Project Course CE:299

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

Under Supervision of

Prof. Vimal Mishra

And

Iqura Malik

by

Himanshu Singh

WATER & CLIMATE LAB
IIT GANDHINAGAR

Introduction

Importance of streamflow simulation in hydrology

SWAT: Robust but limited by assumptions and biases

LSTM: Learns patterns from historical data

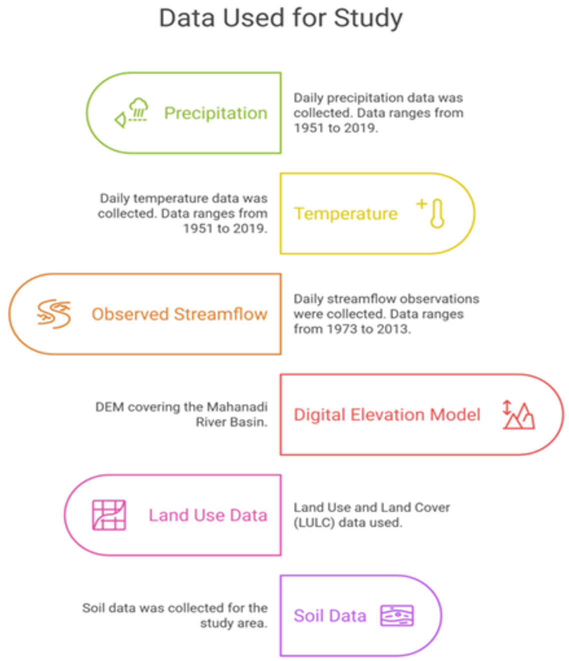
Hybrid SWAT-LSTM model aims to improve prediction accuracy

Objective

To improve streamflow simulation accuracy by coupling the SWAT model with an LSTM-based post-processing approach that reduces errors and bias, particularly during high and low flow periods.

Study Area and Data

This study primarily focuses on the Tikarpara station, a key location within the Mahanadi 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.



