

Development of a Deep Learning-based Error-Updating Model for Improved Streamflow Forecasting Accuracy of a Hydrological Model

Amina Khatun¹, Bhabagrahi Sahoo², and Chandranath Chatterjee¹

¹Agricultural and Food Engineering Department, School of Water Resources, Indian Institute of Technology Kharagpur

²Affiliation not available

February 15, 2024

Development of a Deep Learning-based Error-Updating Model for Improved Streamflow Forecasting Accuracy of a Hydrological Model



Amina Khatun (1), Bhabagrahi Sahoo (2), Chandranath Chatterjee (1)

(1) Agricultural and Food Engineering Department, (2) School of Water Resources, Indian Institute of Technology Kharagpur



PRESENTED AT:



INTRODUCTION

- Reliable forecasting of river-flows with sufficient lead-time aids for developing early warning systems against flood havoc and in regulating reservoir releases for proper water resources management.
- Generally, the meteorological forecasts from Numerical Weather Prediction (NWP) models are widely used as forcings in the rainfall runoff models for operational flood forecasting.
- However, with the inherent discrepancies in the hydrometeorological forecasts, the NWP models cannot accurately represent the physical atmospheric processes at a finer scale.
- Thus, bias-correction of these ensemble rainfall forecasts can help to improve the quality of hydrological model-simulated streamflow forecasts by reducing the biases in the mean rainfall and its variance.
- To date, a plethora of methods dealing with the data-driven, conceptual and physically-based approaches have been developed to model the complex, nonlinear rainfall-runoff process.
- Among the conceptual models, the MIKE11-NAM (Nedbør Afstrømnings Model) is found to be popular to simulate the runoff generation dynamics from small to large catchments with a varied climate as it could be easily integrated with the MIKE11-HD (Hydrodynamic) module to simulate the channel routing process.
- However, despite using improved input forcings and advanced calibration techniques, the outputs of hydrological models still suffer from poor prediction accuracy.
- This may be due to the uncertainties associated with the model structure, its inputs and the parameters.
- This problem is taken care of by adopting a suitable error-updating approach which involves correction of the errors in the hydrological model prediction.
- Among the available deep learning techniques, the Long Short-Term Memory (LSTM), a special type of Recurrent Neural Network (RNN), is the state-of-the-art network structure that can learn or preserve the long term dependencies among the input–output variables.
- Although the application of the LSTMs is rapidly developing in the field of hydrology, being a deep-learning model variant, this has never been tested as an error-updating model till date.

HYDROLOGICAL MODEL

- In large river basins, the channel flows are dominant over the overland flows.
- Therefore, the MIKE11-NAM conceptual model is used herein to simulate the pluvial surface runoff at the sub-catchment scale, which is then routed by using the physically-based MIKE11-HD model with the full Saint Venant equations of continuity and momentum conservation. This integration is termed as MIKE11-NAM-HD.

Calibration and validation

- The MIKE11-NAM-HD is calibrated and validated using the hydrometeorological, viz. mean areal rainfall, mean areal potential evapotranspiration, and streamflow datasets of the monsoon season (June–September) of the years 2000–2007 and 2008–2014, respectively.
-

ERROR-UPDATING MODEL

- This study proposes a nested approach characterized with a cascade of six LSTM models in the first phase of LSTM modelling setup and one LSTM model in the second phase of setup (Fig. 3).
- As a pre-processing of input datasets in the LSTM network, the variance in the error time-series is smoothed using 'hanning', 'hamming', 'bartlett' and 'blackman' smoothing windows, out of which the 'blackman' window is found to be the best.
- The predicted errors generated by each of the sLSTM model are back-transformed to the original domain using the same scaler.
- Each sLSTM model used herein follows a sequence-to-singleoutput procedure in the form of sliding windows considering a total of past five (time-lagged) inputs (errors) to forecast the single δ -day ahead output.

