Enhancing River Channel Dimension Estimation: A Machine Learning Approach Leveraging the National Water Model, Hydrographic Networks, and Landscape Characteristics

Arash Modaresi Rad[1,2*], J. Michael Johnson[1,2*], Zahra Ghahremani[1], James Coll[1,2], Nels Frazier[1,2]

[1]Lynker, Fort Collins, CO, USA

[2]Office of Water Prediction, National Oceanic and Atmospheric Administration, Tuscaloosa, AL, USA

Corresponding authors: Arash Modaresi Rad (amodaresirad@lynker.com) and J. Michael Johnson (jjohnson@lynker.com)

Key Points:

- We present a data-driven approach for predicting channel geometry using a machine learning (ML) method.
- ML predictors include National Water Model flow statistics and a suite of landscape characteristics.
- Our method exceeds the goodness of fit of its predecessors evaluated across regions of the United States.

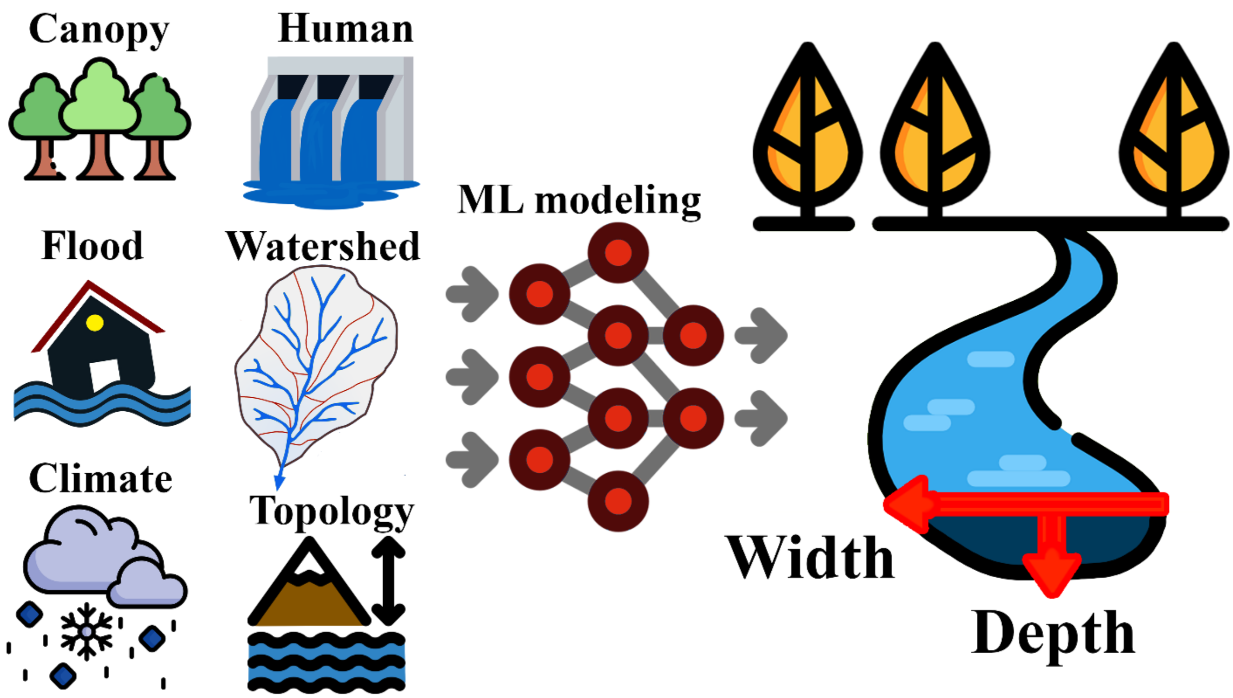
Abstract

Knowledge of bankfull hydraulic geometry represents an essential requirement for various applications, including accurate flood prediction, hydrological routing, river behavior analysis, river management and engineering practices, water resource management, and beyond.

Our work builds upon an extensive body of literature about estimating bankfull top-width and depth at ungauged locations to enhance the understanding of observable factors that affect these parameters. Using more than 200,000 USGS Acoustic Doppler Current Profiler (ADCP) records, we developed a method employing machine learning (ML) using discharge estimates and landscape characteristics from sources, including the National Water Model (NWM), the National Hydrologic Geospatial Fabric network (NHGF), the EPA stream characteristic dataset (StreamCat), and an array of satellite and reanalysis data products. Our method achieved R^2=0.79 predicting bankfull depth (R^2=0.84 for in-channel conditions) and R^2=0.81 predicting bankfull top-width (R^2=0.8 for in-channel conditions) in the testing dataset. The depth predictions showed high skill in plateau regions and low skill in mountainous regions. Our analysis demonstrates the benefit of data-driven modeling in contrast to other global scaling-based or regional statistical methods. In summary, our study illustrates how top-width and depth can be better predicted using ML, reanalysis streamflow simulations, hydrographic networks, and summarized geospatial data.

Keywords: at a station hydraulic geometry, river characteristics, river dimension, hydrofabric, machine learning, National Water Model

Graphical abstract



Plain language summary

Accurately estimating (or generalizing) key characteristics of river channels, such as their top-width and depth, is valuable for tasks like predicting water flow, modeling water-related processes, and mapping flooded areas. Our research builds on existing studies that focus on estimating these important channel characteristics and aims to further develop knowledge and skills in predicting these channel characteristics. In this work, we use over 200,000 historical measurements of channel top-width and depth to develop a machine learning (ML) model to estimate channel top-width and depth. The model uses widely available information from the National Water Model (NWM) discharge and other datasets that represent land surface characteristics, climate, hydrographic connectivity, and human-related structures. The developed model performs well compared to other global, regional, and ML-based methods in the literature within the Continental United States. Validation of the models across different regions indicated

better performance in flatter regions and lower performance in steeper areas. In conclusion, the study highlights the advantages of using ML techniques to estimate channel geometry more accurately, paving the way for improved predictions in unmeasured channels.

Introduction

The use of accurate estimates of channel bankfull depth and top-width improves channel flow routing models (Bindas et al., 2024; Brackins et al., 2021a; Getirana et al., 2013; Han et al., 2020). These dimensions inform generalized cross-sections in large-scale models, and their adequacy can influence hydrological forecasting (Brackins et al., 2021b; Brakenridge et al., 2012; Cohen et al., 2019; Heldmyer et al., 2022) and products that rely on them, like flood map generation (Alfieri et al., 2018; Cohen et al., 2018; Johnson et al., 2019).

Apart from large-scale modeling, precise estimates of bankfull depth and top-width improve flood risk analysis and mapping. These accurate estimations act as a proxy, enhancing the representation of channel volume below standard DEM elevations. (Bates & De Roo, 2000; Sichangi et al., 2018; Yamazaki et al., 2009).

The interest in estimating bankfull depths and top-widths has a long history in academic literature. To date, there have been efforts to present global equations that establish a relation between discharge and bankfull top-width or depth following the geomorphic relation proposed by Leopold and Maddock (1953). One of the earliest of these efforts was proposed by Moody and Troutman (2002) for global channels (Eqs. 1 and 2).

$$w = 7.2Q^{0.5\pm0.02} \quad (1)$$

$$d = 0.27Q^{0.3\pm0.01} \tag{2}$$

where *w* and *d* are bankfull top-width and depth, respectively, and Q is the discharge.

Frasson et al. (2019) proposed an alternative drainage area relation to estimate channel top-width, hypothesizing and affirming that channel top-width is directly associated with catchment area and channel meander wavelength. They aimed to regionalize the above-mentioned global relations by using the Q/A ratio, where A represents the total drainage area and Q is discharge obtained from the Global Runoff Data Center (GRDC). This approach resulted in a global dataset (based on HydroBASIN; Lehner & Grill, 2013) of estimates for bankfull top-width and depth. They subsequently conducted a comparison of these estimates with a set of bankfull top-width records derived from Landsat data. However, these equations are limited to channel reaches below 60°N and top-width greater than the 30m resolution of Landsat scenes. In their validation, they found errors ranging from 8 to 62%. Other researchers have used these equations to compute bankfull estimates of channel top-width and depth to support routing attributes in land surface models (Han et al., 2020; Schumann et al., 2013).

Bieger et al. (2015) proposed regional equations based on regression for different physiographic divisions across the United States and found that annual precipitation and temperature provide additional information that improves channel top-width and depth predictions considerably. A subsequent study by Blackburn-Lynch et al. (2017) developed regional relations for all Hydrologic Landscape Regions (HLR) and physiographic provinces and reported higher goodness of fit (GOF) of the discharge-based relations in comparison to drainage area-based relations. These estimates were used in the development and deployment of WRF-Hydro and its implementation as the NOAA National Water Model. Recently, Neal et al. (2021)

addressed the challenge of missing channel bathymetry more explicitly by demonstrating that improved bed estimations derived from the simplified, gradually varied flow method significantly influenced the dynamics of floodplain inundation and storage during minor flood events.

ML models provide a superior alternative to simple regression models by efficiently learning from multidimensional and complex data, capturing non-linear relationships, and adapting to diverse feature types (Shen, 2018). ML models tailored for learning specific hydraulic or hydrologic variables, like channel dimensions, can be trained effectively using hydrographic datasets such as the National Hydrography Dataset (McKay et al., 2012), the National Hydrographic Geospatial Fabric (Blodgett et al., 2023; Bock et al., 2022) and its derived products (Johnson, 2022), Multi-Error-Removed Improved-Terrain (MERIT) Hydro (Dai Yamazaki et al., 2019), or the Surface Water and Ocean Topography (SWOT) Mission River Database (SWORD; Altenau et al., 2021), to name a few. Estimating channel characteristics on networks such as these provides a means to support a range of hydroscience use cases. In 2023, machine learning applications were introduced by Doyle et al. (2023), which further refined channel top-width and depth estimates within the United States. They achieved these refined estimates by using a random forest model parameterized with the Watershed (Ws) summaries sourced from the EPA StreamCat attributes (Hill et al., 2016). Their method demonstrated the value of using watershed-based predictors to estimate channel dimensions. However, when using a large suite of 96 predictors, high dimensionality (referred to as the 'curse of dimensionality'; Köppen, 2000) and high correlation between variables, such as population and housing density (~ 100%), create collinear predictors within the model, impairing its interpretability post-training (Chan et al., 2022), and make it less extensible. These challenges can lead to model confusion and a distorted representation of the actual responses, even in ML algorithms such as random forest methods (Ghahremanloo et al., 2021).

Aside from statistical and ML predictions, remote sensing is an alternative solution to estimate channel dimensions. There have been numerous studies of automated channel top-width extraction using satellite imagery (Durand et al., 2009; Golly & Turowski, 2017; Monegaglia et al., 2018; Pavelsky & Smith, 2008; Schwenk et al., 2017). In 2017, Isikdogan et al. (2017) developed the RivaMap software that automates the extraction of continental-scale river centerline and top-width for North American rivers using Landsat imagery, and more recently, the RivWidthCloud software using cloud computing (Google Earth Engine) to extract channel top-widths from a vast archive of Landsat imagery (Yang et al., 2019). Global Surface Water Explorer (Pekel et al., 2016) does not provide direct information about channel top-width, but its historical probability map of water occurrence can be used to distinguish riparian floodplain top-width, bankfull top-width, and in-channel top-width.

Concerning channel depth, satellites like ENVISAT and JASON can provide information on water surface elevation through altimeter measurements (Kouraev et al., 2004). The recently launched Surface Water and Ocean Topography (SWOT) satellite shows potential for improving space-based estimates of channel discharge globally (Durand et al., 2020; Emery et al., 2016). SWOT capability to measure water surface elevation can be used to capture variations in channel depth that cannot be directly measured from other remote sensing products such as MODIS, Landsat, and Sentinel products (Pavelsky et al., 2014). The major constraints on all remote sensing approaches are (1) the spatial resolution of the data, (2) the quality of the data (e.g., scan lines and cloud cover), and (3) the computational costs associated with image processing at regional to continental domains. In the case of (1) the 30-meter resolution of Landsat products and algorithms limit estimation to channels greater than 50 m in top-width at the time of imaging (Andreadis et al., 2013). Likewise, changes in water surface elevation are constrained by the 50 m resolution of SWOT, limiting the observable channels to those with top-widths exceeding 100 meters. (Baratelli et al., 2018). While these estimates are critical for the major river systems, supplementary information about the tributary systems that feed them is needed for a wide range of engineering, modeling, and design purposes.