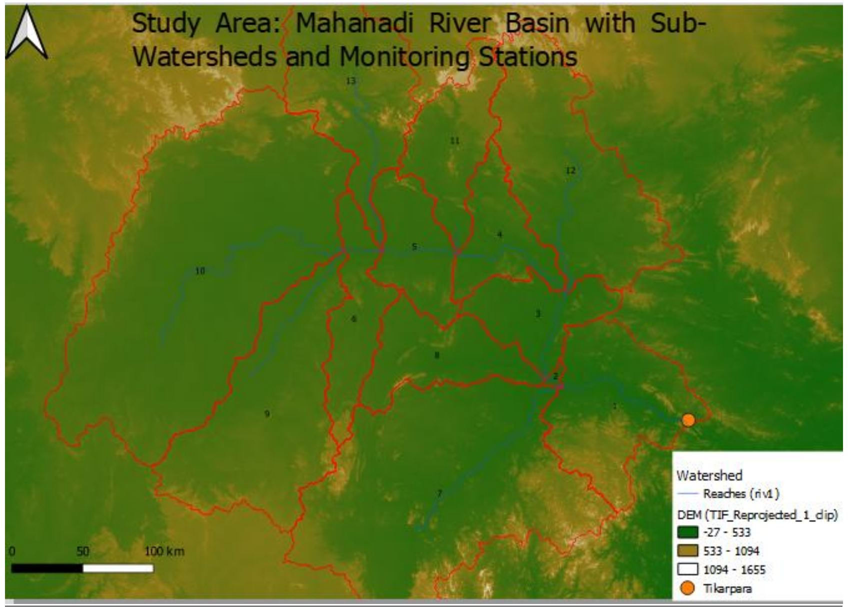
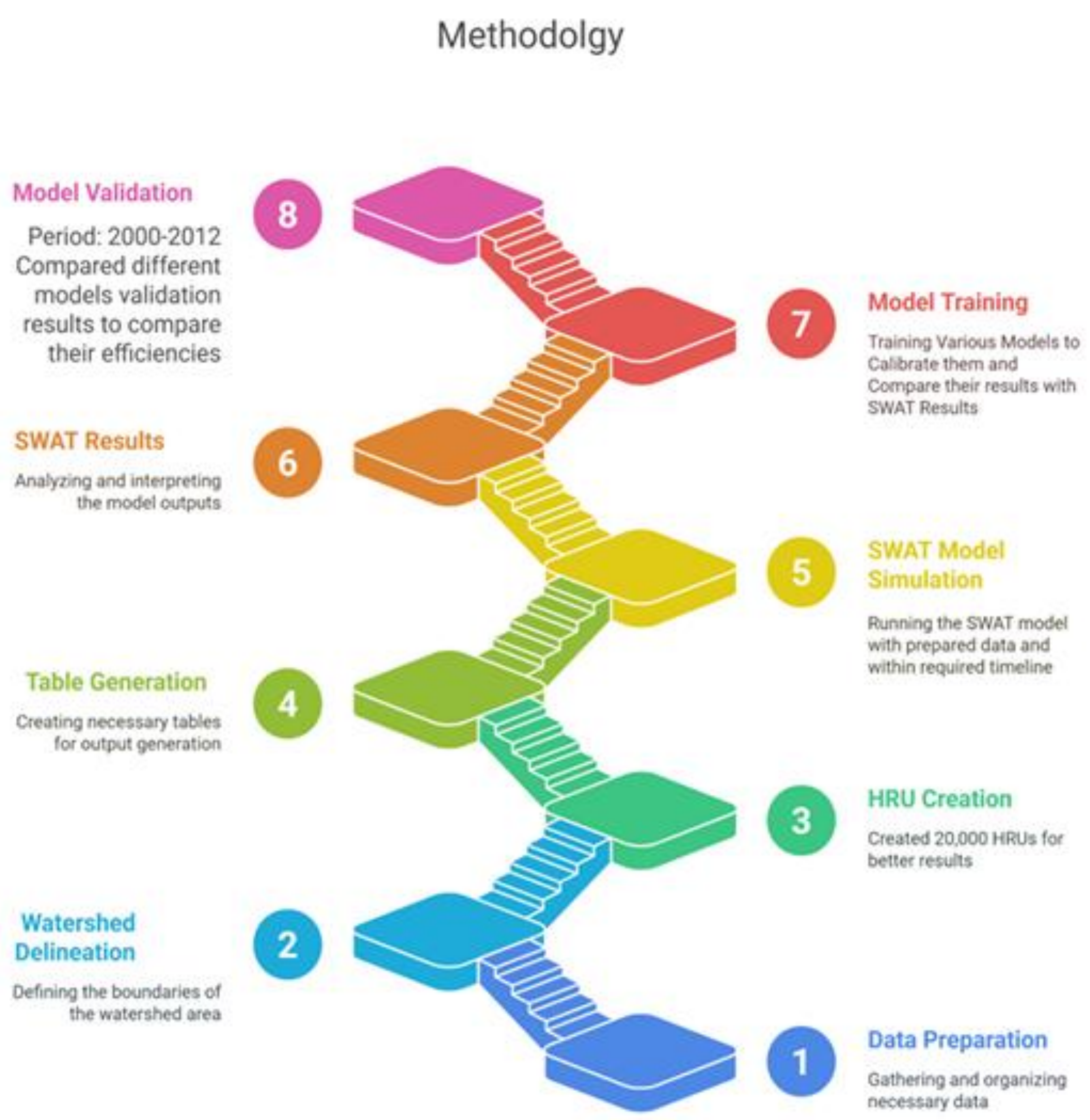


Fig.1 Area of Study

Methodolgy



Methodolgy
Model Validation
8
Period: 2000-2012
Compared different
models validation
results to compare
their efficiencies
7
Model Training
Training Various Models to
Calibrate them and
Compare their results with
SWAT Results
SWAT Results
6
Analyzing and interpreting
the model outputs
5
SWAT Model
Simulation
Running the SWAT model
with prepared data and
within required timeline
Table Generation
4
Creating necessary tables
for output generation
3
HRU Creation
Created 20,000 HRUs for
better results
Watershed
Delineation
2
Defining the boundaries of
the watershed area
1
Data Preparation
Gathering and organizing
necessary data

Results

```
SWAT Model Performance (1951–2013):
NSE      : -0.5447
R²       : -0.5447
RMSE     : 3885.93 m³/s
PBIAS    : 94.67 %
```