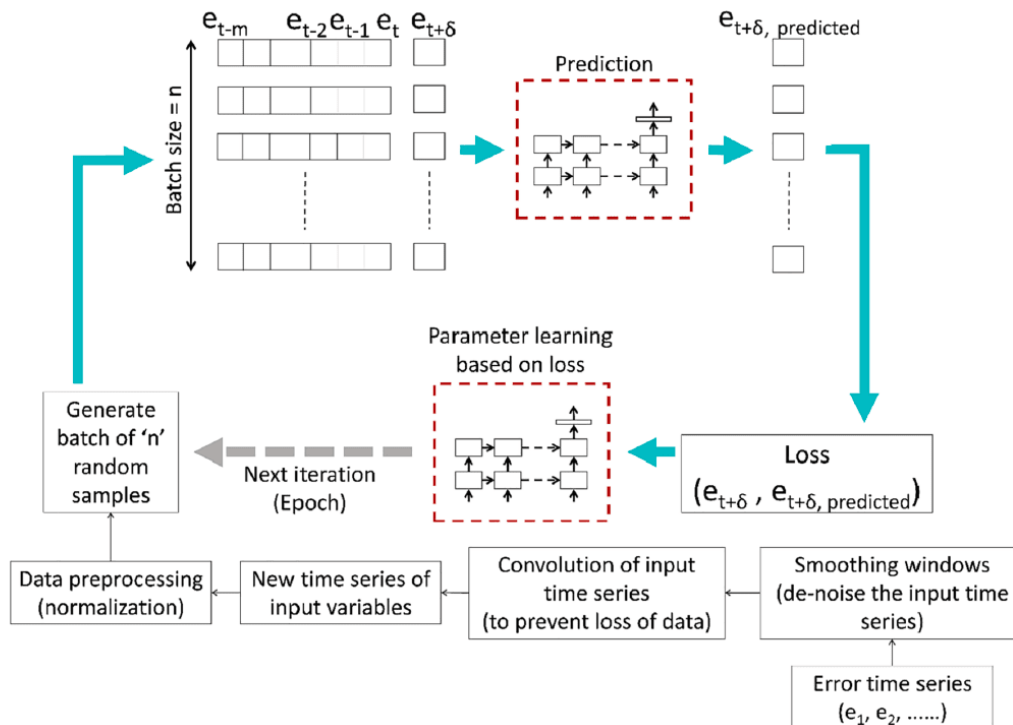


Fig. 3. sLSTM error-updating model.

RESULTS

- During calibration, the MIKE11-NAM-HD model showed satisfactory performance with $NSE = 0.83$, $r = 0.91$, $MAE = 608.65 \text{ m}^3/\text{s}$, $RSR = 0.42$, and $|Evol| = 2.84\%$ (Fig. 2).
- Similarly, during validation, this model performed with $NSE = 0.91$, $r = 0.95$, $MAE = 500.66 \text{ m}^3/\text{s}$, $RSR = 0.30$ and $|Evol| = 1.55\%$.
- Overall, the MIKE11-NAM-HD model performed very well during both the calibration and validation phases.

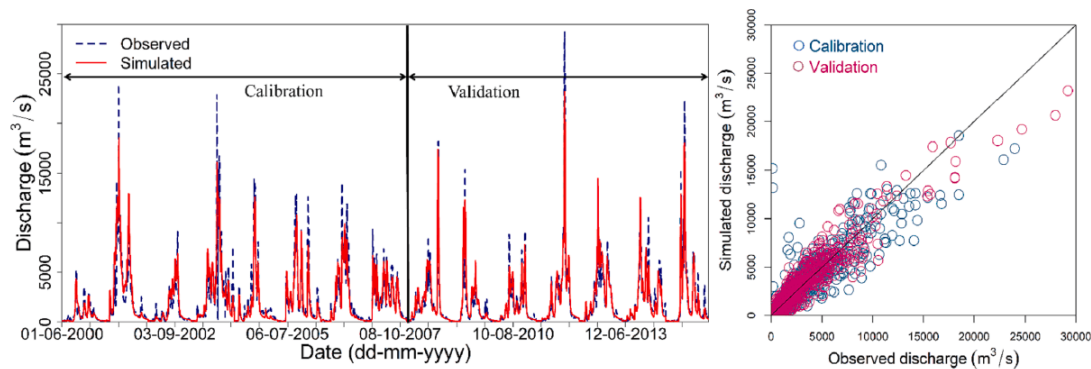


Fig. 2. Calibration and validation of MIKE11-NAM-HD model.

- The bcMIKE-sLSTM (bc = using bias-corrected rainfall) model performs remarkably well at all the lead-times with $NSE = 0.81\text{--}0.92$ and $r = 0.90\text{--}0.96$.
- The bcMIKE-sLSTM model shows 'very good' ($RSR = 0.29\text{--}0.44$) streamflow forecasting skills with an $RSR < 0.5$.
- Also, the bcMIKE-sLSTM is able to capture the peak flood with reasonable accuracy up to 3 and 4 days lead-times.

