Coupling SWAT and LSTM for Improving Daily Streamflow Simulation in Mahanadi Basin

Himanshu Singh[1], Iqura Malik [1] under supervision of Prof. Vimal Mishra [1,2]
[1]Civil Engineering, IIT Gandhinagar, India, [2]Department of Earth Sciences, IIT Gandhinagar, India

WATER & CLIMATE LABORATORY
IIT GANDHINAGAR

1. Introduction

- Why it matters: Daily streamflow forecasts are critical for flood early warning, drought mitigation, reservoir operations and water-resources planning.
- Limitations of SWAT: The process-based SWAT model often under- or over-estimates peak and base flows due to structural uncertainties and parameter uncertainties .
- Rise of ML: Data-driven Long Short-Term Memory (LSTM) networks excel at capturing nonlinear, long-term dependencies in hydrological time series—but lack physical interpretability.
- Our idea: By coupling SWAT with an LSTM that learns SWAT's residual errors, we aim to combine physical realism and ML flexibility.

2. Objectives

- Quantify the standalone performance of SWAT and LSTM in daily discharge simulation.
- Develop a SWAT–LSTM hybrid that corrects SWAT's bias via an LSTM trained on residuals.
- Compare the hybrid against other approaches.
- Assess improvements in high-flow peaks and low-flow baseflow across calibration (1973–2000) and validation (2001–2012).

3. Data and Methods

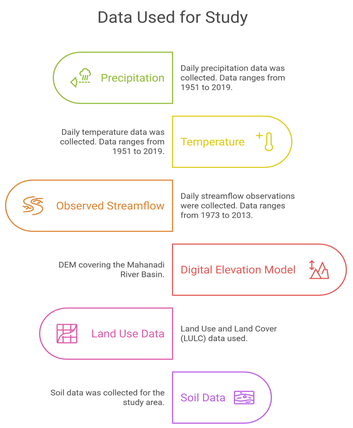


Figure 1. Data Used to Perform the Study.

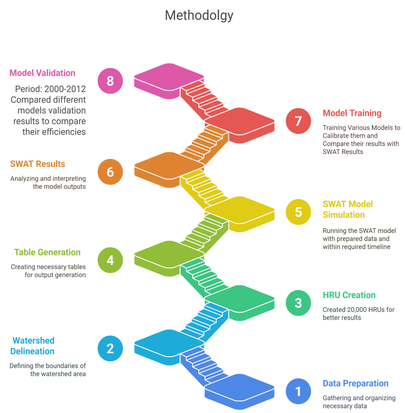
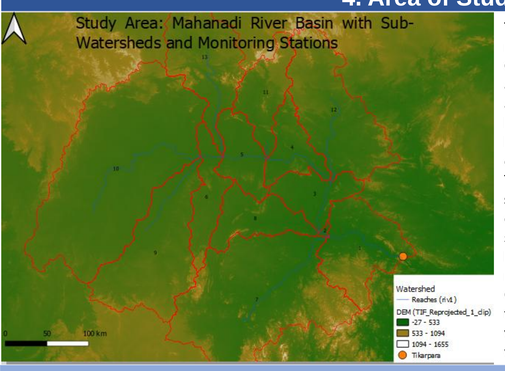


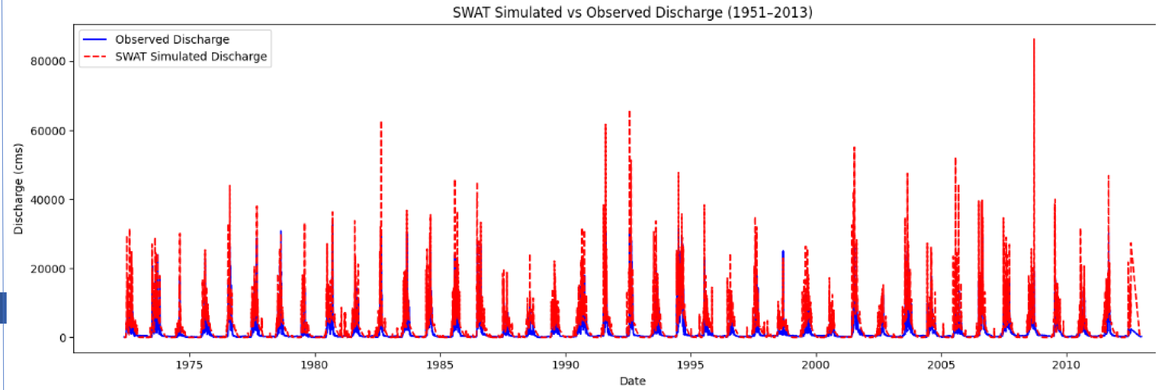
Figure 2. Methodology Used to Perform the Study.