In this research, we test the hypothesis that a meta-learning (ensemble ML) methodology informed by National Water Model (NWM) simulated flow characteristics (such as 100-0.1% annual exceedance probability discharges) and a suite of land surface and climate variables can predict bankfull top-width and depth with GOF equal to or better than previous methods. We hypothesize that the incorporation of NWM simulated flow characteristics and a suite of land surface and climate variables can significantly improve bankfull top-width and depth predictive capability. The rest of the paper is as follows: the methods section will describe the input data used and the model training procedure. The results will discuss the model outputs in relation to the observed data as well as existing global and regional curves and ML approaches. The discussion will highlight areas for improvement, applications of this dataset, and the advantages of taking a hydrofabric-centric approach grounded in evolving federal and international efforts.

2. Methods

2.1. Data

2.1.1 Observations

In this study, we use the HYDRoacoustic dataset in support of Surface Water Oceanographic Topography (HYDRoSWOT; Bjerklie et al., 2020; Canova et al., 2016). This dataset is composed of 200,000+ Acoustic Doppler Current Profiler (ADCP) measurements collected at 10,081 USGS stream gauging locations in the United States. From this, we use the recorded depth, top-width, velocity, and discharge for each campaign.

2.1.2. Data Filtering

With multiple, time-varying observations for the same location, statistical relationships can describe the relationship between streamflow, depth, top-width, and velocity. These relationships have traditionally been described using At a station Hydraulic Geometries (AHG; Leopold & Maddock, 1953; Shen et al., 2016).

The AHG relations are described as:

$$Q = TW \times Y \times V = aQ^b \times cQ^f \times kQ^m = ack \times Q^{b+f+m} \tag{3}$$

Where TW is the top-width, Y is the depth, and V is the velocity at a recorded streamflow Q value. Therefore, individual relationships can be described as follows:

$$TW = aQ^b \tag{4}$$

$$Y = cQ^f \tag{5}$$

$$V = kQ^m \tag{6}$$

and by definition:

$$a \times c \times k = 1 \tag{7}$$

$$b + f + m = 1 \tag{8}$$

Observations of hydraulic data, particularly over long periods, are inherently noisy. To reduce this noise, we leveraged the AHGestimation R package (Johnson et al., 2024) Click or tap here to enter text.which provides the ability to filter data based on statistical outliers prior to using an ensemble-based fitting method to ensure mass is conserved (Eqs 7-8) and error is minimized.

Profiles were only kept if (1) they were made between 2000-2015 (2) they had a depth less than 65 m (corresponding to the Hudson River - the deepest river in the United States) (3) they had a channel top-width of less than 4 km (corresponding to the widest parts of the Mississippi River) (4) their recorded discharge values are in agreement with min and max NWIS records of that site (5) no negative discharge, width, depth, and velocity records. Once reduced, any site that demonstrated an inverse relationship between discharge and depth or had less than 5 profiles were removed. After filtering each site, AHGestimation was used to fit equations 4, 5, and 6 to the filtered HYDRoSWOT data. The AHGestimation (Johnson et al., 2024) uses a combination of ordinary least square, a nonlinear least square, and a genetic algorithm to fit data while ensuring mass preservation (equations 7 and 8). From these fits we calculate the coefficient of determination (R^2). We took an arbitrary threshold of 0.6 that can explain more than half the variability in HYDRoSWOT data and retained sites with a $R^2 > 0.6$ from AHG fit. In total, 4,229 of the 6,226 initial sites were retained.

2.1.3. Bankfull and In-Channel top-width and Depth

To define in-channel and bankfull discharge at the selected HYDRoSWOT sites, we used the widely accepted definitions of discharge at 100 and 50% annual exceedance probability, respectively (Andreadis et al., 2013; Rosgen, 1994; Wilkerson, 2008; Woodyer, 1968). While recognizing that bankfull and in-channel flow vary across river reaches and correlate with different flood recurrence intervals, we simplify our approach by labeling 100 and 50% annual exceedance probability as in-channel and bankfull, respectively. We hypothesized that the 100% annual exceedance probability discharge corresponds to the absence of bathymetry data, wherein a digital elevation model (DEM) generates a flat bottom unable to penetrate water. Subsequently, the 50% annual exceedance probability discharge represents the next higher flood condition, a widely recognized term in the literature concerning the modeling of bankfull width (Andreadis et al., 2013).

Using historical daily NWIS discharge records (retrieved from DeCicco et al., 2023) at each HYDRoSWOT site, we computed an annual maxima series and assumed the series may follow either a generalized extreme value (GEV), generalized Pareto (GP), Log-Pearson Type III (LP3), or generalized gamma distribution (Metzger et al., 2020; Zhang et al., 2021). To identify the distribution that best describes the underlying data, the Kolmogorov-Smirnov (KS) method was used (Ahmad et al., 1988). The selection of the most suitable distribution was determined through the KS test, which involved computing the KS statistic between the candidate distribution and the empirical distribution derived from the observed data. Results showed the majority of sites followed a GEV and GP distribution. Subsequently, the distribution yielding the smallest KS statistic was identified as the best-fitting model and then used to compute bankfull and in-channel discharge. The top-width and depth were obtained using the defined AHG relation. Figure S1 shows several HYDRoSWOT sites with corresponding bankfull and in-channel computed discharges.

The river network used in this study is the National Hydrologic Geospatial Fabric (NHGF; Bock et al., 2022), which is derived from the NHDPlusV2 with modifications to the topology and network characteristics based on feedback from a collection of federal agencies (Blodgett et al., 2023). Using this network, we compiled a list of possible predictors that could be used to explain variability in channel geometry. A list of all considered predictor variables is included in Table S1. These variables were selected based on literature and are composed of hydraulic, hydrologic, and climatological characteristics that affect channel dimensions.

Following the work of Doyle et al. (2023) and Blackburn-Lynch et al. (2017), we look to define a suite of catchment and watershed-level characteristics. Like those efforts, we collected a range of landscape characteristics for all river reaches across the United States using the EPA StreamCat data set (Hill et al., 2016). This dataset includes information on dams, land use, climate, hydrology, geology, and more. The outlet features of each total drainage area were aligned with the reference fabric identifier. In total, we used 58 watershed-level predictors from this dataset. In addition to these precomputed variables, a set of soil, landscape, and weather data was obtained and aggregated to the catchment level using climateR (Johnson & Clarke, 2019). This includes data from TerraClimate (Abatzoglou et al., 2018), POLARIS (Chaney et al., 2019), NLDAS, GLDAS (Rodell et al., 2004), USGS 3DEP, Leaf Area Index, and the Moderate Resolution Imaging Spectroradiometer (MODIS) mission. In total, we built 31 predictors from these sources.

Unlike prior efforts, we wanted to explore the impact of adding network connectivity and streamflow statistics to our predictors. Hydrographic information was taken from the NHGF, and in total, 10 predictors were used that represent general catchment and streamflow characteristics (see Table S1; source: Reference Fabric).

Streamflow statistics were generated for all NHGF reaches from the NWM v2.1 retrospective simulation (Johnson et al., 2023) and include 20-0.1% annual exceedance probability discharge (log Pearson Type 3) and the minimum, 25th, 50th, 75th, and maximum flow percentiles. As this study employed modeled flows, it is essential to note that these are regarded as "synthetic flow percentiles" due to discrepancies with the gauge measurements. In the training phase, although all stations contained floods from both NWIS and NWM, we randomly split sites such that half only contained NWIS bankfull, or in-channel flood, and the other half only contained NWM bankfull, or in-channel flood, and introduced a binary variable as an indicator of the absence of NWIS records. This allows the model to predict top-width and depth using accurate NWIS observations when available and NWM model data otherwise. During the testing phase, bankfull or in-channel floods are only derived from NWM.

As a first-pass quality check, Figure 1 plots the NWM and NWIS one- and two-year flow estimates against each other to assess the skill of the NWM in these lower flood recurrence intervals. While it has been shown that the NWM has less skill in predicting low flows as opposed to high flows (Fang et al., 2023; Johnson et al., 2023), here we test if the addition of greater floods as predictors compensates for this. Notably, the two-year flows show nice agreement ($R2 = 0.79$) with NWIS, while, as expected, the one-year flows show less ($R2 = 0.3$).

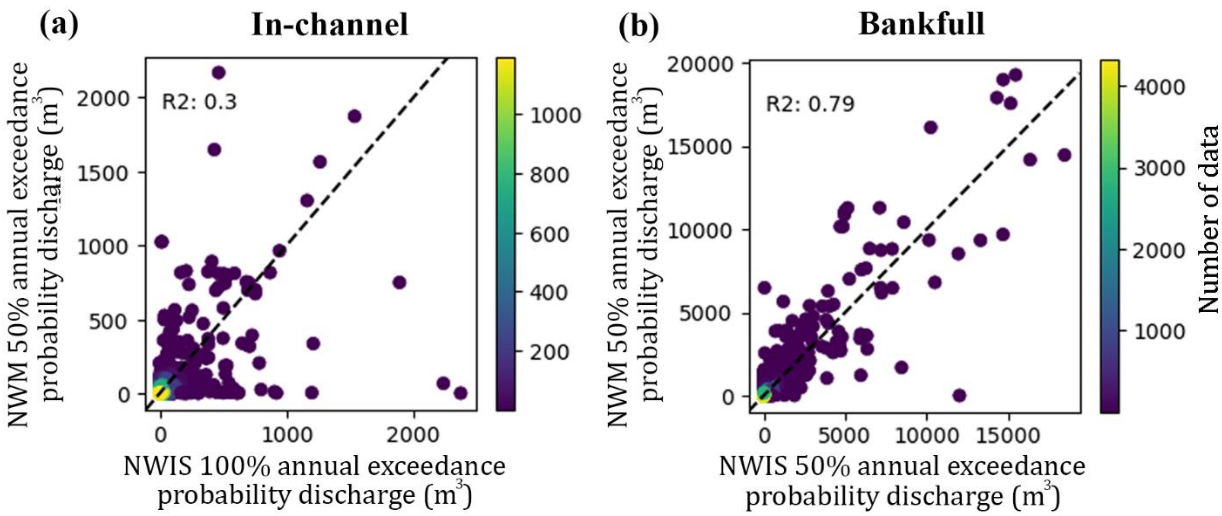


Figure 1. Comparison of NWM and NWIS flood discharges. (a) In-channel or 100% annual exceedance probability flow comparison; and (b) bankfull or 50% annual exceedance probability comparison.

2.2. Modeling

We trained four different machine learning models to predict top-width and depth for both bankfull (50% annual exceedance probability discharge) and in-channel (100% annual exceedance probability discharge) flow conditions across CONUS. This process is diagrammed in Figure 2.

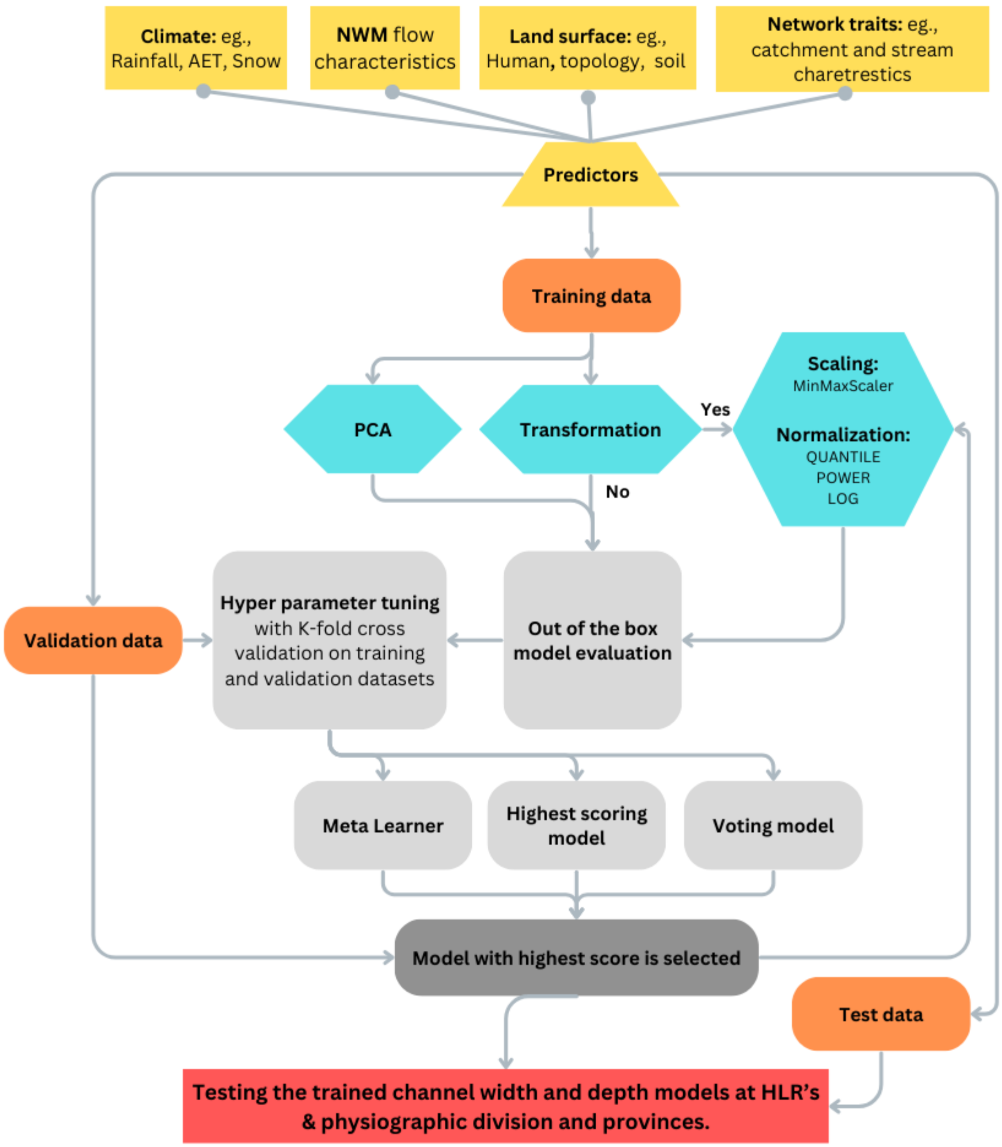


Figure 2. Schematic representation of model training for in-channel and bankfull flow top-width and depth.

