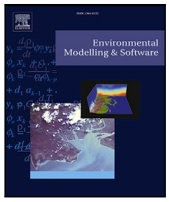
Environmental Modelling and Software 180 (2024) 106136

Contents lists available at ScienceDirect

Environmental Modelling and Software

journal homepage: www.elsevier.com/locate/envsoft



Position Paper

A large dataset of fluvial hydraulic and geometry attributes derived from USGS field measurement records

Seyed Mohammad Hassan Erfani [a,b,c,*,1], Mahdi Erfani [a,d,1], Sagy Cohen [e], Austin R.J. Downey [a,f], Erfan Goharian [a,d]

[a] Department of Civil & Environmental Engineering, University of South Carolina, Columbia, SC 29208, USA
[b] Center for Climate Systems Research, Columbia Climate School, Columbia University, New York, NY, USA
[c] NASA Goddard Institute for Space Studies, New York, NY, USA
[d] Scripps Institution of Oceanography, Center for Western Weather and Water Extremes, University of California San Diego, CA, USA
[e] Department of Geography, University of Alabama, Tuscaloosa, AL 35487, USA
[f] Department of Mechanical Engineering, University of South Carolina, Columbia, SC 29208, USA

ARTICLE INFO

Dataset link: https://github.com/smhassanerfani/ifmha

Keywords:
Hydrology
River channel geometry
Fluvial hydraulic
Large dataset

ABSTRACT

Accurate representation of river channel geometry is important for hydrologic and hydraulic modeling of fluvial systems. Often, channel geometry is estimated using simple rating curves that can be applied across various spatial scales. However, such methods are limited to power law relations that do not employ many potentially relevant catchment and river attributes. This paper introduce a new dataset, IFMHA (Inventory of Field Measurement of Hydraulic Attributes), to enable research studies on channel geometry and streamflow characteristics. IFMHA is derived from the National Water Information System (NWIS) site inventory for surface water field measurements and stream attributes from the National Hydrography Dataset (NHD). IFMHA includes 2,802,532 records from 10,050 sites (NWIS streamgaging stations). The dataset utility is demonstrated here by presenting a series of conceptual models for estimating channel geometry parameters (i.e., channel mean depth, channel maximum depth, wetted perimeter, and roughness) based on the available field attributes within IFMHA. Such a dataset and attributed channel geometry parameters can enhance the performance of operational flood forecasting frameworks (e.g. National Water Model) by providing more accurate initial conditions used in hydrologic and hydraulic routing models.

1. Introduction

Hydrological models are prominent tools for flood forecasting, water resources management, and the study of climate change impacts (Bárdossy and Singh, 2008; Mendoza et al., 2016; Krysanova et al., 2018). Hydrological models often require a routing method to estimate discharge hydrograph through a river channel (Yoo et al., 2017; Zhao et al., 2017; Nguyen-Quang et al., 2018). The accuracy of the representation of channel geometry can significantly impact routing simulations within hydrological models (Orlandini and Rosso, 1998; Hutton et al., 2012; Brackins et al., 2021). Since obtaining accurate channel geometry using surveying methods is costly, simplifying assumptions are often used (Brackins et al., 2021). An example of this approach is the representation of channel geometry in the National Water Model (NWM) developed by the National Oceanic and Atmospheric Administration (NOAA), National Weather Service. At present, the NWM is using a compound channel formulation that incorporates trapezoidal bankfull and rectangular floodplain cross sections as a means of simplifying the behavior of the flow (Gochis et al., 2020).

The importance of proper representation of channel geometry and the lack of surveyed data in most rivers has motivated research in the field of hydraulic geometry, mostly focusing on power-law equations that relate channel width, depth, and velocity with drainage area or discharge (Leopold and Maddock, 1953; Parker et al., 2007; Ferguson, 1986; Dingman, 2007; Stewardson, 2005). In locations where no or limited discharge data is available, drainage area is often used (Dunne and Leopold, 1978; Ames et al., 2009). It has been shown that regional calibration of hydraulic geometry equations provides a more robust estimation of channel parameters (Bieger et al., 2015, 2016). This, however, requires much larger and spatially diverse datasets. Attempting to overcome this challenge, a combined dataset from 50 studies was used to develop regional calibrations for estimation of bankfull hydraulic geometry (Blackburn-Lynch et al., 2017), which are currently

