models of watershed hydrology, V. P. Singh (Ed.), *Water Resources Publications*, 165-214.

Todini, E., 1996. The ARNO rainfall-runoff model. Journal of Hydrology, 175, 339-382.

Todini, E., 2007. Hydrological catchment modelling: past, present and future, *Hydrol. Earth Syst. Sci.*, 11, 468–482, doi:10.5194/hess-11 - 468-2007.

WMO, 2011. Manual on Flood Forecasting and Warning, WMO-No. 1072, World Meteorological Organization, Geneva, Switzerland.

Zhao, R.-J., Zuang, Y.-L., Fang, L.-R., Liu, X.-R. and Zhang, Q.-S., 1980. The Xinanjiang model. Proceedings of the Oxford Symposium, IAHS Publication 129, 351-356