



## Chapter 4

# The Use of Machine Learning in Improved Hydrostatic Load Prediction for Inland Waterways Navigation Infrastructure

Mayank Chadha, Ying Zhao, Zhen Hu, and Michael D. Todd

**Abstract** The miter gate structure is a crucial infrastructure component that regulates waterway traffic and, therefore, can benefit from continuous health monitoring. Our past effort to build a decision-support digital twin of the miter gate structure requires creating a high-fidelity finite element model and its equivalent digital surrogate, capable of modeling the lifecycle of the structure. The primary loading that the miter gate is subjected to includes the upstream and downstream hydrostatic head. Therefore, the damage state of the miter gate depends on the operations of the gate (how many times the gate is opened and closed) as well as the river discharge. In our previous work, we assumed a statistical model to capture the water heads. However, better prognostics of the gate demand better knowledge of future river discharge. Traditional river routing methods are constrained by numerous assumptions and often result in noisy forecasts that diverge from true gauged values. We propose hybrid river discharge forecast models that integrate physics-based prediction models with state-of-the-art machine learning techniques. The proposed hybrid models leverage runoff data from the entire watershed, address the variability of the currently used traditional physics-based model predictions, and utilize advanced data-driven algorithms to forecast river discharge. This integration not only enhances the precision of discharge forecasts but also significantly improves the decision-support systems for miter gate structures and other real-world problems that require discharge forecasts.

**Keywords** Physics-Enhanced Machine Learning · Miter Gates · Hybrid Modeling · Lifecycle Management

## Introduction

Our ongoing research efforts in lifecycle asset management focus on river-related assets, primarily dams and miter gates [1–2]. The goal is to build a comprehensive structural model that is capable of (a) simulating the lifecycle of the structure while considering uncertainties in loading; (b) being updated to reflect the current state of the structure, diagnosed using sensor data obtained through a Structural Health Monitoring (SHM) system (model updating); (c) forecasting the future state of the structure for an assumed loading time series and determining the remaining useful life (prognostics); and (d) guiding decision-making regarding the operations and maintenance of the structure.

In the case of river-based structures, such as dams and miter gates, the primary external load is the upstream and downstream water head, which is directly connected to the river's discharge. Therefore, a key aspect of understanding the future behavior of the structure, albeit probabilistically, is to utilize a reasonable model for the water heads, which, in turn, are connected to the river discharge. Consequently, it is valuable to develop a more accurate discharge predictive model that could feed into the comprehensive structural model and help engineers better understand how the structure might behave in the medium-term future (such as over a year).

In our parallel ongoing research efforts, we have developed better, more reliable, physics-guided, data-driven, and probabilistic river discharge forecast models [3,4]. In this article, we detail the effectiveness of the hybrid discharge forecasting model and outline potential applications in lifecycle asset modeling and decision-making.

---

Mayank Chadha · Michael D. Todd

Department of Structural Engineering, University of California San Diego, La Jolla, CA 92093, USA  
e-mail: [machadha@ucsd.edu](mailto:machadha@ucsd.edu); [mdtodd@ucsd.edu](mailto:mdtodd@ucsd.edu)

Ying Zhao · Zhen Hu

Department of Industrial and Manufacturing Systems Engineering, University of Michigan-Dearborn, Dearborn, MI 48128, USA  
e-mail: [yinzha@umich.edu](mailto:yinzha@umich.edu); [zhennhu@umich.edu](mailto:zhennhu@umich.edu)

## Hybrid Modeling for River Discharge Forecasting

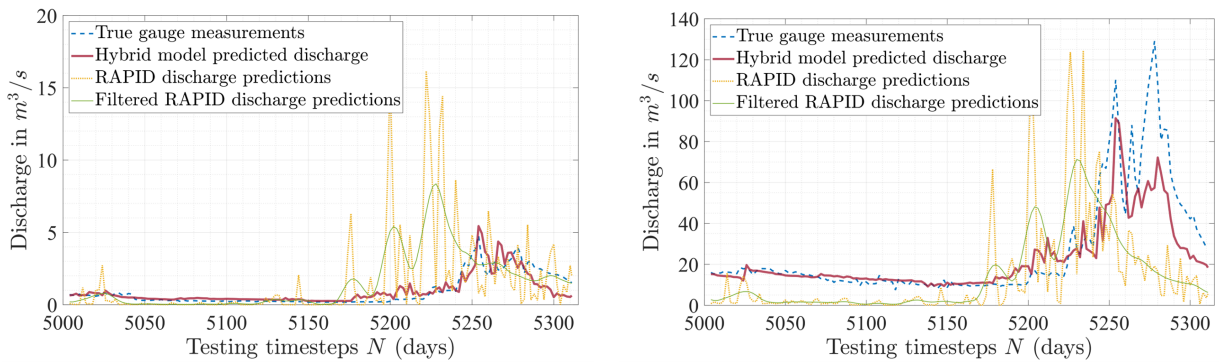
The National Water Information System (NWIS), operated by the U.S. Geological Survey (USGS), is the primary source of streamflow data in the United States. However, as of 2023, only 10,996 sites have automated monitoring systems to track current stream conditions. This is a small fraction considering the U.S. has around 250,000 rivers and over 3.5 million miles of waterways. The scarcity of discharge measurements from gauged locations makes stream and waterway predictions challenging. As a result, alternative methods, such as the Routing Application for Parallel Computation of Discharge (RAPID) model, are often used to estimate flows [5–8].