* Corresponding author at: Center for Climate Systems Research, Columbia Climate School, Columbia University, New York, NY, USA.
E-mail address: mohammad.erfani@columbia.edu (S.M.H. Erfani).
[1] Contributed equally.

https://doi.org/10.1016/j.envsoft.2024.106136
Received 17 October 2023; Received in revised form 20 April 2024; Accepted 28 June 2024
Available online 10 July 2024

being used in NWM to estimate reach-averaged geometry (Gochis et al., 2020). Moreover, taking advantage of satellite measurements, the Surface Water and Ocean Topography (SWOT) mission by NASA and CNES, focuses on providing several river data products such as river width and water surface elevation (Lettenmaier et al., 2015; Altenau et al., 2021). However, remote sensing products such as those offered by SWOT, are not without limitations as they cannot observe below the water surface and the small rivers are unobserved (Mersel et al., 2013; Nickles et al., 2020).

In recent years, the development of large datasets has opened new opportunities in various fields of water resources engineering. An example of this is the Catchment Attributes and Meteorology for Large-sample Studies (CAMELS) Dataset (Newman et al., 2015; Kratzert et al., 2023), allowing for benchmarking various types of hydrological models and deriving robust conclusions on hydrologic processes (Kratzert et al., 2019; Lees et al., 2021; Feng et al., 2022). Furthermore, development of such datasets has enabled researchers in the field to understand regional variability in catchment behavior, regional model performance, analyzing extreme events and model uncertainty (Lane et al., 2019; Sawicz et al., 2011; Gudmundsson et al., 2019; Westerberg et al., 2016). While the utility of such datasets has been shown, the number of publicly available large datasets is still limited (Coxon et al., 2020), underlining the importance of developing new and improved datasets. However, care must be taken when using large datasets as data inconsistencies, sometimes referred to as disinformative data, can lead to model calibration and inference issues (Beven, 2024; Kauffeldt et al., 2013; Clerc-Schwarzenbach et al., 2024).

In the context of river and channel geometry, a number of studies were published focusing on the development of relevant datasets including, the National Hydrography Dataset Plus (NHDPlusV2) (McKay et al., 2012), USGS HYDRoacoustic dataset in support of the Surface Water Oceanographic Topography satellite mission (HYDRoSWOT) (Canova et al., 2016; Bjerklie et al., 2020), and satellite-based datasets, providing opportunities for the development of data-driven approaches for estimation of river geometry (Allen and Pavelsky, 2015; Lin et al., 2020; Li et al., 2022). One recent application involved developing a new dataset, the hydraulic geometry (HyG), to better characterize the effect of channel parameters on NWM-simulated streamflow using a set of regression-based regionalizations (Heldmyer et al., 2022). The HyG dataset, which is currently unpublished, is a recent compilation of field discharge measurements gathered from gauges located throughout the continental United States.

The HYDRoSWOT dataset is a compilation of geometry and hydraulic observations from over 220,000 cross-sectional measurements obtained from 10,081 United States Geological Survey (USGS) streamgaging sites. This dataset is constrained to measurements recorded by Acoustic Doppler Current Profiler. HYDRoSWOT contains a variety of channel geometry parameters and flow characteristics such as mean depth, mean velocity, discharge, stage, water-surface width, maximum depth, and maximum velocity (Canova et al., 2016). However, by filtering zero, missing, suspicious values, and outliers, the number of records significantly diminished. For example, Bjerklie et al. (2020) imposed restrictions on HYDRoSWOT based on some hydraulic conditions which resulted in a subset including only 20,625 records, representing 3519 gaging stations. The low number of records in many locations, combined with the inherent uncertainties and inconsistencies of measurements recorded irregularly over long periods of time, underline the benefits of developing larger and improved datasets.

In this study, a new channel hydraulic attributes dataset is presented; the Inventory of Field Measurement of Hydraulic Attributes (IFMHA). IFMHA is based on the National Water Information System (NWIS) surface water field measurements and attributes from the National Hydrography Dataset Plus Version 2.1 (NHDPlusV2.1). It includes 2,802,532 records from 10,050 sites (USGS gaging stations) collected from the entire field measurements data for each site. Compared to HYDRoSWOT, IFMHA includes a much larger number of measurements as well as additional channel geometry data, which allows for a more robust analysis. IFMHA includes five important channel geometry and streamflow characteristics (i.e., flow rate, velocity, stage, channel width, and cross-section area). Moreover, we present a series of analyses based on IFMHA in which additional channel parameters (i.e., channel mean depth, maximum depth, wetted perimeter, and roughness) are estimated by conceptual models. The high predictability of these models allows us to not only estimate additional attributes but also fill gaps in the existing channel geometry attributes, stemming from missing values in the NWIS field measurements data across the US.