2.2.1. Feature Space Reduction

Given the large predictor set selected for this study (116), we used two approaches to reduce the number of variables while retaining impactful information. This helps increase model generalization and interpretability while reducing computational requirements, noise, and dimensionality (Köppen, 2000). The first involves using SHapely Additive Explanations (SHAP; Lundberg & Lee, 2017) to quantify individual predictor contributions to the target variable. SHAP values reveal the contribution of each feature to the disparity between the ML model's prediction and the baseline prediction (ML model's average prediction). Positive values signify a feature contributes to increasing the model's prediction, while negative values denote the opposite. The magnitude of the SHAP value indicates the strength of the feature's influence. This process ranks feature importance, and progressively prunes the least significant predictors in a series of model retraining and validation cycles, and results in the identification of the balance between complexity and predictive skill.

To address predictor collinearity, we perform correlation analysis following Chan et al. (2022). We grouped variables based on their correlation and applied Principal Component Analysis (PCA) to each cluster to create new composite features (Sharma et al., 2015), which we use as inputs into the model (Figure S2). We specified a target explained variance of 95% to ensure the retained components collectively account for most of the dataset's variability while reducing dimensionality (Cruz-Cárdenas et al., 2014). PCA results across the different categories and their impact on generated components are shown in supplementary materials (Figures S3-11). We employed the "elbow" heuristic strategy to identify an optimal feature count, which involves iteratively excluding less informative features while monitoring the coefficient of determination (R-squared, denoted as 'R^2') (Liu & Deng, 2020). Upon reaching a plateau in the R^2 value (suggesting the optimal number of features), we identified a subset of 15 features while the R^2 value remained close to 0.80, signifying these as the most relevant predictors.

2.2.2. Modeling Channel Geometry

The training process involves an out-of-the-box evaluation by training 40 different ML models, which include neural network, tree-based, and support vector machine approaches to select top performing models in prediction of top-width and depth as separate models (train on training data and test on validation data; Figure 2). Then, R^2 and Root Mean Square Error (RMSE) metrics were used in the objective function to compare the predicted to observed top-width or depth values, and the 10 top-performing models were selected prior to hyperparameter tuning (see Table S2). The selected models were then fine-tuned using a 5-fold cross-validation with 3 repeats and a systematic exploration of hyperparameter combinations. Ultimately, four different ML models were developed for the estimation of top-width and depth under bankfull and in-channel conditions.

Next, we built a meta-learner and a voting model on the fine-tuned models to harness collective intelligence. Both the meta-learner and voting model, in this case, use the predictions of the top 10 fine-tuned selected models, with the exception that a meta-learner has a meta-model stacked on top to learn from the 10 individual base-level models. These models leverage the strengths of the top 10 diverse models while minimizing the impact of weaknesses or overfitting that may occur in any single model. Therefore, we hypothesized that they could be more accurate than the individual models.

Applying this final step helps capture unique insights from the data, reduces overfitting risk, and enhances overall model performance by leveraging the strengths of the top ten models. The entire process is conducted on training (70%) and validation (10%) splits of our datasets, reserving testing (20%) split only for model performance comparison. During the training and validation phases, the ML configuration with no data transformation or scaling for both predictors and target variables yielded the highest R^2. This is consistent with the inherent properties of tree-based algorithms that are non-parametric and base their decisions on splitting criteria. However, in the case of neural networks, we identified power transformation and scaling (among the tested log, power, and quantile transformations) as the most suitable techniques during training.

Our out-of-box evaluation showed that XGBRegressor, RandomForestRegressor, HistGradientBoostingRegressor, LGBMRegressor, ElasticNet, MLPRegressor, BayesianRidge, ARDRegression, KNeighborsRegressor, and BaggingRegressor were often among the top 10 candidates for our ensemble approach. The highest R^2 was obtained through the meta-learner ensemble method. A comprehensive list of the models tested, and an example of the outputs used in the selection process are provided in Table S2. Following the completion of model training, we utilize the reference fabric aggregated features outlined in section 2.1.4 (Channel Predictors) to predict channel width and depth under both bankfull and in-channel conditions for approximately 2.7 million river reaches across the CONUS.

2.3. Hydrological Traits Impact on Channel Geometry

We delve deeper into the impact of features on model performance by examining the top predictors identified through feature importance analysis. This involves categorizing the values of significant features into quartiles and assessing the model's skill (R^2) in predicting channel geometry, including channel width and depth under both in-channel and bankfull conditions. By doing so, we aim to uncover any potential biases of the model towards favoring certain hydrological characteristics.

To assess the impact of varying magnitudes of predictor variables on channel geometry, we categorized sites based on influential predictors into four quantile ranges: 0-25%, 25-50%, 50-

75%, and 75-100% for each predictor. Subsequently, we evaluated how the model's skill (R^2) evolves across these quantile ranges, thereby revealing any potential modeling bias towards specific predictor variables. The presence of discernible patterns in model performance (such as positive or negative trends from lower to higher quartiles) indicates potential bias in certain river segments towards either high or low values of the respective attribute. We restricted our analysis to the four most influential variables identified by SHAP.

We utilize SHAP values to comprehend how crucial features influence model outputs (i.e., channel width and depth). This involves plotting the SHAP value of a feature against its value for all examples in the dataset. By doing so, we illustrate how predicted channel dimensions change as the feature of interest varies, offering insight into complex interactions. This approach may uncover feature (hydrological) thresholds that influence the model's estimation of channel depth and width, deviating from simple monotonic relationships between feature values and channel dimensions. Instead of consistently increasing or decreasing with changes in feature values, these thresholds may lead to instances where certain increases in feature values result in a reversal of the model's predicted behavior.

2.4. Literature Comparisons

Here, we compare the performance of our proposed ML models to other modeling efforts documented in the literature. The goodness-of-fit metrics employed to compare the proposed machine learning model with existing literature include the R^2 coefficient (permitting direct comparison to previous studies), Kling–Gupta efficiency (KGE), and the normalized root mean square error (NRMSE). We compare our results to that of Blackburn-Lynch et al. (2017), which introduced regional relations between discharge and channel top-width and depth based on hydrologic landscape regions (HLR; Figure S12; Wolock et al., 2004) and physiographic provinces. We took the reported R^2 values from Blackburn-Lynch et al. (2017) and juxtaposed them with those predicted by our ML approach across all test sites, grouped in individual HLR and Physiographic provinces.

Further we compare our model skill to that of global equations founded on drainage area and discharge, as introduced by Frasson et al. (2019) and Andreadis et al. (2013), respectively. To facilitate this comparison, we applied these equations to our HYDRoSWOT dataset, using the total upstream catchment area along with 100 and 50% annual exceedance probability discharge as input. We also compare our predictions to the recently developed ML method of Doyle et al. (2023), where we extracted the predicted top-width from their study using the values that are published in StreamCat data (Hill et al., 2016). Next, we aligned the stream segment IDs (COMIDs) from their dataset with ours to ensure consistency in comparing stream segments. Subsequently, we evaluated our predictions of channel top-widths and depths alongside those of Doyle et al. (2023) against the observed HYDRoSWOT data.

3. Results

3.1. In-channel and Bankfull Model Performance

For channel top-width, the model R^2 was approximately 0.82 for bankfull and in-channel conditions, and for channel depth, the model R^2 was approximately 0.80 (Figure 3).

Figure 3 shows the performance of all four models on unseen test data. The top-width models are slightly more accurate (in terms of R^2) than the depth models, a trend that is consistent with previous studies (Booker & Dunbar, 2008). In general, the models underpredict both top-width and depth as channel top-width and depth increase (scatter dots located on the right side of the 1:1 line in Figure 3. c-f). This observation can be attributed to the skewed nature of hydrologic observations, with proportionately fewer data points collected from larger channels (in HYDRoSWOT), resulting in less training data for the machine learning model (Krawczyk, 2016).

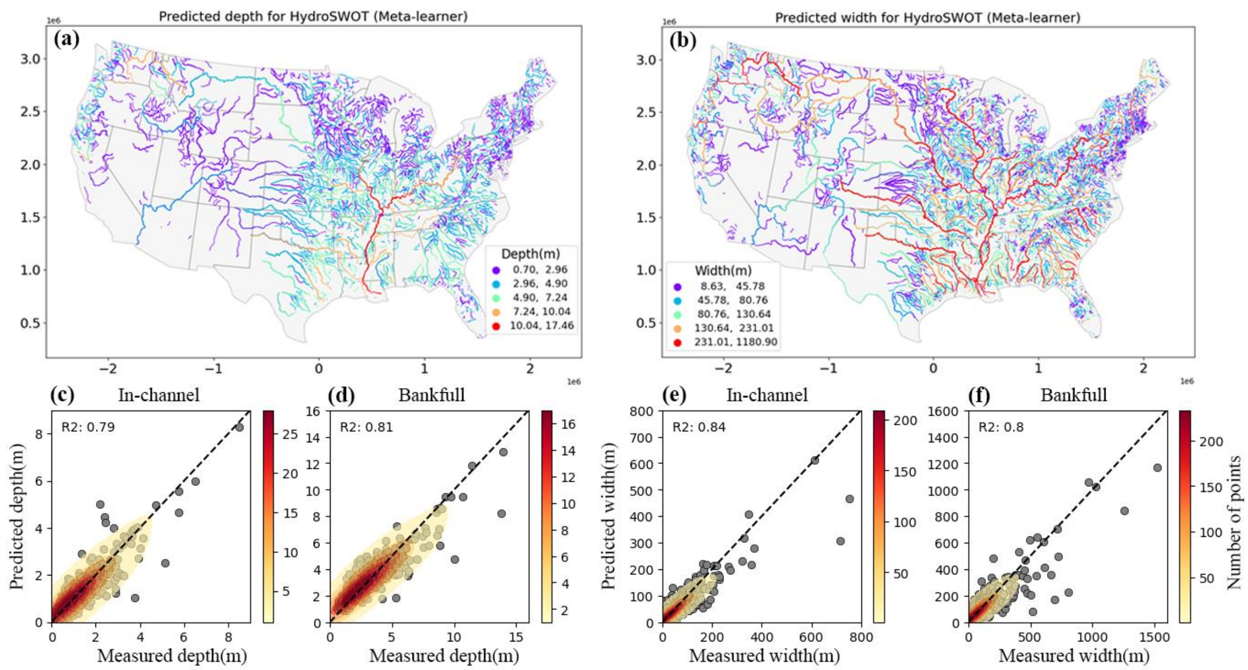


Figure 3. Predicted bankfull depth and top-width for all used stations in HYDRoSWOT data mapped to reference fabric flowlines using stream segment IDs (COMID). using meta-learner (subplots a and b, respectively). Subplots c and d represent the meta-learner channel depth GOF for in-channel and bankfull flow conditions. Subplots e and f represent the meta-learner channel top-width GOF for in-channel and bankfull flow conditions.