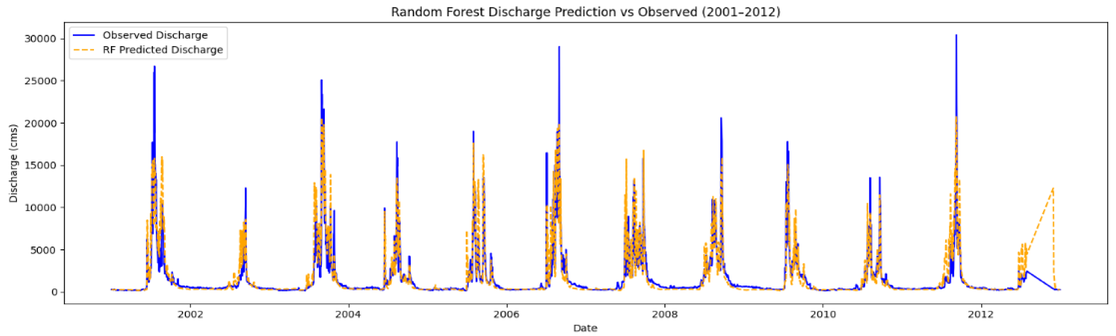
4. Area of Study

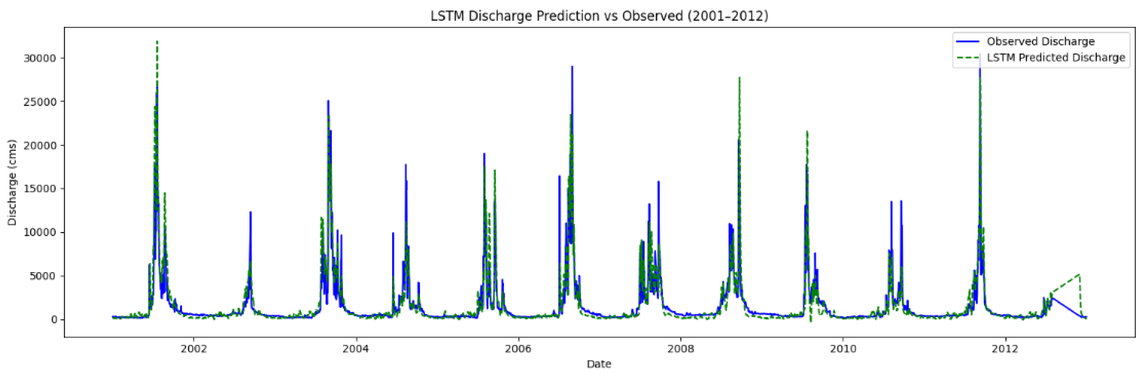


The Mahanadi River Basin is one of the major river basins in India, covering parts of Chhattisgarh and Odisha, and spans an area of approximately 141,600 square kilometers. It plays a crucial role in agriculture, water supply, and hydropower generation in the region. The basin has been divided into 13 sub-watersheds for detailed hydrological analysis.
This study primarily focuses on the **Tikarpara station**, a key location within the basin that lies downstream of most tributaries. Tikarpara is strategically important due to its proximity to the Hirakud Dam and its relevance in flood monitoring and water resource management. Observed and simulated hydrological data from this station have been used for model calibration, validation, and performance assessment in this work.

4. Results







Model	Calibration period (1973-2000)				Validation period (2001-2012)			
	NSE	R2	RMSE (m3/s)	PBIAS (%)	NSE	R2	RMSE (m3/s)	PBIAS (%)
SWAT	-0.5447	-0.5447	3885.93	94.67		N/A		
SWAT-LSTM	0.7815	0.7815	1530.8	-1.2	0.7811	0.7811	1322.83	-5.15
Random Forest	0.976	0.976	502.57	-0.13	0.7636	0.7636	1374.55	7.44

5. Conclusions

The SWAT-LSTM hybrid model performs much better than the SWAT model alone. It gives high accuracy with NSE values of 0.78 in calibration and 0.75 in validation, while SWAT alone gives a poor score of –0.54. It also reduces the error in flow prediction, bringing the bias down to –1.2% and –5.1%, compared to SWAT's large overestimation of +94.7%. It predicts both flood peaks and low flows well, with RMSE around 1530 m³/s in calibration and 1307 m³/s in validation.

By combining the physical knowledge of SWAT with the learning ability of LSTM, this model gives more accurate and balanced streamflow results useful for managing floods and droughts.

In the future, this hybrid approach can be extended to other river basins for wider application. Further improvements can include incorporating real-time data and testing with other machine learning models to enhance prediction accuracy and adaptability under changing climate conditions.

6. References

- Mei, Z., Peng, T., Chen, L. et al. Coupling SWAT and LSTM for Improving Daily Streamflow Simulation in a Humid and Semi-humid River Basin. Water Resource Manage 39, 397–418 (2025). https://doi.org/10.1007/s11269-024-03975-w
- https://github.com/coconutgrapefruit/SWAT-LSTM_ver4

7. Scope of Improvement

1. The current SWAT-LSTM model is trained on a single basin. It still needs to be validated on other basins .
2. Exploring other hybrid configurations (e.g., CNN-LSTM, Transformer-based models, Physics-Informed Neural Networks) may improve interpretability and performance.
3. Integrating the hybrid model with real-time data from satellites or sensors can enhance practical water resource forecasting.

Email: 23110134@iitgn.ac.in