2. Dataset overview

IFMHA includes 2,802,532 field measurements collected at 10,050 USGS gage sites across the US. The time span of these measurements is from 05-05-1845 to 10-24-2022 (our most recent data retrieval date). Notably, IFMHA expands upon the HYDRoSWOT dataset by incorporating the entire record of field measurements within NWIS including different types of field measurement methods (see Table 1). The number of unique gage sites included in IFMHA is 31 sites less than HYDRoSWOT (17 of these stations are operated by an agency other than USGS, and the data for the other 14 stations are not available on the USGS website). IFMHA augmented the NWIS field measurements by adding relevant attributes reported in the NHDPlus V2.1 dataset.

As mentioned, IFMHA consists of three main sources of data including HYDRoSWOT, NWIS field measurements, and NHDPlus V2.1 datasets. The order and process of incorporating each dataset for the compilation of IFMHA are depicted in Fig. 1. The flowchart outlines the sequential steps involved in dataset adoption. The following paragraphs will explain each data sources mentioned in Fig. 1.

2.1. HYDRoSWOT

HYDRoSWOT is an extensive collection of USGS cross-section surveys for supporting Surface Water Oceanographic Topography (SWOT). The dataset comprises over 200,000 records of USGS Acoustic Doppler Current Profiler discharge measurements collected from the USGS streamgaging network. It offers a range of essential fields, including mean and maximum depth, velocity, discharge, stage, and water-surface width. Since USGS site numbers are common to both IFMHA and HYDRoSWOT datasets, constant stream and catchment features such as station names, geographic coordinates, drainage area, and steam types were extracted from HYDRoSWOT (see Table 2). For this purpose, first, the desired features were selected in HYDRoSWOT, then they were attributed to IFMHA using the USGS streamgage identifier (site number).

2.2. NWIS field measurements

The USGS conducts periodic field measurements including cross-section surveys for thousands of gage sites. Several methods are used by the USGS to conduct streamflow measurements, but the USGS principally uses mechanical current meters and hydroacoustic meters (Turnipseed and Sauer, 2010). Cross-section surveys at gaging stations generally include measurement of width, depth, velocity, the direction of flow, and recording of field notes. These observations are used to calculate discharge using the 'velocity-area' method (i.e., discharge is computed as the product of the area and velocity). The USGS provides public access to these measurements through the NWIS web interface under the category of surface water field measurement for each USGS site.

For the compilation of IFMHA, field measurements were queried in a tab-separated format for each NWIS station and then parsed and converted into a unified data frame comprised of records for all sites. For this purpose, URL links for each USGS site station were first created

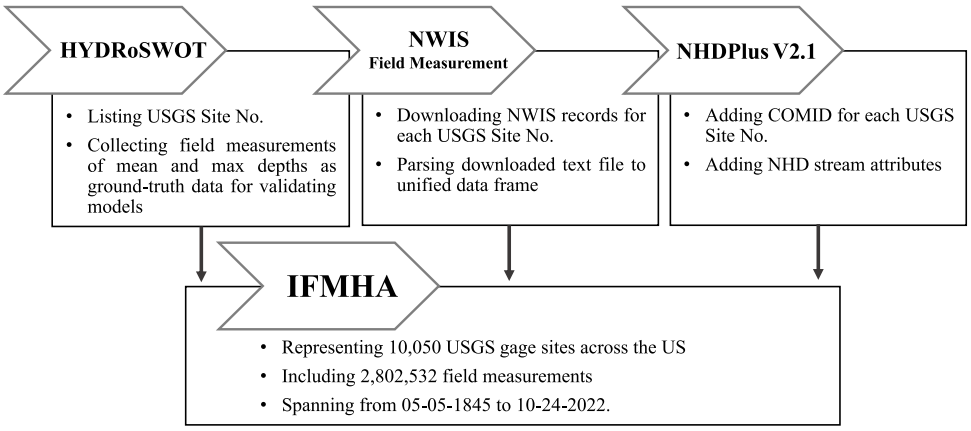


Fig. 1. Flowchart illustrates the incorporation of three key data sources, namely HYDRoSWOT, NWIS field measurements, and NHDPlus V2.1 datasets, for the compilation of IFMHA. The desired variables from each data source were linked to IFMHA using the USGS site number and common identifiers (COMID).

