SATELLITE-BASED FRAMEWORK FOR RIVER DISCHARGE ESTIMATION: A HYBRID APPROACH INTEGRATING SENTINEL-1, SENTINEL-2 AND ALTIMETRY DATA

ORTA DOĞU TEKNİK ÜNİVERSİTESİ
MIDDLE EAST TECHNICAL UNIVERSITY

Ceren Y. Tural [1], Koray K. Yilmaz, [1] and Angelica Tarpanelli[2]

[1]Middle East Technical University (METU), Geological Eng. Dept., Türkiye (yazigulu@metu.edu.tr; yilmazk@metu.edu.tr)

[2] Research Institute for Geo-Hydrological Protection, National Research Council, Italy (angelica.tarpanelli@cnr.it)

This presentation participates in OSPP

Outstanding Student & PhD candidate Presentation contest

MOTIVATION

- ✓ Climate change – increase frequency & severity of droughts & floods - necessitates innovative & reliable techniques for continuous monitoring of discharge to effectively manage risk.
- ✓ Objective: Develop a data-driven river discharge estimation algorithm that generalizes well across diverse basins with varying hydroclimatic conditions.
- ✓ We utilized remotely-sensed water area, water indices, band values from optical, SAR, and altimetry satellites, in-situ discharge observations collected during field campaigns, and both machine learning (e.g., Random Forest Regression (RFR) (Breiman, 2001)) and deep learning method (e.g., Long-Short Term Memory (LSTM) (Hochreiter, S., & Schmidhuber, J., 1997) to estimate river discharge.

DATASETS & PLATFORM

Initial Approach:

- **Google Earth Engine (GEE):** A powerful cloud-based spatial analysis platform providing publicly available satellite data in Javascript and Python language, without downloading images.
- **Remotely-Sensed (RS) Data Sources:**

- **Sentinel-1** Synthetic Aparture Radar (SAR): Water indices and backscatter bands extracted via GEE
- **Sentinel-2** Multispectral Images (Level-1C): Water indices and reflectance bands extracted via GEE

- **Ground Truth Discharge Measurements:**

- The US Geological Survey (USGS, 2024) - Mississippi River
- The Regional Agency for Environmental Protection - Po River
- State Hydraulic Works of Türkiye (DSİ - Devlet Su İşleri, 2024) - Kizilirmak River

Ongoing Work:

- **Altimetry Data:** Sentinel-3 and Sentinel-6 from Hydroweb-next platform.
- **GLO-30 DEM:** Elevation & Slope features extracted via GEE
- **Seasonality Indicators:** sine/cos transformations (day-of-year) & monthly cycle

SITE SELECTION

Study areas were selected based on their representation of diverse hydrological flow regimes. To support the development of a generalized and transferable river discharge estimation model, one area was chosen from each of three different basins: Memphis along the Mississippi River (USA), Pontelagoscuro along the Po River (Italy), and the Sogutluhan along Kizilirmak River (Türkiye) **(Figure 1)**.

Additionally, these regions are characterized by substantial alluvial sediment deposits, which allow for lateral water movement in response to discharge surges, making them ideal for studying surface water dynamics and river expansion behavior.

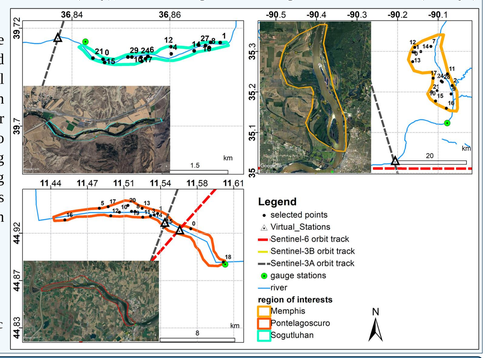


Figure 1: The geographic locations of study areas

INITIAL APPROACH: RF MODELING

Model Development: The RFR model was initially trained and validated using 5-fold cross-validation. To evaluate temporal generalizability, the same model was later trained using a time-based split. The overall workflow is illustrated in **Figure 2**, while the performance results and visualized results across all study regions are presented in **Table 1** and **Figure 3,** respectively.

In **Table 1,** RMSE and logRMSE are complementary: RMSE emphasizes flood dynamics, logRMSE captures low-flow reliability. R^2 presents the overall model fit.