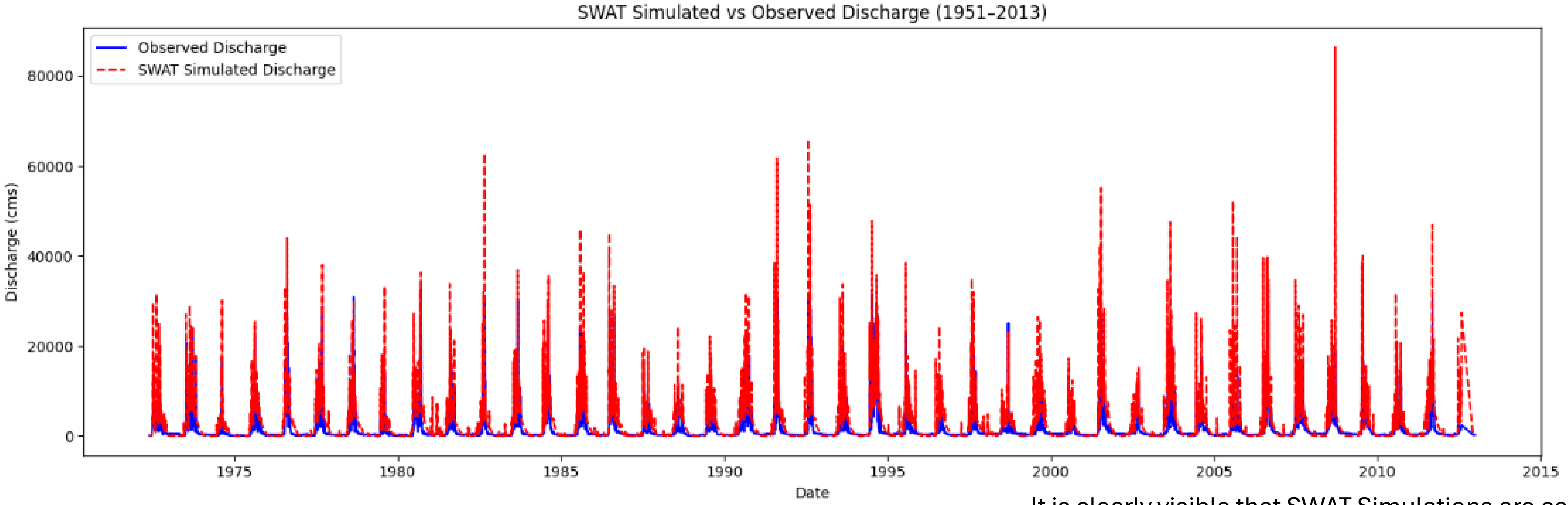


Fig.2 SWAT simulations vs Observed Values

It is clearly visible that SWAT Simulations are capturing variations well but there is huge difference in values.

Continued.....

Results

Calibration(1973-2000) and Validation(2001-2012)

LSTM RMSE: 1322.83
LSTM NSE: 0.7811
LSTM R² : 0.7811

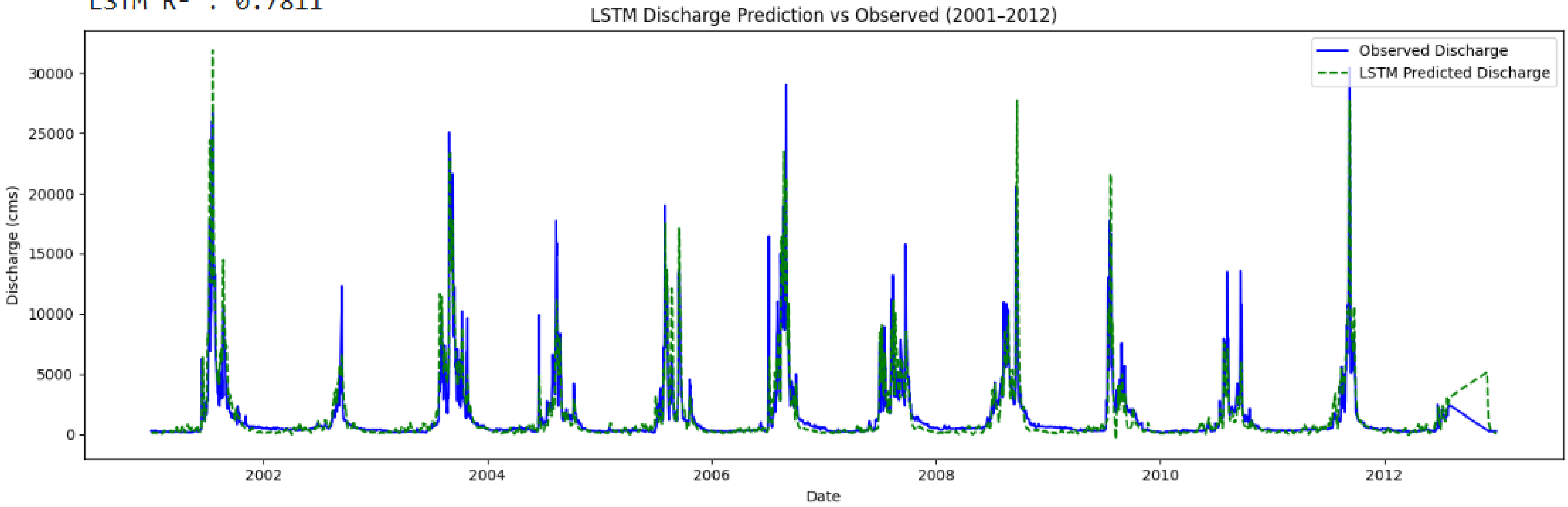


Fig.3 LSTM Predictions vs Observed (2001-2012)

We can see this model is capturing trends and values both a lot better than SWAT simulations.