Table 1
The different types of USGS streamflow field measurement methods used in IFMHA.

Method	Description	Frequency	Percent
UNSPE	Unspecified	1 204 516	43.19%
QSCMM	Discharge, measured, midsection method	792 169	28.40%
QADCP	Discharge, measured, ADCP from moving boat	437 418	15.68%
OTHER		347 104	12.45%
QFLUM	Discharge, measured, flume	3888	0.14%
QIDIR	Discharge, measured, indirect method	2959	0.11%
QVOLM	Discharge, measured, volumetric	900	0.03%
NONE		4	0.00%
ACOUS	Acoustic doppler current profile	2	0.00%
ESTIM	Estimated	1	0.00%

Table 2
The HYDRoSWOT column names and definitions.

Field code	Description
site_no	Site number
station_nm	Station name
dec_lat_va	Decimal latitude (degrees)
dec_long_va	Decimal longitude (degrees)
drain_area_va	Drainage area (square mile)
contrib_drain_area_va	Contributing drainage area (square mile)
site_tp_cd	Site type code

using the list of site numbers in the HYDRoSWOT dataset. Then, field measurements for each site station were queried from the NWIS Web Interface using the URL links. Finally, the requested text file of each site was parsed into a unified data frame. The field codes and definitions of attributes provided in the NWIS web interface are shown in Table 3. The Python script for retrieving a list of active streamgages and their associated field measurements from the NWIS web service, as well as parsing the raw files into a unified data frame, can be found on the public IFMHA GitHub page.

2.3. NHDPlus V2.1 National Seamless Geodatabase

National Hydrography Dataset Plus (NHDPlus) is a national geospatial surface water framework including the features and capabilities of the National Hydrography Dataset (NHD), the National Elevation Dataset (NED), and the Watershed Boundary Dataset (WBD). NHDPlus integrates the vector NHD stream network and WBD hydrologic unit boundaries with the NED gridded land surface. In this study, the desired field attributes were extracted from the gage location (GageLoc) and NHD flow line network (NHDFlowline) geospatial databases and attributed to IFMHA using the USGS site number and the common identifiers (COMID) of the NHD stream features (see Table 4). NHDFlowline represents different types of streamflows (i.e., stream/river, canal/ditch, pipeline, artificial path, coastline, and connector) which are included in the NHDPlus surface water network, and GageLoc contains the locations of stream flow gages on the NHDFlowline features. For those USGS sites where there was no correspondence in the GageLoc geospatial database, NHDPlus reach attributes were assigned to the nearest site within a 300 meters buffer (McKay et al., 2012) using the Nearest Neighbor Analysis tool in QGIS.

3. Methodology

In addition to the aforementioned field attributes, IFMHA provides estimates for four channel geometry parameters (i.e., channel mean depth, maximum depth, wetted perimeter, and roughness) using conceptual models. The models adopted to estimate these channel parameters are explained in this section. The proposed models are evaluated using the mean and maximum depth values measured from observations of HYDRoSWOT. For this purpose, records with missing and zero values for the desired field attributes (i.e., site type, discharge, channel cross-section area, width, mean depth, and max depth) were excluded. Then, sites that include negative discharge observations are omitted as they, most likely, represent control sections and tidal areas. Finally, only sites categorized as channels or streams are selected, excluding other types such as canals, estuaries, and coastal. After these filtrations, the subset of HYDRoSWOT used in the following analyses includes 47,522 records within 5914 sites (see Table 5).

A similar filtering procedure was adopted for the IFMHA dataset. First, missing and zero values for the desired field attributes including site type, discharge, channel cross-section area, width, and slope were removed. Similar to the HYDRoSWOT subset, sites that include negative discharge observations or are categorized as site types other than natural streams were excluded. For meaningful statistical analysis, only sites with over 50 observations were selected. Eventually, 2,064,782 observations (represented by 6498 sites) remained for the analyses. The geographical distribution of sites and the corresponding number of records are shown in Fig. 2.

Table 3
The NWIS web interface tab-separated field column names and definitions.