RS Data Extraction from selected points & Merging in-situ measurements → RS Data Extraction & Merging in-situ measurements → Data Preprocessing - detection of missing value & outliers - elimination of highly correlated variables → Model Training 1. K-Fold CV with *GridSearch* 2. Time-split with *GridSearch* Hypertuning → Model Evaluation with Performance Metrics: *RMSE*, *logRMSE*, R^2 → Explainable AI – *Shapley Addictive exPlanations* (SHAP) for feature importance

Figure 2: The RFR workflowfor river discharge estimation with two different models

REFERENCES

Breiman, L. (2001) Random Forests. Machine Learning, 45, 5-32.http://dx.doi.org/10.1023/A:1010933404324

Long short-term memory. Neural Computation, 9(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

European Space Agency. (2024). Sentinel-1. Retrieved from https://sentinel.esa.int/web/sentinel/missions/sentinel-1

European Space Agency. (2024). Sentinel-2. Retrieved from https://sentinel.esa.int/web/sentinel/missions/sentinel-2

Google Earth Engine. (2024). About Earth Engine. Retrieved from https://developers.google.com/earth-engine

Hochreiter, S., & Schmidhuber, J"urgen. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780.

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems (NeurIPS), 30.

U.S. Geological Survey, 2024, Earthquake Lists, Maps, and Statistics, accessed Mar 30, 2024 at URL https://waterdata.usgs.gov/nwis/measurements

INITIAL APPROACH: RESULTS

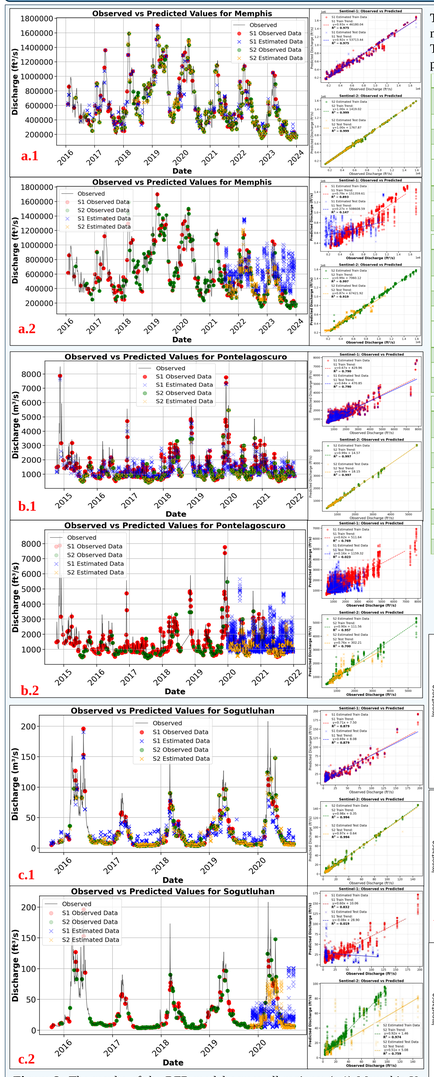


Figure 3: The results of the RFR model across all regions; (a.1) Memphis K-Fold CV, (a.2) Memphis Time-split, (b.1) Pontelagoscuro K-Fold CV, (b.2) Pontelagoscuro Time-split, (c.1) Sogutluhan K-Fold CV, (c.2) Sogutluhan Time-split

Table 1: The performance metrics of K-Fold CV and Time-split models across all proposed areas

Study Site	Memphis	
Performance Metrics	**Sentinel-1 K-Fold**	**Sentinel-1 Time-split**
RMSE	67532.64	311984.66
logRMSE	0.15	0.72
R^2	0.975	**0.147**
Performance Metrics	**Sentinel-2 K-Fold**	**Sentinel-2 Time-split**
RMSE	12401.09	249685.71
logRMSE	0.02	0.15
R^2	0.999	**0.919**

Study Site	Pontelagoscuro	
Performance Metrics	**Sentinel-1 K-Fold**	**Sentinel-1 Time-split**
RMSE	456.16	749.82
logRMSE	0.30	0.47
R^2	0.790	**0.023**
Performance Metrics	**Sentinel-2 K-Fold**	**Sentinel-2 Time-split**
RMSE	49.39	280.76
logRMSE	0.04	0.16
R^2	0.996	**0.700**

Study Site	Sogutluhan	
Performance Metrics	**Sentinel-1 K-Fold**	**Sentinel-1 Time-split**
RMSE	12.65	40.62
logRMSE	0.59	1.35
R^2	0.879	**0.019**
Performance Metrics	**Sentinel-2 K-Fold**	**Sentinel-2 Time-split**
RMSE	3.36	24.53
logRMSE	0.18	0.53
R^2	0.994	**0.759**

While the RF model performed well under K-Fold CV validation with high R^2, low RMSE during peak flows, and consistent logRMSE during low-flow periods, its performance declined when using a time-based split, as illustrated in **Figure 3** and supported by the metrics presented in **Table 1**.