3.2. Impactful Predictors

Figure 4 shows the significant features for the prediction of bankfull channel top-width and depth. It visualizes the SHAP values of each feature across all samples (dots on the plot). It also arranges features based on the cumulative sum of SHAP value magnitudes across all samples, and utilizes SHAP values to illustrate the distribution of each feature's impact on the model output. The most influential feature identified is bankfull discharge and the 0^{th} flood frequency principal component (Figure 4; NWM Flood pc 0). Focusing on top-width (Figure 4a), the most pivotal variables are the cumulative lengths of all upstream flowlines (Arbolatesu), the topological wetness index (TWI), and actual evapotranspiration (AET). All three variables exhibit a positive relationship with channel top-width, as indicated by SHAP importance values.

Concerning the feature importance of the depth prediction (Figure 4b), the most influential variables include bankfull flow estimates, mean catchment elevation, channel slope (Slope), and the base flow index. Upon comparing the distinctions between the two, it's evident that channel depth is influenced by topographical factors (elevation and slope), groundwater recharge (base flow index), and soil texture. Conversely, channel width is influenced by cumulative stream lengths (Arbolatesu), catchment area (Totdasqkm), TWI, and precipitation.

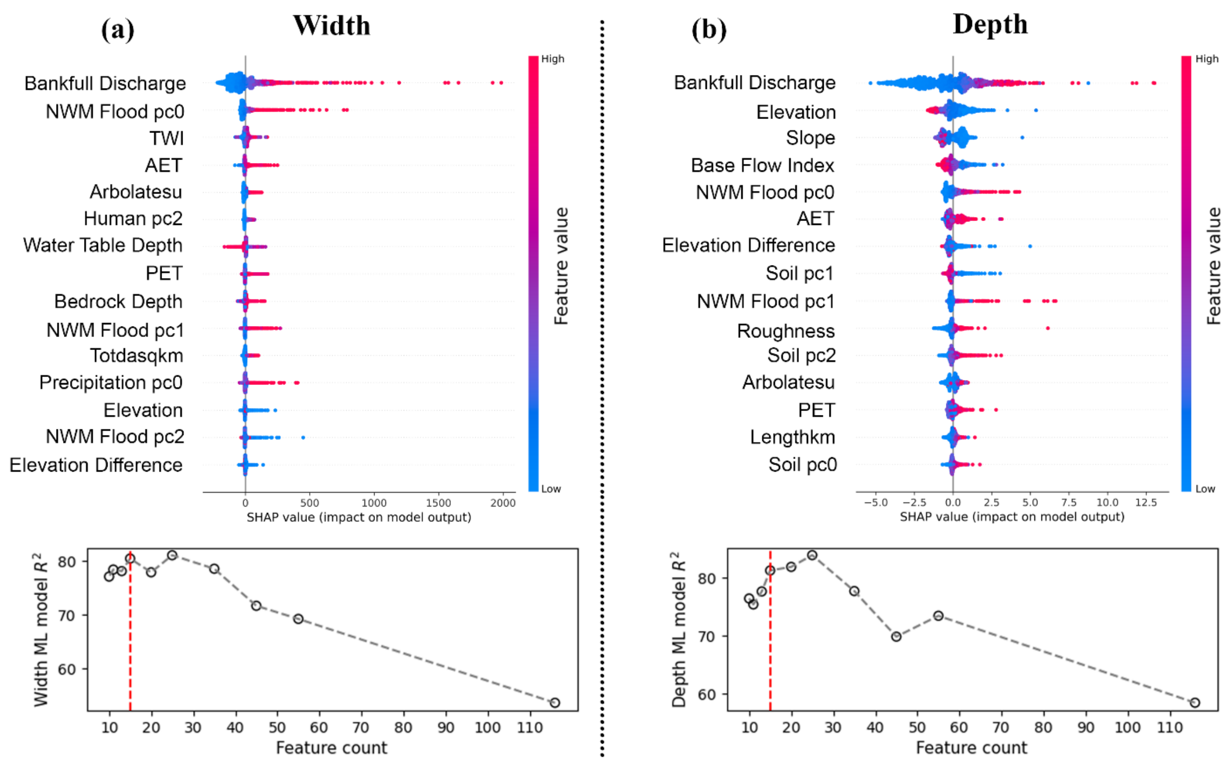


Figure 4. SHAP features importance values, distribution, and impact on model prediction for bankfull top-width (a) and depth (b). All NWM Flood principal components (PCs) are uncorrelated representations (PCs are orthogonal to each other) of 20-0.1% annual exceedance probability discharge. Human PC2 is a representation of the total reservoir volume in the catchment. Soil PC0 represents soil moisture, Soil PC1 represents % of clay and silt, and Soil PC2 represents % of sand.

3.3. Impact of Predictors Magnitude on Model Performance

Next, we examined the four most influential features in each model by partitioning their values into quartiles, allowing us to assess model performance across segments as shown in Figure 5. In the case of channel top-width, both bankfull discharge and the flood frequency PC0 show similar patterns when assessing model performance (Figure 5a, b), showing that as channel top-width increases, modeling errors decrease. A pattern was also found in the arbolate sum (Arbolatesu; Figure 5c). The topographic wetness index (TWI; Figure 5d) values do not appear to have a strong pattern between quartiles and performance.

Looking at the depth model in Figure 5e, we see that as the channel slope decreases, model GOF (R^2) increases. Model performance declines when bankfull discharge is smaller (Figure 5f). Higher elevations (Figure 5g) also tend to negatively impact model GOF (R^2), and the model showcases its highest performance in regions dominated by runoff rather than groundwater contributions (baseflow index; Figure 5h).

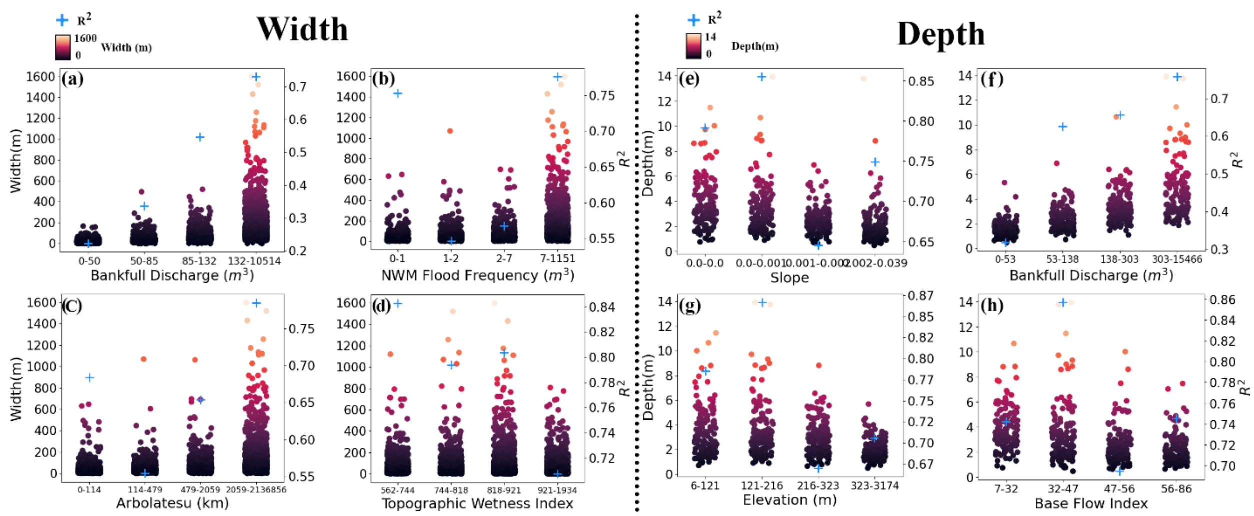


Figure 5. ML model performance across different quantiles of the most influential variables for both depth and width models. Within each subplot each dot represents a station belonging to a quantile of an important feature a, b, c, and d belong to grouped performances of top-width and subplots e, f, g, and h belong to depth. Th model skill (R^2) is shown as cyan plus sign.

3.4. Hydrological thresholds

It has been theorized that channel geometry follows stable hydrologic regimes (Rosgen, 1994). To elucidate this, we investigate the interrelation between key variables and their respective influence on model prediction, as depicted in Figure 6. In Figure 6a, the significance of the Topographic Wetness Index (TWI) in forecasting channel top-width is quantified. The absolute magnitude of the SHAP value serves as an indicator of the importance of each site (dots on the plot) in machine learning. A positive SHAP value suggests an increase in channel dimension, while a negative value indicates the opposite. The illustration reveals that channels with elevated bankfull discharge (depicted by purple dots) exhibit particularly informative data, characterized by high positive or negative SHAP values, during the machine learning model's training phase. Interestingly, there is an inflection point near TWI = 810. Here, high discharges (pink dots) linked with TWI < 810 result in a decrease in channel top-width (indicated by negative SHAP values), while values above TWI > 810 correspond to an increase in channel top-width. This is also observed when looking at bankfull discharge high (purple) and low (blue) values, as higher discharge values do not linearly contribute to greater channel dimensions. This suggests that channel morphology shifts after TWI = 810 and is most evident in larger channels (or higher discharge/pink dots). Similarly, an inflection point of AET = 65, as shown in Figure 6b, illustrates that areas characterized by AET < 65 contribute to shallower channels, and AET greater than this threshold is an indicator of deeper channels. The soil PC2, which is a representative of clay and silt components (see supplementary materials in Figure S3), is evenly scattered and shows no pattern.

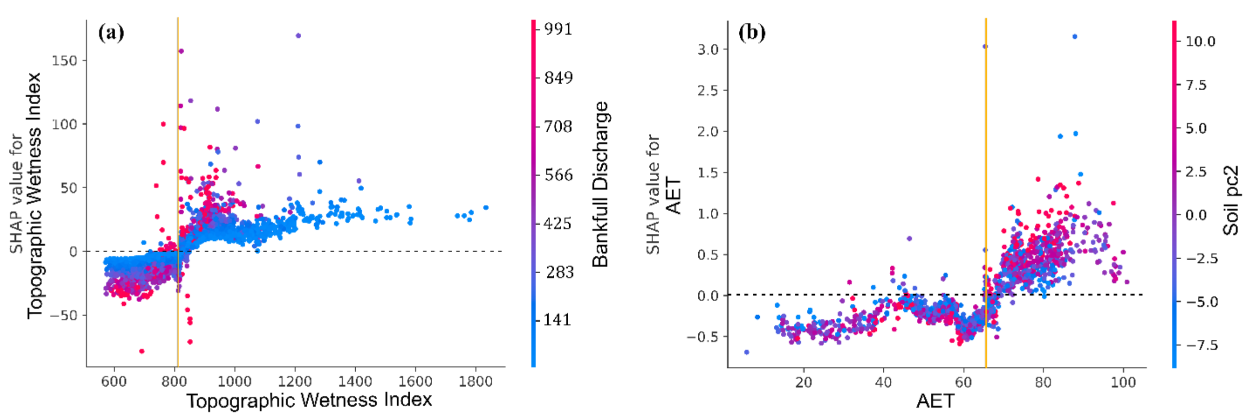


Figure 6. Representation of the contribution of each important feature and its inner correlation and impact on the top-width and depth models. (a) represents the impact of the topological wetness index on the channel top-width ML model and its correlation with the NWM 50% annual exceedance probability discharge. (b) represents the impact of AET on the channel depth ML model and its correlation with % of soil clay and silt content.

3.5. Performance in Hydrologic Landscape Regions

. It reveals a pattern where channel depth models show lower GOF (R^2) in the prediction of bankfull depth in regions with impermeable bedrock (HLR 16, 17, 18, and 19; Carlier et al., 2018; Santhi et al., 2008). Figures 7b and c show that the GOF of regional drainage area-based equations provided by Blackburn-Lynch et al. (2017) over the CONUS is poor (R^2 ~ 0.2-0.3). We took the reported R^2 values from Blackburn-Lynch et al. (2017) and compared it to the one predicted by the proposed ML approach in Figures 7d and e. Figures 7d and e, highlight the potential correlation between model performance, the distribution of the target variable, and the quantity of available points for machine learning. Results showed that the proposed depth model shows a significant improvement in R^2 across all HLR and physiographic provinces (see supplementary Figure S13; except for physiographic provinces 1 (Adirondack) and 23 (Southern Rocky Mountains)). In hydrological regions, HLR 16, characterized as humid mountains with permeable soils and impermeable bedrock, the ML model did not show improved GOF compared to the Blackburn-Lynch et al. (2017) study ($R^2 \cong 0.60$ of our model vs. $R^2 \cong 0.62$ reported). Moreover, regions with impermeable bedrock have lower range of measured channel depths (HYDRoSWOT measurements; as depicted in Figure 7d). Furthermore, we provide the improvements of ML-based approaches over discharge-based regional equations by Bieger et al. (2015) in supplementary information (Figure S13).