Field code	Description
agency_cd	Agency code
site_no	Site number
measurement_nu	Measurement number
measurement_dt	Date of measurement (MM-DD-YYYY)
tz_cd	Time zone offset, an ANSI SQL/92 time zone offset string
q_meas_used_fg	Flag indicates if the discharge measurement is marked used
party_nm	An indication of who made the measurement
site_visit_coll_agency_cd	What agency made the measurement at the site visit
discharge_va	The computed discharge (cfs)
gage_height_va	Gage height as shown on the inside staff gage at the site (ft)
gage_va_change	The amount the gage height changed during the measurement (ft)
gage_va_time	The amount of time elapsed during the measurement (decimal hours)
measured_rating_diff	Measurement rating codes, the relative quality of the measurement
control_type_cd	Condition of the rating control at the time of the measurement
discharge_cd	The adjustment code for the measured discharge
chan_nu	The channel number
chan_name	The channel name
meas_type	The channel measurement type
streamflow_method	The channel discharge measurement method
velocity_method	The channel velocity measurement method
chan_discharge	The channel discharge (cfs)
chan_width	The channel width (ft)
chan_velocity	The mean velocity (ft/s)
chan_area	The channel area (ft^2)
chan_stability	The stability of the channel material
chan_material	The channel material
chan_evenness	The channel evenness from bank to bank
long_vel_desc	The longitudinal velocity description
horz_vel_desc	The horizontal velocity description
vert_vel_desc	The vertical velocity description
chan_loc_cd	The channel location code
chan_loc_dist	The channel location distance

Table 4
The NHDPlus V2 column names, definitions, and sources.

Field code	Description	Source
Source_Fea	Gage ID/USGS NWIS Site Number	GageLoc
FLComID	ComID of the NHDFlowline feature on which the gage is located	GageLoc
State_CD	State Code	GageLoc
State	State Abbreviation	GageLoc
DASqMi	NWIS Drainage Area (in square miles)	GageLoc
DASqKm	NWIS Drainage Area (in square kilometers)	GageLoc
LatSite	NWIS Latitude	GageLoc
LonSite	NWIS Longitude	GageLoc
ComID	Common identifier of the NHD feature	NHDFlowline
Ftype	NHD Feature Type	NHDFlowline
StreamOrde	Modified Strahler Stream Order	NHDFlowline
Slope	Slope of flowline (meters/meters) based on smoothed elevations	NHDFlowline

Table 5
The statistics for discharge and channel geometry parameters of HYDRoSWOT subset.

Statistics	Discharge (m^3/s)	Area (m^2)	Width (m)	Mean depth (m)	Max depth (m)
Count	47,522	47,522	47,522	47,522	47,522
Mean	311.50	318.76	72.91	2.04	3.00
STD	1671.69	1354.10	121.31	2.22	3.28
Min	0.0015	0.10	0.78	0.16	0.21
25%	6.77	18.12	21.10	0.82	1.25
50%	22.66	49.19	37.38	1.36	2.07
75%	88.58	147.31	74.43	2.39	3.48
Max	63,925.25	33,166.39	4987.12	36.20	62.69

3.1. Mean depth

In order to estimate the mean depth, the "Mean Value Theorem for Integrals" is applied. This theorem states that the average value of a continuous function $f(x)$ over a closed interval $[a, b]$ is equal to the average value of $f(x)$ over $[a, b]$.

$$f(c) = \frac{1}{(b-a)} \int_a^b f(x)\,dx \tag{1}$$

where, $f(c)$ is the average value of the function $f(x)$, and c is a point in an interval $[a, b]$. The schematic for the calculation of mean depth is illustrated in Fig. 3.

The performance of the proposed model was assessed using the HYDRoSWOT subset. Firstly, for each record, the mean depth value was estimated using the corresponding width and cross-section area. The results were then compared against the actual mean depth measurements. Eventually, the model was used to calculate the mean depth for IFMHA records. To do this, the values of the channel cross-section