Continued.....

Results

Calibration(1973-2000) and Validation(2001-2012)

```
Random Forest RMSE: 1374.55
Random Forest NSE: 0.7636
Random Forest R² : 0.7636
```

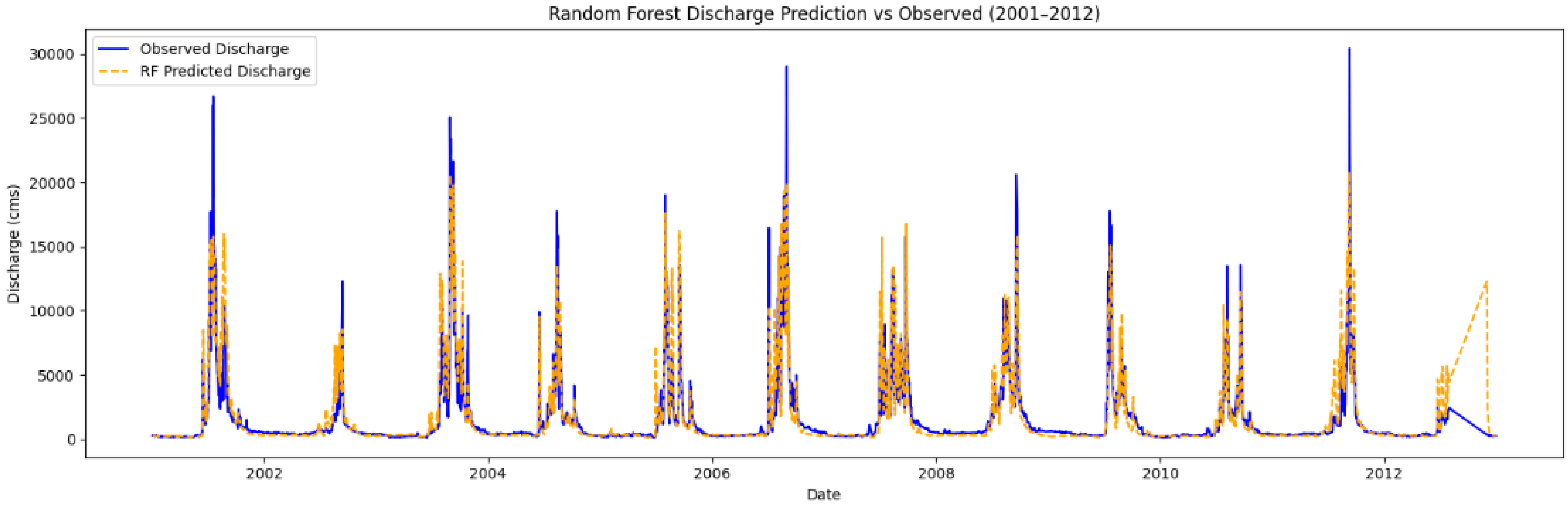


Fig.4 Random Forest Predictions vs Observed (2001-2012)

We can see the RF model performs poorer than the SWAT-LSTM model. This happens because RF does not capture the temporal dependencies and sequence patterns effectively, especially during high peak flow events. RF tends to underpredict peak discharges and overpredict baseflows, which results in less accurate discharge simulation compared to the LSTM model

Conclusion

- Standalone SWAT needs calibration which depends on tuning different Parameters which takes time and is a difficult process.
  NSE = –0.54 and PBIAS = +95% over 1951–2013 indicates severe overestimation of both flood peaks and low flows.

- We can improve calibration of SWAT by coupling it with Machine Learning.

- Data-driven ML models deliver major gains
  LSTM: NSE = 0.78, R^2 = 0.78, RMSE =1323 m³/s, PBIAS = 5%
  Random Forest: NSE =0.76, R^2 =0.76, RMSE =1375 m³/s, PBIAS = 7%

- LSTM edges out RF in capturing both high-flow peaks and baseflows during 2001–2012

Key takeaway:
Pure process-based models struggle with complex nonlinearity, whereas ML (especially LSTM) can effectively learn residual errors and reduce bias

Implication:
Coupling SWAT with an LSTM error-corrector harnesses physical realism and data flexibility for more reliable streamflow forecasts.

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

Thank You