For this study, we considered building a river discharge prediction model for rivers with available gauge data (true observations). Readers are referred to Kratzert et al. [9] for methods of dealing with ungauged basins. To develop the hybrid model for predicting discharge over an extended period, we utilized the following data sources: (a) RAPID historical discharge predictions for the river of interest (a physics-based approximate model); (b) historical gauge measurement time series data for the river of interest (true values); and (c) runoff data for all rivers within the specified basin.

We perform three primary data processes on the mentioned datasets: (a) We use the Hanning filter to smoothen RAPID model predictions; (b) For  $N_{\text{rivers-in-basin}}$  numbers of rivers in the basin of interest and  $N_{\text{total-days}}$  number of days for which runoff data is available, the dimension of the runoff data is typically very large  $N_{\text{rivers-in-basin}} \times N_{\text{total-days}}$ . Since there is a correlation in the runoff of rivers within the same watershed (neighborhood effect), we apply Singular Value Decomposition (SVD) to reduce the dimensions of the runoff dataset; (c) To further reduce the dimensionality of all three datasets, we downsample the number of days when the machine learning (ML) method selected for building the hybrid model deems it necessary.

To build the hybrid ML predictive model, we deployed two different algorithms. The first is Delta Learning, which involves three major steps: (a) training a surrogate for RAPID that is informed by both past RAPID data and past runoff data, along with forecasted values for the near future available from Noah-MP; (b) using the true discharge value and RAPID surrogate outputs to calculate the error (the delta term) in the enhanced-RAPID predictions relative to the true discharge value; and (c) building an ML model that takes past RAPID data, past runoff data, and future forecasted data to predict the delta term. The ML models developed in the first and third steps can then be used to make discharge predictions in a recurrent manner. More details can be found in Section 3.1 of Zhao et al. [4]. The second algorithm is Data Augmentation, which also involves three steps for training: (a) training a RAPID surrogate similar to the Delta Learning algorithm; (b) although the RAPID surrogate in the previous step was trained on RAPID and runoff data, we obtain the output time series using the RAPID surrogate by replacing the past RAPID data with true discharge values, along with the runoff data, yielding an augmented prediction vector; and (c) finally, using the true discharge data, runoff data, and the augmented prediction vector from the previous step, we train an ML model capable of directly predicting the gauge measurements. More details can be found in Section 3.2 of Zhao et al. [4].

The hybrid ML models were implemented in a basin located in Colorado, consisting of 431 watersheds with a river running through them. Among these 431 watersheds, six have gauges, and these become rivers of interest. Figure 1 illustrates the prediction results for three of these rivers, built using the GP-NARX and Data Augmentation algorithm.



**Fig. 1** Discharge prediction using the hybrid ML models: (left) a case of a river with low discharge; (right) a case of a river with high discharge.

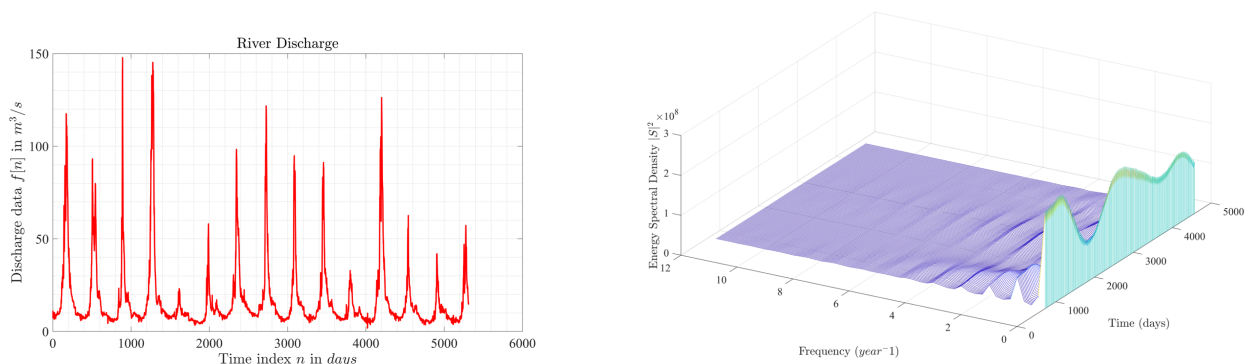
In addition to GP-NARX, we have also investigated more complex ML methods such as long short-term memory (LSTM) networks, bidirectional LSTM (Bi-LSTM), and NARX-Net. Readers are referred to Zhao et al. [4] for more details on the

implementation, results, and comparative performance of the various ML methods. As part of this ongoing effort, we have also developed the capability to obtain uncertainty bounds in the discharge predictions.