This revealed **limited generalizability** to **unseen** temporal patterns. Since our goal is to develop a robust and transferable algorithm applicable across diverse hydroclimatic regimes, **we required a model that can adapt to sequential dynamics and data variability.** Therefore, we transitioned to a time-aware LSTM framework, capable of learning temporal dependencies and maintaining performance across folds and hydrological conditions (e.g., dry, wet, transition periods).

Feature selection for both models was guided by SHAP analysis to enhance interpretability, which identifies the most influential features contributing to model training (Lundberg & Lee, 2017) **(Figure 4)**.

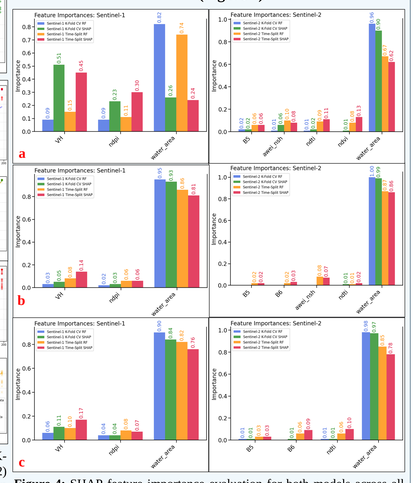


Figure 4: SHAP feature importance evaluation for both models across all areas; (a) Memphis, (b) Pontelagoscuro, and (c) Sogutluhan

ADVANCED APPROACH: LSTM MODELING

Model Development: The LSTM model was trained and validated using 5-fold cross-validation *TimeSeriesSplit* to evaluate temporal generalizability. The overall workflow is illustrated in **Figure 5**. The LSTM model, integrated with a transfer learning mechanism, provides the final evaluation on the entire sequential time series in **Fold 5**.

Fold 1 serves as the *initial* training phase. Therefore, its primary purpose is not to deliver high predictive performance, but rather to allow the model to learn the underlying temporal patterns. For this reason, Fold 1 results are not visualized.

In contrast, **Fold 5** represents the *final* evaluation stage. Thanks to the cumulative learning transferred from previous folds, the test accuracy in Fold 5 may even exceed the training accuracy, highlighting cumulative knowledge transfer learning, not overfitting similar to Pontelagoscuro region Dataset 2 and Sogutluhan region Dataset 1evaluation in **Figure 6**.

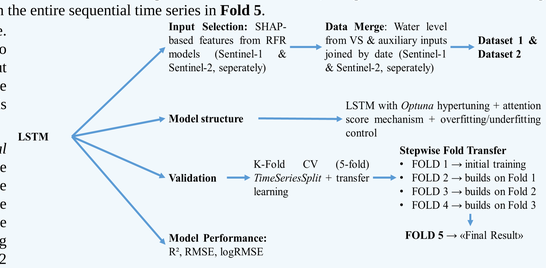


Figure 5: The workflow of step-wise fold transfer for time-aware LSTM model

Full-sequence river discharge estimation (Fold 5) indicated that both datasets capture dischare dynamic more effectively than the RF Time-split model **(Figure 6).** Dataset 2 slightly outperformed Dataset 1 in most metrics, particularly for peak flows, indicated by preliminary performance results **(Figure 7)**.

ADVANCED APPROACH: PRELIMINARY RESULTS

This section presents preliminary results of the time-aware LSTM model applied across all regions (**Figure 6)** using two distinct input configurations (**Dataset 1:** satellite-based data with Sentinel-1, altimetry water level, auxilary data with ground-truth measurement; **Dataset 2:** satellite-based data with Sentinel-2, altimetry water level, auxilary data with ground measurement). Additionally, scatter plots along full time-series of final results (Fold 5) illustrate the agreement between observed and predicted discharge values for both datasets with metric score evaluation in **Figure 7** with the preliminary metric results.