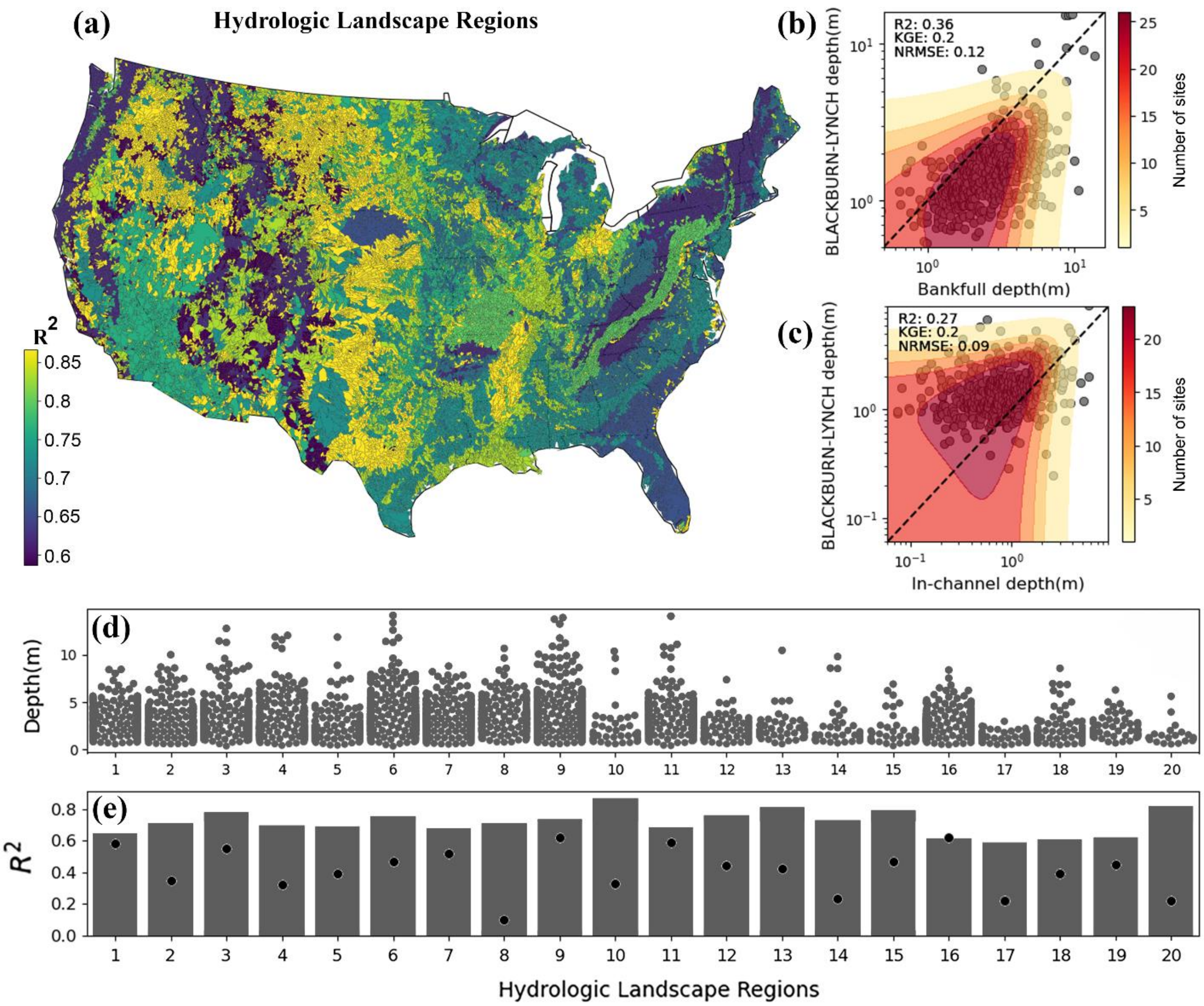


Figure 7. Depth ML model GOF (R^2) comparison with literature and across hydrologic landscape regions (a). (b) the GOF (R^2) of Blackburn-Lynch et al. (2017) derived depth from NWIS bankfull depth measurements. (c) GOF (R^2) of Blackburn-Lynch et al. (2017) derived depth from NWIS in-channel depth measurements. (d) shows the population of sites and distribution of depth per HLR in HYDRoSWOT. (e) proposed ML model GOF (R^2) comparison

(in columns) to Blackburn-Lynch et al. (2017) (in black dots) across hydrologic landscape regions.

Performing the same operations for the top-width ML model (Figure 8a) and comparing model skill (R^2) to Blackburn-Lynch et al. (2017) (Figure 8b and c), we see an overall higher R^2 performance in the tested ML model (Figure 8d and e and Figure S14). Looking at HLRs, we see that the lowest GOF comes from region 19, which is described as very humid mountains with permeable soils and impermeable bedrock with $R^2 \cong 0.75$. The highest GOF based on R^2 are seen in all plains (HLR 1-8) and plateau regions (i.e., HLR 9-13; Vidon & Hill, 2004). All these plateaus are categorized as surface flow response (in contrast to groundwater flow response) regions (Wolock et al., 2004). On the other hand, the mountainous regions in HLR 16-19 have the lowest GOF (in terms of R^2). Compared to depth estimates, the drainage area-based equations by Blackburn-Lynch et al. (2017) have a better skill in the prediction of top-width under in-channel conditions (Figure 8c; R^2 ~ 0.5) but are not appropriate for bankfull flows (Figure 8c; R^2 ~ 0.2).

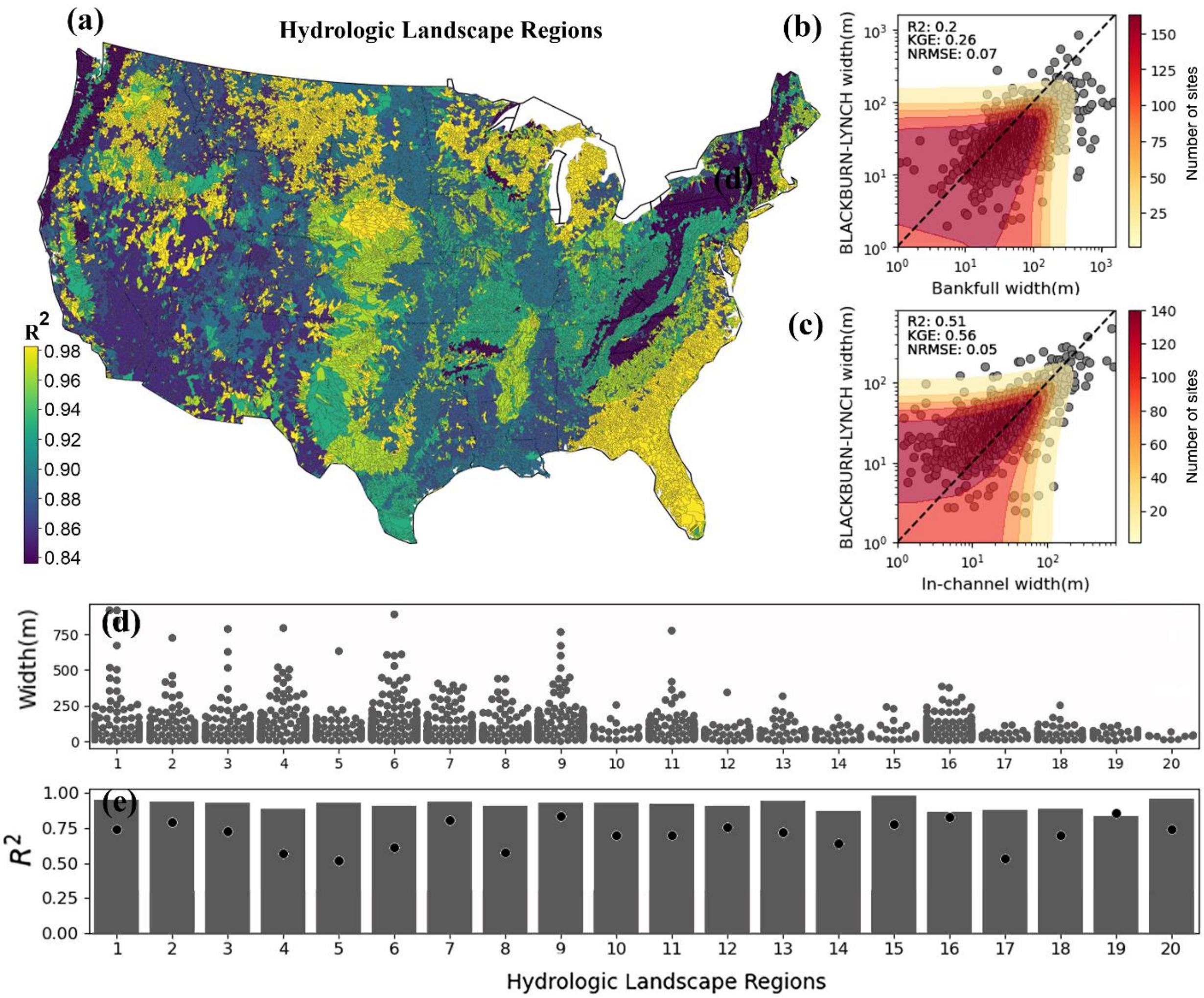


Figure 8. Width ML model GOF (R^2) comparison with literature and across hydrologic landscape regions (a). (b) the GOF (R^2) of Blackburn-Lynch et al. (2017) derived width from NWIS bankfull width measurements. (c) GOF (R^2) of Blackburn-Lynch et al. (2017) derived width from NWIS in-channel width measurements. (d) shows the population of sites and distribution of width per HLR in HYDRoSWOT (e) proposed ML model GOF (R^2) comparison (in columns) to Blackburn-Lynch et al. (2017) (in black dots) across hydrologic landscape regions.

3.6. Comparative Model Assessment

Following the substantial enhancement in performance when compared to regionally based discharge equations, we compare our results to global discharge and drainage area based questions and recent ML model by Doyle et al. (2023). Figure 9. shows the comparison of the global models and the recently proposed ML-based approaches, where ML models show a clear superiority over global relations.

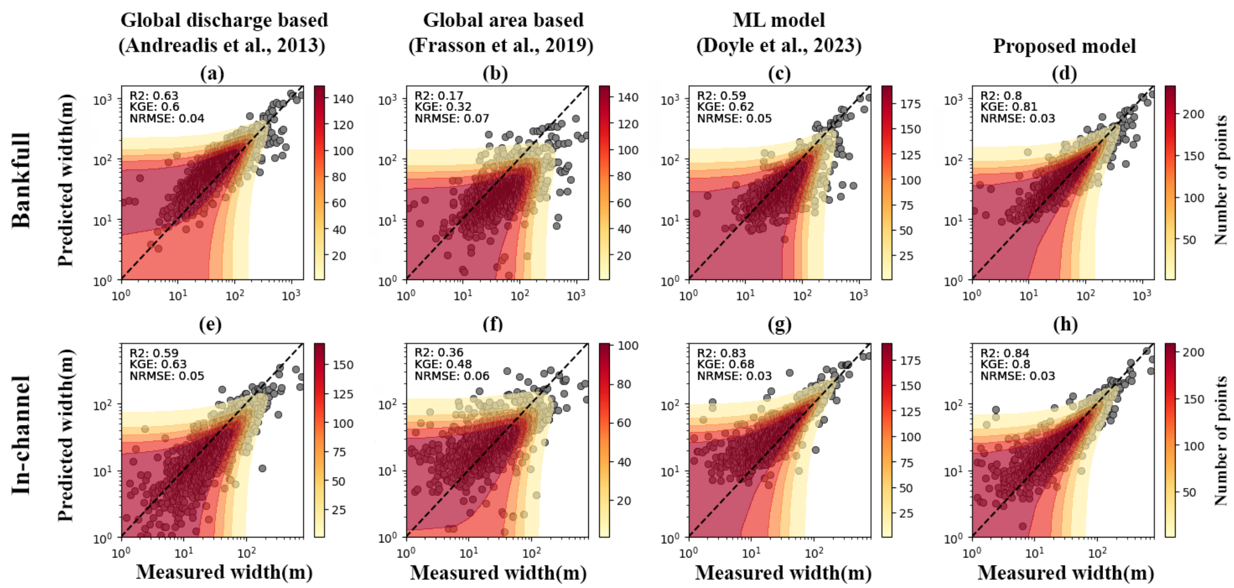


Figure 9. Channel top-width at bankfull and in-channel flow conditions model performance comparison between (a, e) global drainage area-based equations, (b, f) global discharge-based equations, (c, g) the ML model proposed by Doyle et al. (2023), and (d, h) our proposed model.

Using the same process, we can derive and compare our proposed depth model to other models in the literature (Figure 10).