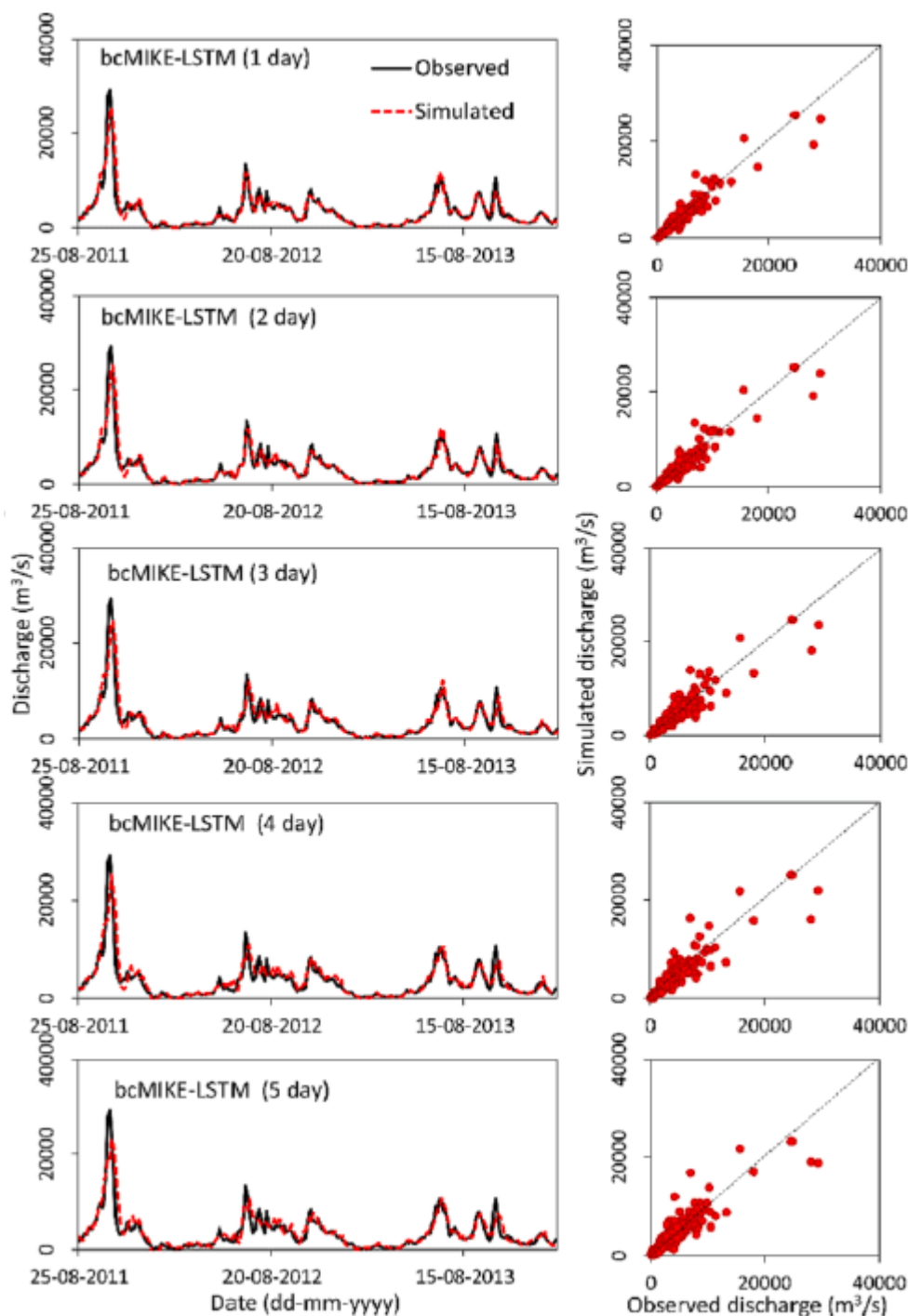


Fig. 4. Performance of the bcMIKE-sLSTM model.