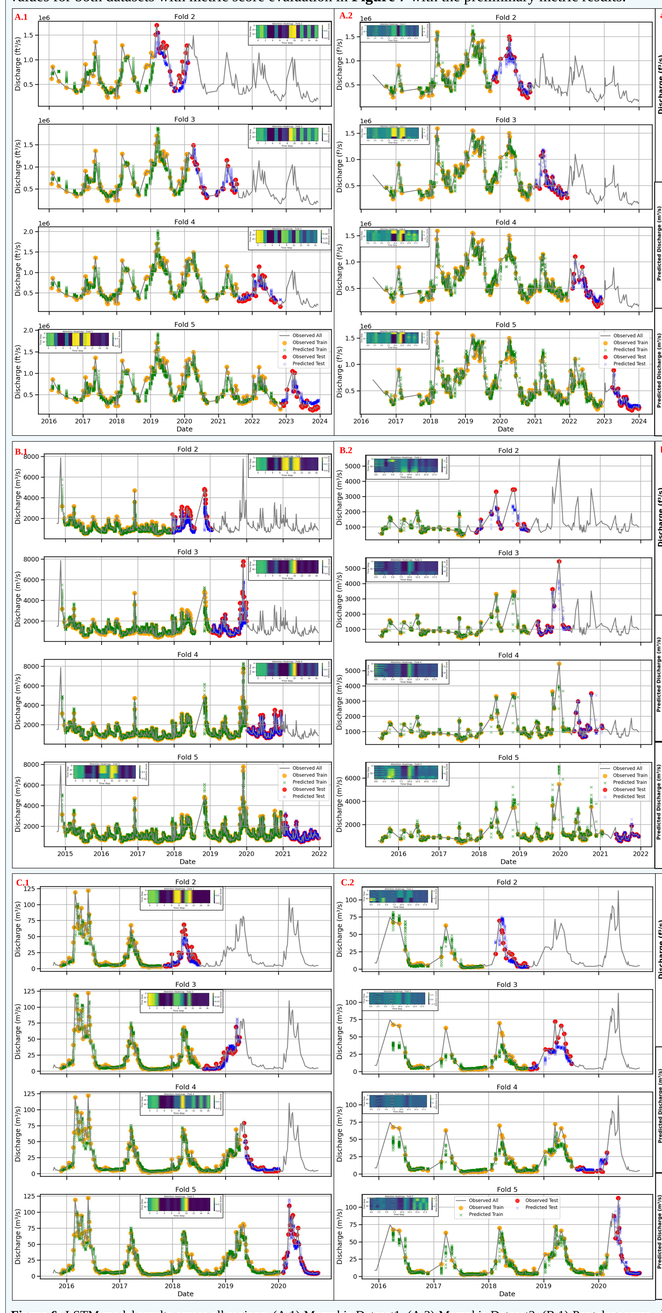
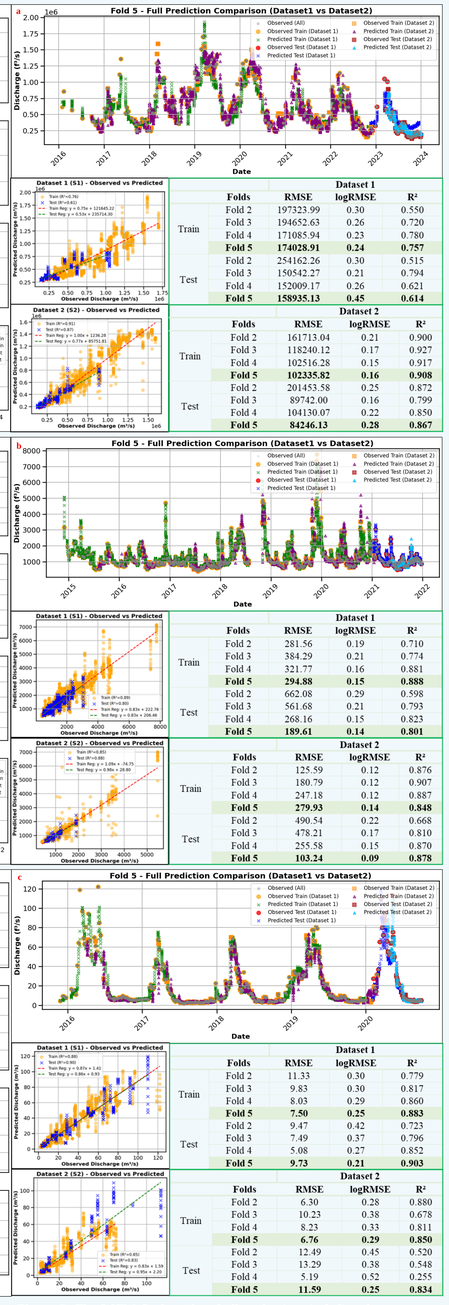


Figure 6: LSTM model results across all regions; (A.1) Memphis Dataset1, (A.2) Memphis Dataset2, (B.1) Pontelagoscuro Dataset1, (B.2) Pontelagoscuro Dataset2, and (C.1) Sogutluhan Dataset1, (C.2) Sogutluhan Dataset2. Time-series predictions across Fold 2 to Fold 5 are shown in the top panels.