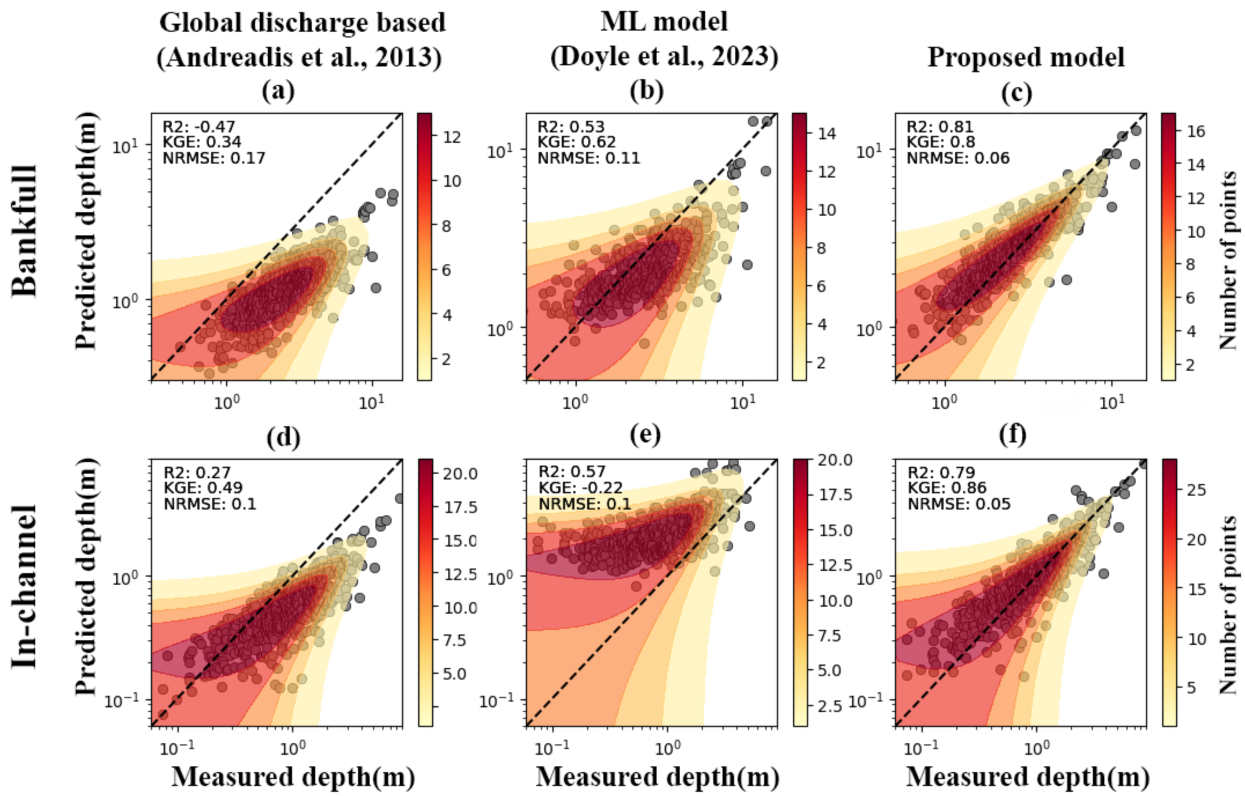


Figure 10. Channel depth at bankfull and in-channel flow conditions model performance comparison between (a, e) global drainage area-based equations, (b, f) global discharge-based equations, (c, g) the ML model proposed by Doyle et al. (2023), and (d, h) our proposed model.

With improved representation of discharge characteristics (from NWM), the proposed ML model has higher GOF (in terms of R^2) compared to its predecessors.

4. Discussion

4.1. Significance of Discharge in Channel Geometry Modeling

The outcomes of the SHAP features importance analysis suggest that the machine learning models have effectively captured the variations in channel dimensions while preserving the faithful representation of diverse hydrological and hydraulic processes and their interrelations. Conventional methods often employ empirical equations that relate these dimensions to drainage areas (Dunne & Leopold, 1978). More recent studies have focused on the application of bankfull discharge rather than drainage areas for developing channel dimension relations (Bieger et al., 2015). In Figure 6a we saw a nonlinear relationship between bankfull discharge/TWI and channel dimensionality. This non-linear relation was also reported by Erikson et al. (2024) recent study that showed regions exhibiting non-linear discharge scaling, bankfull channel dimensions increase more rapidly with drainage area compared to areas with linear discharge scaling. This suggests that the recurrence interval of the characteristic discharge determining channel geometry may be greater in regions with non-linear discharge scaling than in those with linear scaling.

We found the most influential feature identified for both depth and width is bankfull/in-channel discharge and is aligns well with the feature importance analysis conducted by Doyle et al. (2023) for wetted top-width and thalweg depth, as well as studies where authors employed in-situ measurements to establish relationships between discharge and channel characteristics using satellite imagery (Bjerklie, 2007; Bjerklie et al., 2005; Zakharova et al., 2020). Our findings underscore the significant role of using 50% annual exceedance probability discharge derived from NWM retrospective data when NWIS-derived 50% annual exceedance probability discharge is unavailable (the most important predictor shown Figure 4). The usage significantly enhances the GOF of prediction of channel top-width and depth, as indicated by the increase in R^2 value. It consistently (in width and depth under both flow conditions) demonstrates a strong positive correlation between higher magnitudes of the flood (Bankfull discharge and NWM

Flood PCs), and broader, deeper channels, aligning with the established literature (Bjerklie et al., 2005; Brown & Pasternack, 2014; Wohl & Wilcox, 2005).

Interestingly, global equations based on drainage area show significantly lower GOF (R^2), while methods grounded in discharge demonstrate notably higher levels of GOF. However, the use of hydraulic geometry curves does require a comprehensive understanding of the bankfull or in-channel discharge at specific locations of interest. To overcome this issue, Doyle et al. (2023) used the mean runoff from RunoffWs (StreamCat data; McCabe & Wolock, 2011) that is generated based on a 4 km × 4 km water-balance model for a period of 1900–2008. We improved upon this by replacing the water balance model with more accurate NWM retrospective data and derived various discharge characteristics by computing discharges of 20-0.1% annual exceedance probability and long-term average flow percentiles from NWM v2.1 historical archives (Michael et al., 2023). Lin et al. (2020) also identified discharge as the primary significant characteristic both in bankfull top-width and depth estimations. However, Doyle et al. (2023) identified the watershed area as the primary significant feature for their models in feature importance analysis. Contrarily, in our study, we demonstrated that a more accurate discharge from NWM (compared to Doyle et al.'s (2023) water balance model) can result in higher model GOF and the identification of discharge as the primary significant feature.

4.2. Other influential variables in Channel Geometry Modeling

Other most influential variables for the modeled depth were identified to be channel slope (Slope), mean catchment elevation (also reported by Lin et al. (2020) as the third important variable), and the base flow index. All these variables demonstrate a negative relationship with channel depth, suggesting that regions characterized by lower elevations and mild slopes (e.g., coastal areas) combined with lower groundwater to runoff contribution (low BFICat values) favor deeper channels (Biswal et al., 2023; Harman et al., 2008). These insights contribute to a comprehensive understanding of the multifaceted relationships between the examined features and channel dimensions (Morel et al., 2020).

In the case of width models, other influential variables were namely Arbolatesu, TWI, and AET that showed a positive relationship with channel top-width, as indicated by SHAP importance values. Practically, this can be interpreted to mean that regions with a high Topological Wetness Index, higher vegetation density and coverage (AET), or extended upstream drainage area tend to possess broader channels.

4.3. Utility and Limitations

The utility of bankfull top-width and depth estimates is evidenced in a broad body of literature and their application in large-scale modeling (such as the National Water Model) and observation (such as the Surface Water and Ocean Topography (SWOT) mission). Across the hydrological and hydraulic sciences, these estimates serve applications across studies encompassing research, engineering, modeling, and mapping applications. As model domains and resolutions expand, the capability to precisely estimate these attributes within high-resolution and dynamic hydrographic networks is becoming more important and useful (Archfield et al., 2015; Wood et al., 2011). Despite advancements in remote sensing applications for monitoring, the current resolution do not yet align with the necessary level of detail. Hence, there persists a demand for modeling and estimating these characteristics across diverse and heterogeneous landscapes.

To date, there have been efforts to generalize these relationships with respect to hydrographic (drainage area) and landscape (StreamCat) traits. These efforts encompass global and regional regression models, along with the application of random forest techniques across an extensive array of predictor variables at the catchment level. In this work, we aimed to enhance prior approaches by introducing two distinctive augmentations.

The initial addition was to incorporate comprehensive hydrographic details and historical streamflow model simulations from NWM aggregated based on reference fabric apart from other landscape attributes and drainage area. This establishes a dynamic, updatable channel geometry dataset for United States reference flow (Blodgett et al., 2023), notably applied in numerous integrated federal modeling initiatives. Consequently, our findings underscored elevation, slope, arbolate sum (Arbolatesu), drainage area, and length as pivotal predictors in estimating in-channel and bankfull flow conditions. The aggregation scale of these predictors emphasizes the potentially apparent yet significant role of network configuration in shaping bankfull flow conditions. Moreover, the incorporation of streamflow statistics derived from a comprehensive hydrologic model (National Water Model 2.1) demonstrated notable predictive power in establishing bankfull conditions. While Figure 1 highlights the level of agreement between the one-year (R^2 = 0.3) and two-year (R^2 = 0.79) flow conditions with the observed record, it is evident that the patterns in the data were able to help the ML model deduce more accurate estimates (based on R^2) across the entire network.

The subsequent addition to this research was to introduce a more robust ML method to address this challenge. We minimized the complexity of the model by retaining only the most impactful predictors. This process helped reduce the number of model predictors from 116 to 15. This represents a major improvement over previous efforts such as Doyle et al. (2023), who performed a similar test but concluded with 96 predictors, many of which exhibit high collinearity. Also, diligent attempts were undertaken to minimize data covariance and refine the input parameter space. A comprehensive evaluation encompassed a broad spectrum of "out-of-the-box" (40) and fine-tuned models (10), culminating in the development of a meta-learner model that harnessed the expertise and diversity of the collective models. These efforts yielded four models that exhibited enhanced predictive capabilities compared to existing methodologies. Beyond model development, this approach was applied to the entire NHGF network, resulting in outcomes integrated into the core data product. This integration actively improves data, making it more findable, accessible, interoperable, and reusable within the expanding hydrologic science data system.

Despite the numerous advantages, certain limitations warrant consideration. One limitation is the training data and its general applicability, which are constrained to USGS sites only. These sites are often situated in locations with high banks that confine the flow for ease of measurement and, as such, may not precisely represent the geometry of the entire reach. Also in estuaries, particularly near the river mouth where it widens into the coastal zone, there are no recorded ADCP measurements in HYDRoSWOT, and model accuracy is unreliable. This limitation in training dataset also result in unreliable estimates near manmade structures (e.g., dams and ponds). Another limitation stems from the primary significant feature of ML models, or the 100% annual exceedance probability discharge derived from the NWM, where the NWM diminished skill (as shown in Figure 1) in this exceedance probability, impacting the overall model GOF. While this method offers consistent reach-averaged channel geometry data for the entire CONUS, its resolution is constrained to the reach scale. The consequences of averaged reach-scale resolution can vary depending on the application. On one hand, it may lead to limitations such as overlooking intricate channel features like braiding, incision, and aggradation. On the other hand, it can offer benefits such as simplifying complexity and reducing computational efforts and can be fused with DEM derived products to represent the entire river reach. Achieving a resolution lower than reach scale would necessitate the utilization of satellite imagery for accurate width estimation. However, even with satellite imagery, resolving depth accurately at resolutions lower than reach scale remains a challenge in the absence of bathymetric related information.

Moreover, although our modeling strategy effectively reduced 112 variables into 15 components, the reliance on source data restricts the scalability of this approach to different geographical locations. Substantial data challenges associated with using existing NWM retrospective data (Johnson J. Michael et al., 2023) and its sheer size currently make this dataset globally unavailable, although less accurate global hydrological models can be substituted. Consequently, our model's applicability could be constrained in global regions where a comprehensive long-term historic modeled record is unavailable. The ML modeling method's applicability extends beyond a singular regression relation, thereby potentially posing challenges in its adaptation to new networks due to the complexity of the parameter set and model deployment. Nonetheless, envisioning future applications within the United States, the comprehensive coverage across our domain, coupled with the seamless integration into the continually evolving data system, holds significant promise for advancing the evolution and dissemination of hydrofabric data.