STUDY AREA AND DATA USED

Study area

- The Hirakud reservoir catchment of the upstream Mahanadi River basin comprises of 83,400 km² area that is located in between 19°30'N–23°35'N latitudes and 80°30'E–84°30'E longitudes in eastern India (Fig. 1).
- It is characterized by flat to moderate slopes having a minimum temperature of about 4–12°C (December–January) and a maximum temperature of about 42–45°C occurring in May.
- This tropical rainfed catchment receives an average rainfall of about 1400 mm annually, 75% of which occurs during the southwest monsoon season during June to September with long series of zero rainfall events during the non-monsoon season.
- The Hirakud dam is one of the largest earthen dams in the world, in operation since 1957, which is constructed with a live storage capacity of 5818×10⁶ m³.
- The current live storage capacity of the reservoir is about 4823×10⁶ m³.
- The area downstream of the Hirakud dam is mostly flood-prone.

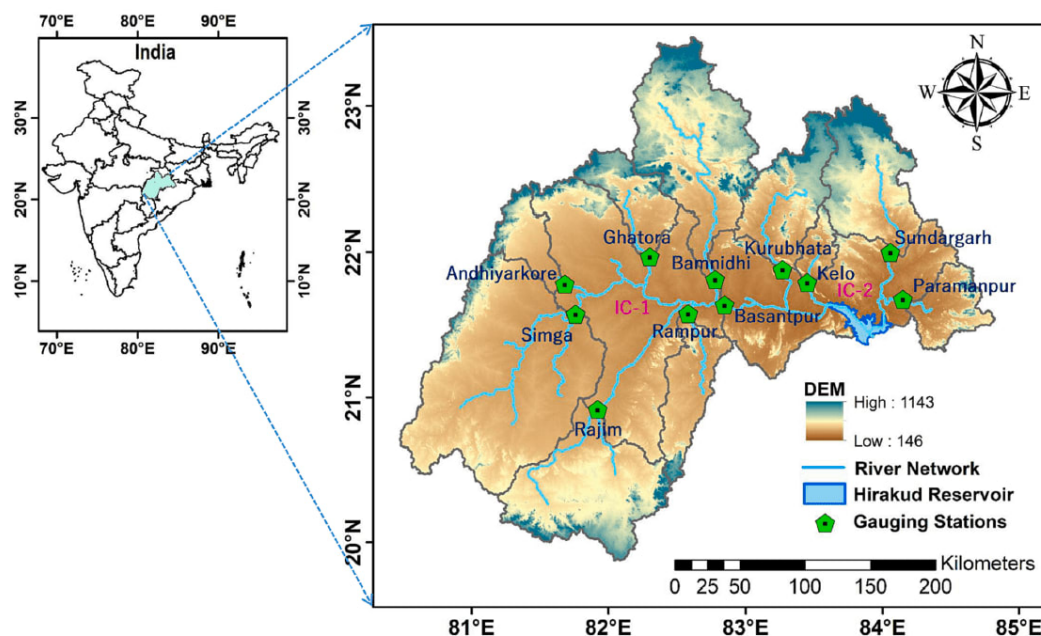
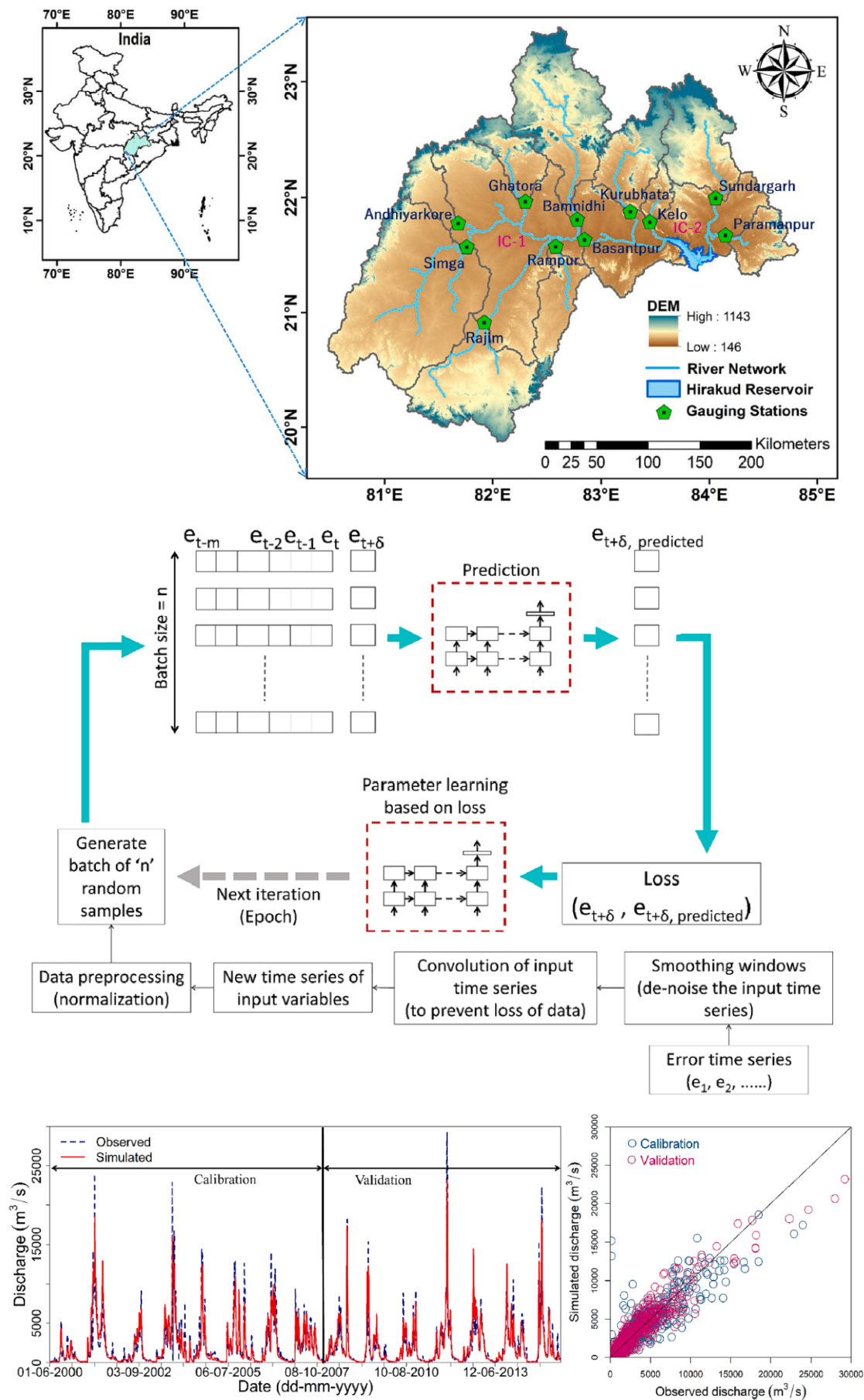
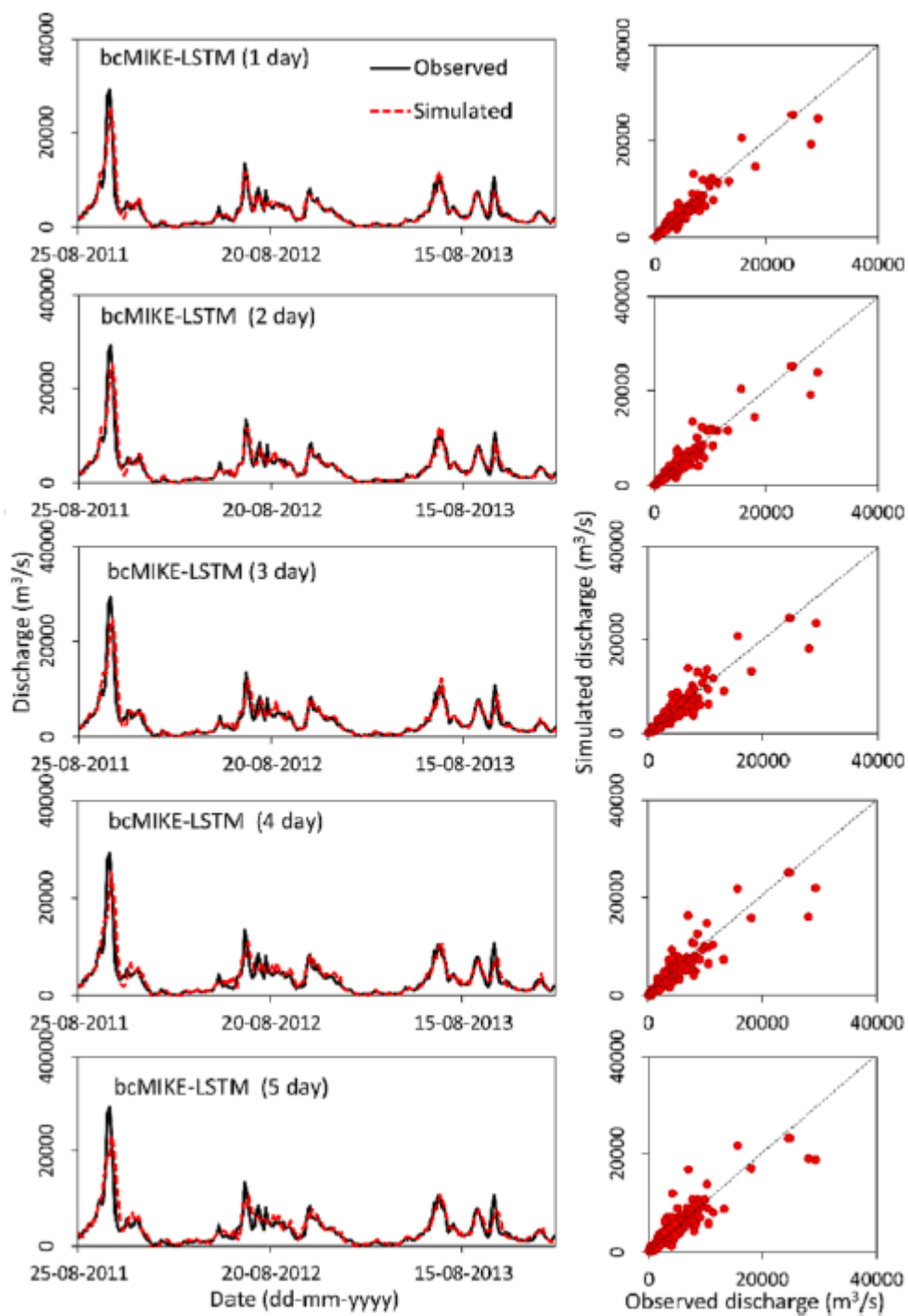