## Leveraging Discharge Forecasting Capability to Lifecycle Modeling and Decision-Making

River discharge and gauge height share a direct relationship with the upstream water head (and, by extension, with the downstream head controlled by the dam). Forecasting discharge over an extended period allows for the capture of (a) seasonality in loading (see illustration in Fig. 2); (b) intensity of loading; (c) uncertainty bounds in loading.

In our ongoing research effort, we plan to investigate the use of discharge prediction capabilities in (a) projecting the behavior of the structure in the near future as part of prognostics and RUL estimation; (b) assisting in modeling seasonality for the loading model used in Structural Health Monitoring (SHM) system design and lifecycle optimization; and (c) scheduling operations and maintenance plans. For instance, the seasonality in the discharge, as shown in the figure below, is crucial to include when simulating the structure's lifecycle.



**Fig. 2** Seasonality in river discharge: (a) river discharge time series; (b) spectrogram of river discharge data showing annual, semi-annual, and quarterly seasonality.

Other applications in river and asset management include: (a) early warning for flooding and decision-making for dam and miter gate operations; and (b) providing support for dredging operations.

## Conclusion

In this brief article, we discussed different algorithms and machine learning methods that are both physics-based and data-driven, which can be used to forecast river discharge more accurately than the currently existing physics-based approximations and for a longer period of time. This capability could have several implications for improving lifecycle simulations of river-based assets where the loading depends on river discharge. In addition, such predictive models can provide early warnings for floods and potentially save lives and property by enabling well-informed operational decisions.

## References

1. Yang, Yichao, Mayank Chadha, Zhen Hu, Manuel A. Vega, Matthew D. Parno, and Michael D. Todd. "A probabilistic optimal sensor design approach for structural health monitoring using risk-weighted f-divergence." *Mechanical Systems and Signal Processing* 161 (2021): 107920.
2. Vega, Manuel A., Zhen Hu, Yichao Yang, Mayank Chadha, and Michael D. Todd. *Diagnosis, prognosis, and maintenance decision making for civil infrastructure: Bayesian data analytics and machine learning*. Springer International Publishing, 2022.
3. Zhao, Ying, Mayank Chadha, Nicholas Olsen, Elissa Yeates, Josh Turner, Guga Gugaratshan, Guofeng Qian, Michael D. Todd, and Zhen Hu. "Machine learning-enabled calibration of river routing model parameters." *Journal of Hydroinformatics* 25, no. 5 (2023): 1799–1821.
4. Zhao, Ying, Mayank Chadha, Dakota Barthlow, Elissa M. Yeates, Charles J. Mcknight, Natalie P. Memarsadeghi, Guga Gugaratshan, Michael D. Todd, and Zhen Hu. "Physics-enhanced machine learning models for streamflow discharge forecasting." *Journal of Hydroinformatics* (2024).

5. McCarthy, Gerald T. *The unit hydrograph and flood routing*. Army Engineer District, Providence, 1939.
6. David, Cédric H., David R. Maidment, Guo-Yue Niu, Zong-Liang Yang, Florence Habets, and Victor Eijkhout. “River network routing on the NHDPlus dataset.” *Journal of Hydrometeorology* 12, no. 5 (2011): 913–934.
7. David, Cédric H., James S. Famiglietti, Zong-Liang Yang, Florence Habets, and David R. Maidment. “A decade of RAPID—Reflections on the development of an open source geoscience code.” *Earth and Space Science* 3, no. 5 (2016): 226–244.
8. Tavakoly, Ahmad A., Alan D. Snow, Cédric H. David, Michael L. Follum, David R. Maidment, and Zong-Liang Yang. “Continental-scale river flow modeling of the Mississippi River Basin using high-resolution NHDPlus dataset.” *JAWRA Journal of the American Water Resources Association* 53, no. 2 (2017): 258–279.
9. Kratzert, Frederik, Daniel Klotz, Mathew Herrnegger, Alden K. Sampson, Sepp Hochreiter, and Grey S. Nearing. “Toward improved predictions in ungauged basins: Exploiting the power of machine learning.” *Water Resources Research* 55, no. 12 (2019): 11344–11354.