Anticipating advancements within the United States Next Generation Water Resource Model initiative led by NOAA and Cooperative Institute for Research to Operations in Hydrology (CIROH), the availability of this information will certainly help many ongoing efforts. Notably, the team spearheading the development of routing infrastructure for the NextGen system relies on approximations of general and natural cross-sectional data to facilitate Muskingum-Cunge and Diffusive Wave routing methods. The first of these can be improved or derived from a quality estimate of bankfull top-width and depth using existing methods like those in WRF-hydro or other channel geometry estimators (e.g., parabolic forms such as in Dingman, 2009). Additionally, the Flood Inundation Mapping team can incorporate estimated depths into their synthetic rating curve workflow (Johnson et al., 2019; Zheng et al., 2018), addressing the current limitations in accounting for bathymetry within DEM-derived hydraulic states. The inclusion of an estimated bankfull depth can help better estimate the volume of water in the channel and reduce instances of systematic over- or under-prediction (Johnson et al., 2019). While the effectiveness of these diverse use cases awaits validation, their accessibility through fundamental geospatial products supporting USGS, NOAA, and CIROH efforts will facilitate simplified testing and evaluation processes.

5. Conclusions

The measurements of channel top-width and depth, as well as their ratio, particularly at bankfull flow conditions, play a pivotal role in the fields of hydrology and river science (Bjerklie, 2007). These metrics hold profound importance for several key reasons. First and foremost, they are fundamental to characterizing the physical attributes and morphology of river systems, providing critical insights into the geometry of channels. (Luo et al., 2007)

Also, NASA's Surface Water and Ocean Topography (SWOT; Biancamaria et al., 2016) mission objectives, which encompass the precise measurement of channel discharge, water surface elevation, and variations in channel dimensions, rely on the calibration and validation of its remote sensing instruments, and the proposed method could play a pivotal role in ensuring the mission's accuracy and its valuable contributions to our comprehension of Earth's surface water dynamics. In smaller rivers and tributaries, estimating channel dimensions presents a considerable challenge, primarily due to the restricted spatial resolution of satellite data. Nonetheless, as showcased in this study, ML approaches trained on hydrographic information and model simulated discharge values provide an alternative method to adequately capture channel dimensions.

Our study showcases the efficacy of the current ensemble ML model in minimizing the necessary predictors while mitigating concerns related to multicollinearity and model confusion. Leveraging the extensive dataset of NOAA's Office of Water Prediction's (OWP) NWM flow characteristics encompassing 2.7 million reaches, we have achieved notably high GOF (reported as R^2) in predicting bankfull depth and top-width. Nonetheless, it is pertinent to acknowledge that the model's GOF diminishes with reduced discharge levels. This result again points to the challenge of getting equal skill in large and small segments of the network. This pronounced effect in smaller rivers and tributaries suggests a potential association with NWM's limited skill in representing base flow, a key determinant of channel dimension within ML models (as shown in Figure 5, there is a correlation between lower discharge and reduced model GOF). Certain constraints include the exclusive use of training data from USGS sites, primarily situated in areas with elevated banks, restricting flow and potentially not faithfully representing the complete reach geometry. We conducted extensive model evaluations across different hydrologic landscape regions and physiographic provinces and divisions. Our comprehensive analysis substantiates the capacity of our approach to augment the existing models documented in scientific literature.

Acknowledgement

We sincerely appreciate the invaluable contributions of the reviewers whose insightful feedback has significantly enhanced the quality of this paper. We thank Dr. Fred Ogden for his constructive comments and review of this manuscript. We also appreciate the generous support from the National Oceanic and Atmospheric Administration (NOAA), Office of Water Prediction (OWP) which has been instrumental in conducting impactful research in hydraulic geometry. Their funding, resources, and guidance have significantly contributed to our work's success.

Data Availability

The datasets used in this study, including the StreamCat data (Hill et al. 2016), Reference Fabric (Bock et al., 2022), HYDRoSWOT (Canova et al., 2016), NWM (Johnson et al., 2023), and ClimateR (Johnson & Clarke, 2019) are publicly available.

The ML predictions for channel depth and width are available at:

Modaresi Rad, A. (2024). Machine Learning-Derived Channel Width and Depth for the National Hydrologic Geospatial Fabric in CONUS, HydroShare, http://www.hydroshare.org/resource/d147fcf554a54b2aaa4f146f85da0e03

The developed code for this research are available at: https://github.com/LynkerIntel/bankfull_W_D

Disclaimer

The views expressed in this article do not necessarily represent the views of NOAA or the United States.

References

Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A., & Hegewisch, K. C. (2018). TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. *Scientific Data*, *5*(1), 1–12.

Ahmad, M. I., Sinclair, C. D., & Werritty, A. (1988). Log-logistic flood frequency analysis. *Journal of Hydrology*, *98*(3–4), 205–224.

Alfieri, L., Cohen, S., Galantowicz, J., Schumann, G. J. P., Trigg, M. A., Zsoter, E., et al. (2018). A global network for operational flood risk reduction. *Environmental Science & Policy*, *84*, 149–158.

Altenau, E. H., Pavelsky, T. M., Durand, M. T., Yang, X., Frasson, R. P. de M., & Bendezu, L. (2021). The Surface Water and Ocean Topography (SWOT) Mission River Database (SWORD): A global river network for satellite data products. *Water Resources Research*, *57*(7), e2021WR030054.

Andreadis, K. M., Schumann, G. J., & Pavelsky, T. (2013). A simple global river bankfull width and depth database. *Water Resources Research*, *49*(10), 7164–7168.

Archfield, S. A., Clark, M., Arheimer, B., Hay, L. E., McMillan, H., Kiang, J. E., et al. (2015). Accelerating advances in continental domain hydrologic modeling. *Water Resources Research*, *51*(12), 10078–10091.

Baratelli, F., Flipo, N., Rivière, A., & Biancamaria, S. (2018). Retrieving river baseflow from SWOT spaceborne mission. *Remote Sensing of Environment*, *218*, 44–54.

Bates, P. D., & De Roo, A. P. J. (2000). A simple raster-based model for flood inundation simulation. *Journal of Hydrology*, *236*(1–2), 54–77.

Biancamaria, S., Lettenmaier, D. P., & Pavelsky, T. M. (2016). The SWOT mission and its capabilities for land hydrology. *Remote Sensing and Water Resources*, 117–147.

Bieger, K., Rathjens, H., Allen, P. M., & Arnold, J. G. (2015). Development and evaluation of bankfull hydraulic geometry relationships for the physiographic regions of the United States. *JAWRA Journal of the American Water Resources Association*, *51*(3), 842–858.

Bindas, T., Tsai, W., Liu, J., Rahmani, F., Feng, D., Bian, Y., et al. (2024). Improving River Routing Using a Differentiable Muskingum-Cunge Model and Physics-Informed Machine Learning. *Water Resources Research*, *60*(1), e2023WR035337.

Biswal, S., Sahoo, B., Jha, M. K., & Bhuyan, M. K. (2023). A hybrid machine learning-based multi-DEM ensemble model of river cross-section extraction: Implications on streamflow routing. *Journal of Hydrology*, *625*, 129951.

Bjerklie, D. M. (2007). Estimating the bankfull velocity and discharge for rivers using remotely sensed river morphology information. *Journal of Hydrology*, *341*(3–4), 144–155.

Bjerklie, D. M., Moller, D., Smith, L. C., & Dingman, S. L. (2005). Estimating discharge in rivers using remotely sensed hydraulic information. *Journal of Hydrology*, *309*(1–4), 191–209.

Bjerklie, D. M., Fulton, J. W., Dingman, S. L., Canova, M. G., Minear, J. T., & Moramarco, T. (2020). Fundamental hydraulics of cross sections in natural rivers: Preliminary analysis of a large data set of acoustic Doppler flow measurements. *Water Resources Research*, *56*(3), e2019WR025986.

Blackburn-Lynch, W., Agouridis, C. T., & Barton, C. D. (2017). Development of regional curves for hydrologic landscape regions (HLR) in the contiguous United States. *JAWRA Journal of the American Water Resources Association*, *53*(4), 903–928.

Blodgett, D., Johnson, J. M., & Bock, A. (2023). Generating a reference flow network with improved connectivity to support durable data integration and reproducibility in the coterminous US. *Environmental Modelling & Software*, *165*, 105726.

Bock, A. R., Blodgett, D. L., Johnson, J. M., Santiago, M., & Wieczorek, M. E. (2022). PROVISIONAL: National Hydrologic Geospatial Fabric Reference and Derived Hydrofabrics: U.S. Geological Survey data release, . *Preprints 2022, 2022120390.*

Booker, D. J., & Dunbar, M. J. (2008). Predicting river width, depth and velocity at ungauged sites in England and Wales using multilevel models. *Hydrological Processes: An International Journal*, *22*(20), 4049–4057.

Brackins, J., Moragoda, N., Rahman, A., Cohen, S., & Lowry, C. (2021a). The Role of Realistic Channel Geometry Representation in Hydrological Model Predictions. *JAWRA Journal of the American Water Resources Association*, *57*(2), 222–240.

Brackins, J., Moragoda, N., Rahman, A., Cohen, S., & Lowry, C. (2021b). The Role of Realistic Channel Geometry Representation in Hydrological Model Predictions. *JAWRA Journal of the American Water Resources Association*, *57*(2), 222–240.

Brakenridge, G. R., Cohen, S., Kettner, A. J., De Groeve, T., Nghiem, S. V, Syvitski, J. P. M., & Fekete, B. M. (2012). Calibration of satellite measurements of river discharge using a global hydrology model. *Journal of Hydrology*, *475*, 123–136.

Brown, R. A., & Pasternack, G. B. (2014). Hydrologic and topographic variability modulate channel change in mountain rivers. *Journal of Hydrology*, *510*, 551–564.

Canova, M. G., Fulton, J. W., & Bjerklie, D. M. (2016). USGS HYDRoacoustic dataset in support of the Surface Water Oceanographic Topography satellite mission (HYDRoSWOT). *US Geological Survey [Data Set], Https://Doi. Org/10.5066/F7D798H6*.

Carlier, C., Wirth, S. B., Cochand, F., Hunkeler, D., & Brunner, P. (2018). Geology controls streamflow dynamics. *Journal of Hydrology*, *566*, 756–769.

Chan, J. Y.-L., Leow, S. M. H., Bea, K. T., Cheng, W. K., Phoong, S. W., Hong, Z.-W., & Chen, Y.-L. (2022). Mitigating the multicollinearity problem and its machine learning approach: a review. *Mathematics*, *10*(8), 1283.

Chaney, N. W., Minasny, B., Herman, J. D., Nauman, T. W., Brungard, C. W., Morgan, C. L. S., et al. (2019). POLARIS soil properties: 30-m probabilistic maps of soil properties over the contiguous United States. *Water Resources Research*, *55*(4), 2916–2938.

Cohen, S., Praskievicz, S., & Maidment, D. R. (2018). Featured collection introduction: National water model. *JAWRA Journal of the American Water Resources Association*.

Cohen, S., Raney, A., Munasinghe, D., Loftis, J. D., Molthan, A., Bell, J., et al. (2019). The Floodwater Depth Estimation Tool (FwDET v2. 0) for improved remote sensing analysis of coastal flooding. *Natural Hazards and Earth System Sciences*, *19*(9), 2053–2065.

Cruz-Cárdenas, G., López-Mata, L., Villaseñor, J. L., & Ortiz, E. (2014). Potential species distribution modeling and the use of principal component analysis as predictor variables. *Revista Mexicana de Biodiversidad*, *85*(1), 189–199.

DeCicco, L., Hirsch, R., Lorenz, D., Watkins, D., & Johnson, M. (2023). dataRetrieval: R packages for discovering and retrieving water data available from U.S. federal hydrologic web services. .

Dingman, S. L. (2009). *Fluvial hydraulics*. oxford university press.

Doyle, J. M., Hill, R. A., Leibowitz, S. G., & Ebersole, J. L. (2023). Random forest models to estimate bankfull and low flow channel widths and depths across the conterminous United States. *JAWRA Journal of the American Water Resources Association*.

Dunne, T., & Leopold, L. B. (1978). *Water in environmental planning*. Macmillan.

Durand, M., Rodriguez, E., Alsdorf, D. E., & Trigg, M. (2009). Estimating river depth from remote sensing swath interferometry measurements of river height, slope, and width. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *3*(1), 20–31.

Durand, M., Chen, C., de Moraes Frasson, R. P., Pavelsky, T. M., Williams, B., Yang, X., & Fore, A. (2020). How will radar layover impact SWOT measurements of water surface elevation and slope, and estimates of river discharge? *Remote Sensing of Environment*, *247*, 111883.