Dataset 1

	Folds	RMSE	logRMSE	R^2
Train	Fold 2	197323.99	0.30	0.550
	Fold 3	194652.63	0.26	0.720
	Fold 4	171085.94	0.23	0.780
	Fold 5	**174028.91**	**0.24**	**0.757**
Test	Fold 2	254162.26	0.30	0.515
	Fold 3	150542.27	0.21	0.794
	Fold 4	152009.17	0.26	0.621
	Fold 5	**158935.13**	**0.45**	**0.614**

Dataset 2

	Folds	RMSE	logRMSE	R^2
Train	Fold 2	161713.04	0.21	0.900
	Fold 3	118240.12	0.17	0.927
	Fold 4	102516.28	0.15	0.917
	Fold 5	**102335.82**	**0.16**	**0.908**
Test	Fold 2	201453.58	0.25	0.872
	Fold 3	89742.00	0.16	0.799
	Fold 4	104130.07	0.22	0.850
	Fold 5	**84246.13**	**0.28**	**0.867**

Dataset 1

	Folds	RMSE	logRMSE	R^2
Train	Fold 2	281.56	0.19	0.710
	Fold 3	384.29	0.21	0.774
	Fold 4	321.77	0.16	0.881
	Fold 5	**294.88**	**0.15**	**0.888**
Test	Fold 2	662.08	0.29	0.598
	Fold 3	561.68	0.21	0.793
	Fold 4	268.16	0.15	0.823
	Fold 5	**189.61**	**0.14**	**0.801**

Dataset 2

	Folds	RMSE	logRMSE	R^2
Train	Fold 2	125.59	0.12	0.876
	Fold 3	180.79	0.12	0.907
	Fold 4	247.18	0.12	0.887
	Fold 5	**279.93**	**0.14**	**0.848**
Test	Fold 2	490.54	0.22	0.668
	Fold 3	478.21	0.17	0.810
	Fold 4	255.58	0.15	0.870
	Fold 5	**103.24**	**0.09**	**0.878**

Dataset 1

	Folds	RMSE	logRMSE	R^2
Train	Fold 2	11.33	0.30	0.779
	Fold 3	9.83	0.30	0.817
	Fold 4	8.03	0.29	0.860
	Fold 5	**7.50**	**0.25**	**0.883**
Test	Fold 2	9.47	0.42	0.723
	Fold 3	7.49	0.37	0.796
	Fold 4	5.08	0.27	0.852
	Fold 5	**9.73**	**0.21**	**0.903**

Dataset 2

	Folds	RMSE	logRMSE	R^2
Train	Fold 2	6.30	0.28	0.880
	Fold 3	10.23	0.38	0.678
	Fold 4	8.23	0.33	0.811
	Fold 5	**6.76**	**0.29**	**0.850**
Test	Fold 2	12.49	0.45	0.520
	Fold 3	13.29	0.38	0.548
	Fold 4	5.19	0.52	0.255
	Fold 5	**11.59**	**0.25**	**0.834**

Figure 7: The LSTM final results of all study sites, including the agrements between observed and estimated dicharge for each proposed area; (a) Memphis, (b) Pontelagoscuro, and (c) Sogutluhan.

TAKEAWAYS & OUTLOOK

- Attention layers improve interpretability by identifying the temporal relevance of input variables across folds.
- The fold-wise transfer learning approach enhances temporal generalization and helps stabilize predictions across unseen sequences.
- Due to LSTM's lookback structure and satellite data frequency limitations, integration with hydrologic models is essential for achieving daily and stable discharge estimates.
- Expand the methodology to basins with diverse climate and topography.
- Combine LSTM with other machine learning or physically-based models to improve generalizability, especially in mountainous or morphologically dynamic areas.
- Integrate hydrologic modeling components to enable daily discharge estimation.
- Operationalize the approach for automated, large-scale application in ungauged or data-scarce basins.