Fig. 1. Upper Mahanadi River basin

Data used

- Based on the period of data availability, the short-to-medium range forecasts of meteorological variables, such as rainfall and temperature from the IMD-MME forecast datasets over six years (2008–2013) is used in this study.
- The rainfall forecasts are bias-corrected using a newly developed hybrid Copula and enhanced Kohonen Self-Organizing Map based bias-correction technique.
- The temperature datasets are used to estimate the potential evapotranspiration employing the Hargreaves temperature method.





CONCLUSIONS

- The bcMIKE-LSTM framework outperforms all other model variants in forecasting the overall discharge time-series at 1–5 days lead-times ($NSE = 0.81–0.92$) as well as the high flood peaks.
-

AUTHOR INFO

Amina Khatun

Then Ph.D. Candidate in the Agricultural and Food Engineering Department, IIT Kharagpur, India.

Currecntly, Assistant Professor (Agricultural Engineering), College of Horticulture & FSR, Assam Agricultural University, Nalbari, Assam, India.

TRANSCRIPT

ABSTRACT

Recent advancements in the deep learning models, such as the Long Short-Term Memory (LSTM) networks are gaining popularity in hydrological applications. In this study, an LSTM-based error-updating model is developed to forecast the streamflow prediction errors of a hydrological model, namely, the MIKE11-NAM-HD (MIKE). The daily raw rainfall forecasts from the ensemble rainfall forecast products (IMD-MME) of the India Meteorological Department (IMD) up to a 3-days lead-time are bias-corrected using a hybrid copula-enhanced Kohonen Self-Organizing Map-based bias-correction technique. Both the raw and bias-corrected rainfall forecasts, along with the evapotranspiration forecasts are forced as the meteorological inputs to the MIKE model for the upper reaches of the Mahanadi River basin in eastern India. A smoothing-based LSTM (sLSTM) error-updating model is trained and tested using the errors in the MIKE-simulated streamflows. A nested approach is followed in developing the sLSTM model to obtain improved accuracy in the daily streamflow forecasts. The results indicate inter-comparable performance of the MIKE-LSTM models with the raw (NSE=0.92) and bias-corrected (NSE=0.92) rainfall forecasts at 1-day lead time. However, as the lead-time increases from 2-3 days, the performance of the MIKE-LSTM model with the raw rainfall forecasts deteriorates to produce a Nash-Sutcliffe Efficiency (NSE) of 0.87 (2-day) and 0.67 (3-day). The MIKE-LSTM model with the bias-corrected rainfall forecasts outperforms with an NSE of 0.90 and 0.87 at 2-day and 3-day lead-times, respectively. Along with the overall time-series, the MIKE-LSTM model forced with the bias-corrected rainfall forecasts is also able to capture the annual maximum peaks in the testing period with reasonable accuracy.