Emery, C. M., Biancamaria, S., Boone, A., Garambois, P.-A., Ricci, S., Rochoux, M. C., & Decharme, B. (2016). Temporal variance-based sensitivity analysis of the river-routing component of the large-scale hydrological model ISBA–TRIP: Application on the Amazon Basin. *Journal of Hydrometeorology*, *17*(12), 3007–3027.

Erikson, C. M., Renshaw, C. E., & Magilligan, F. J. (2024). Spatial variation in drainage area—Runoff relationships and implications for bankfull geometry scaling. *Geomorphology*, *446*, 108998.

Fang, S., Johnson, J. M., Yeghiazarian, L., & Sankarasubramanian, A. (2023). Improved National-Scale Above-normal Flow Prediction for Gauged and Ungauged Basins using a Spatio-temporal Hierarchical Model. *Water Resources Research*.

Frasson, R. P. de M., Pavelsky, T. M., Fonstad, M. A., Durand, M. T., Allen, G. H., Schumann, G., et al. (2019). Global relationships between river width, slope, catchment area, meander wavelength, sinuosity, and discharge. *Geophysical Research Letters*, *46*(6), 3252–3262.

Getirana, A. C. V, Boone, A., Yamazaki, D., & Mognard, N. (2013). Automatic parameterization of a flow routing scheme driven by radar altimetry data: Evaluation in the Amazon basin. *Water Resources Research*, *49*(1), 614–629.

Ghahremanloo, M., Lops, Y., Choi, Y., & Yeganeh, B. (2021). Deep learning estimation of daily ground-level NO2 concentrations from remote sensing data. *Journal of Geophysical Research: Atmospheres*, *126*(21), e2021JD034925.

Golly, A., & Turowski, J. M. (2017). Deriving principal channel metrics from bank and long-profile geometry with the R package cmgo. *Earth Surface Dynamics*, *5*(3), 557–570.

Han, M., Mai, J., Tolson, B. A., Craig, J. R., Gaborit, É., Liu, H., & Lee, K. (2020). Subwatershed-based lake and river routing products for hydrologic and land surface models applied over Canada. *Canadian Water Resources Journal/Revue Canadienne Des Ressources Hydriques*, *45*(3), 237–251.

Harman, C., Stewardson, M., & DeRose, R. (2008). Variability and uncertainty in reach bankfull hydraulic geometry. *Journal of Hydrology*, *351*(1–2), 13–25.

Heldmyer, A., Livneh, B., McCreight, J., Read, L., Kasprzyk, J., & Minear, T. (2022). Evaluation of a new observationally based channel parameterization for the National Water Model. *Hydrology and Earth System Sciences*, *26*(23), 6121–6136.

Hill, R. A., Weber, M. H., Leibowitz, S. G., Olsen, A. R., & Thornbrugh, D. J. (2016). The Stream-Catchment (StreamCat) Dataset: A database of watershed metrics for the conterminous United States. *JAWRA Journal of the American Water Resources Association*, *52*(1), 120–128.

Isikdogan, F., Bovik, A., & Passalacqua, P. (2017). RivaMap: An automated river analysis and mapping engine. *Remote Sensing of Environment*, *202*, 88–97.

Johnson, J. M. (2022). National Hydrologic Geospatial Fabric (hydrofabric) for the Next Generation (NextGen) Hydrologic Modeling Framework.

Johnson, J Michael, Munasinghe, D., Eyelade, D., & Cohen, S. (2019). An integrated evaluation of the national water model (NWM)–Height above nearest drainage (HAND) flood mapping methodology. *Natural Hazards and Earth System Sciences*, *19*(11), 2405–2420.

Johnson, J Michael, Coll, J., Clarke, K. C., Afshari, S., Saksena, S., & Yeghiazarian, L. (2022). Determining Feature Based Hydraulic Geometry and Rating Curves using a Physically Based, Computationally Efficient Framework. https://doi.org/https://doi.org/10.20944/preprints202212.0390.v1

Johnson J. Michael, Blodgett David L., Clarke Keith C., & Pollak Jon. Restructuring and serving web-accessible streamflow data from the NOAA National Water Model historic simulations, Scientific data § (2023).

Johnson, J Michael, Afshari, S., & Rad, A. M. (2024). AHGestimation: An R package for computing robust, mass preserving hydraulic geometries and rating curves. *Journal of Open Source Software*, *9*(96), 6145.

Johnson J.M. (2023). AHGEstimation: Tools for Estimating Physically-Based, Computationally Efficient Feature Based Hydraulic Geometry and Rating Curves. *JOSS*.

Johnson, J.M., Fang, S., Sankarasubramanian, A., Rad, A. M., Kindl da Cunha, L., Jennings, K., et al. (2023). Comprehensive analysis of the NOAA National Water Model: A call for heterogeneous formulations and diagnostic model selection. . *Journal of Geophysical Research - Atmospheres*.

Johnson, M. , & Clarke, K. C. (2019). climateR: An R package for getting point and gridded climate data by AOI.

Köppen, M. (2000). The curse of dimensionality. In *5th online world conference on soft computing in industrial applications (WSC5)* (Vol. 1, pp. 4–8).

Kouraev, A. V, Zakharova, E. A., Samain, O., Mognard, N. M., & Cazenave, A. (2004). Ob'river discharge from TOPEX/Poseidon satellite altimetry (1992–2002). *Remote Sensing of Environment*, *93*(1–2), 238–245.

Krawczyk, B. (2016). Learning from imbalanced data: open challenges and future directions. *Progress in Artificial Intelligence*, *5*(4), 221–232.

Lehner, B., & Grill, G. (2013). Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems. *Hydrological Processes*, *27*(15), 2171–2186.

Leopold, L. B., & Maddock, T. (1953). *The hydraulic geometry of stream channels and some physiographic implications* (Vol. 252). US Government Printing Office.

Lin, P., Pan, M., Allen, G. H., de Frasson, R. P., Zeng, Z., Yamazaki, D., & Wood, E. F. (2020). Global estimates of reach-level bankfull river width leveraging big data geospatial analysis. *Geophysical Research Letters*, *47*(7), e2019GL086405.

Liu, F., & Deng, Y. (2020). Determine the number of unknown targets in open world based on elbow method. *IEEE Transactions on Fuzzy Systems*, *29*(5), 986–995.

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

Luo, X.-L., Zeng, E. Y., Ji, R.-Y., & Wang, C.-P. (2007). Effects of in-channel sand excavation on the hydrology of the Pearl River Delta, China. *Journal of Hydrology*, *343*(3–4), 230–239.

McCabe, G. J., & Wolock, D. M. (2011). Independent effects of temperature and precipitation on modeled runoff in the conterminous United States. *Water Resources Research*, *47*(11).

McKay, L., Bondelid, T., Rea, A., Johnston, C., Moore, R., Dewald, T., et al. (2012). User Guide (Data Model Version 2.1).

Metzger, A., Marra, F., Smith, J. A., & Morin, E. (2020). Flood frequency estimation and uncertainty in arid/semi-arid regions. *Journal of Hydrology*, *590*, 125254.

Monegaglia, F., Zolezzi, G., Güneralp, I., Henshaw, A. J., & Tubino, M. (2018). Automated extraction of meandering river morphodynamics from multitemporal remotely sensed data. *Environmental Modelling & Software*, *105*, 171–186.

Moody, J. A., & Troutman, B. M. (2002). Characterization of the spatial variability of channel morphology. *Earth Surface Processes and Landforms: The Journal of the British Geomorphological Research Group*, *27*(12), 1251–1266.

Morel, M., Booker, D. J., Gob, F., & Lamouroux, N. (2020). Intercontinental predictions of river hydraulic geometry from catchment physical characteristics. *Journal of Hydrology*, *582*, 124292.

Neal, J., Hawker, L., Savage, J., Durand, M., Bates, P., & Sampson, C. (2021). Estimating river channel bathymetry in large scale flood inundation models. *Water Resources Research*, *57*(5), e2020WR028301.

Pavelsky, T. M., & Smith, L. C. (2008). RivWidth: A software tool for the calculation of river widths from remotely sensed imagery. *IEEE Geoscience and Remote Sensing Letters*, *5*(1), 70–73.

Pavelsky, T. M., Durand, M. T., Andreadis, K. M., Beighley, R. E., Paiva, R. C. D., Allen, G. H., & Miller, Z. F. (2014). Assessing the potential global extent of SWOT river discharge observations. *Journal of Hydrology*, *519*, 1516–1525.

Pekel, J.-F., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, *540*(7633), 418–422.

Rodell, M., Houser, P. R., Jambor, U. E. A., Gottschalck, J., Mitchell, K., Meng, C.-J., et al. (2004). The global land data assimilation system. *Bulletin of the American Meteorological Society*, *85*(3), 381–394.

Rosgen, D. L. (1994). A classification of natural rivers. *Catena*, *22*(3), 169–199.

Santhi, C., Allen, P. M., Muttiah, R. S., Arnold, J. G., & Tuppad, P. (2008). Regional estimation of base flow for the conterminous United States by hydrologic landscape regions. *Journal of Hydrology*, *351*(1–2), 139–153.

Schumann, G., Neal, J. C., Voisin, N., Andreadis, K. M., Pappenberger, F., Phanthuwongpakdee, N., et al. (2013). A first large-scale flood inundation forecasting model. *Water Resources Research*, *49*(10), 6248–6257.

Schwenk, J., Khandelwal, A., Fratkin, M., Kumar, V., & Foufoula-Georgiou, E. (2017). High spatiotemporal resolution of river planform dynamics from Landsat: The RivMAP toolbox and results from the Ucayali River. *Earth and Space Science*, *4*(2), 46–75.

Shen, C. (2018). A transdisciplinary review of deep learning research and its relevance for water resources scientists. *Water Resources Research*, *54*(11), 8558–8593.

Shen, C., Wang, S., & Liu, X. (2016). Geomorphological significance of at-many-stations hydraulic geometry. *Geophysical Research Letters*, *43*(8), 3762–3770.

Sichangi, A. W., Wang, L., & Hu, Z. (2018). Estimation of river discharge solely from remote-sensing derived data: an initial study over the Yangtze river. *Remote Sensing*, *10*(9), 1385.

Vidon, P. G. F., & Hill, A. R. (2004). Landscape controls on the hydrology of stream riparian zones. *Journal of Hydrology*, *292*(1–4), 210–228.

Wilkerson, G. V. (2008). Improved Bankfull Discharge Prediction Using 2-Year Recurrence-Period Discharge 1. *JAWRA Journal of the American Water Resources Association*, *44*(1), 243–257.

Wohl, E. E., & Wilcox, A. (2005). Channel geometry of mountain streams in New Zealand. *Journal of Hydrology*, *300*(1–4), 252–266.

Wolock, D. M., Winter, T. C., & McMahon, G. (2004). Delineation and evaluation of hydrologic-landscape regions in the United States using geographic information system tools and multivariate statistical analyses. *Environmental Management*, *34*, S71–S88.

Wood, E. F., Roundy, J. K., Troy, T. J., Van Beek, L. P. H., Bierkens, M. F. P., Blyth, E., et al. (2011). Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water. *Water Resources Research*, *47*(5).

Woodyer, K. D. (1968). Bankfull frequency in rivers. *Journal of Hydrology*, *6*(2), 114–142.

Yamazaki, D, Oki, T., & Kanae, S. (2009). Deriving a global river network map and its sub-grid topographic characteristics from a fine-resolution flow direction map. *Hydrology and Earth System Sciences*, *13*(11), 2241–2251.

Yamazaki, Dai, Ikeshima, D., Sosa, J., Bates, P. D., Allen, G. H., & Pavelsky, T. M. (2019). MERIT Hydro: A high-resolution global hydrography map based on latest topography dataset. *Water Resources Research*, *55*(6), 5053–5073.

Yang, X., Pavelsky, T. M., Allen, G. H., & Donchyts, G. (2019). RivWidthCloud: An automated Google Earth Engine algorithm for river width extraction from remotely sensed imagery. *IEEE Geoscience and Remote Sensing Letters*, *17*(2), 217–221.

Zakharova, E., Nielsen, K., Kamenev, G., & Kouraev, A. (2020). River discharge estimation from radar altimetry: Assessment of satellite performance, river scales and methods. *Journal of Hydrology*, *583*, 124561.

Zhang, H., Chen, L., & Singh, V. P. (2021). Flood frequency analysis using generalized distributions and entropy-based model selection method. *Journal of Hydrology*, *595*, 125610.

Zheng, X., Maidment, D. R., Tarboton, D. G., Liu, Y. Y., & Passalacqua, P. (2018). GeoFlood: Large-scale flood inundation mapping based on high-resolution terrain analysis. *Water Resources Research*, *54*(12), 10–13.