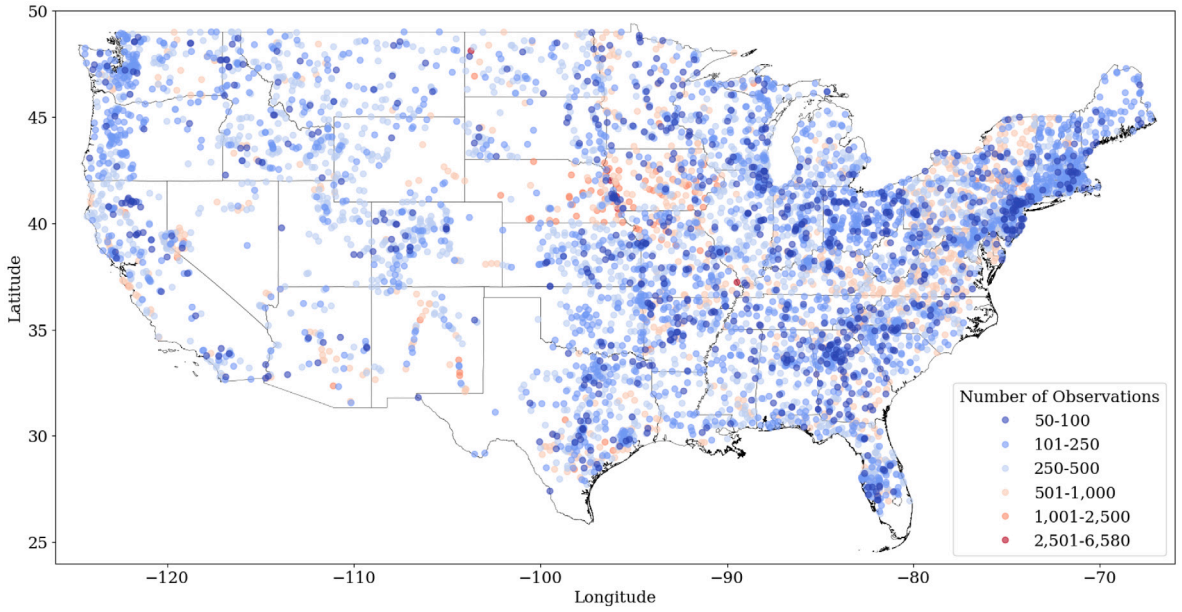
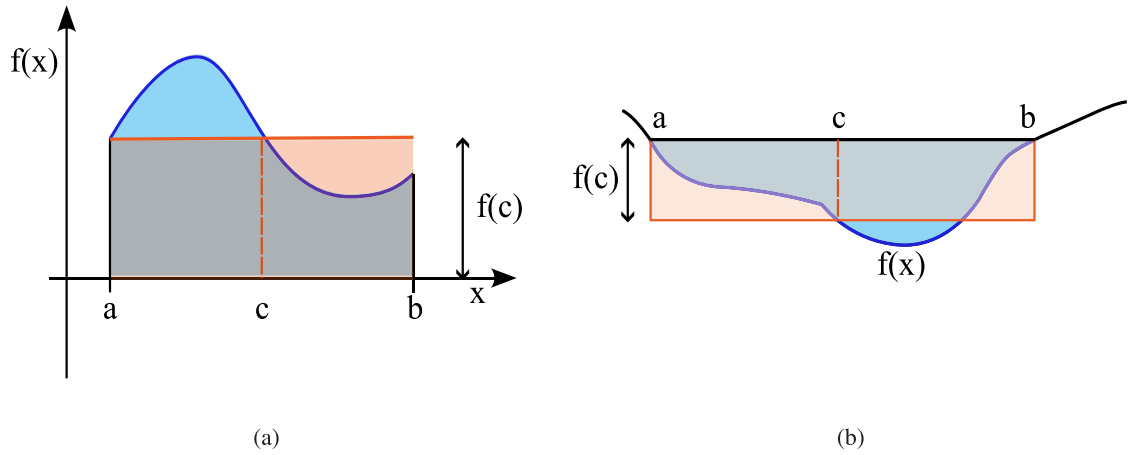


Fig. 2. Map of the distribution of streamgages across the continental United States.



(a) (b)

Fig. 3. Using "Mean Value Theorem for Integrals" to estimate mean depth, (a) is a mathematical illustration of the theorem, and (b) is equivalent parameters in channel cross-section geometry.

area were divided by their corresponding channel width. The results of mean depth are reported and discussed in Section 4.1.

3.2. Maximum depth (Thalweg)

From HYDRoSWOT, Bjerklie et al. (2020) found that among several cross-sectional hydraulic ratios, the ratio of maximum to mean depth offers the smallest coefficient of variation, indicating these measures are the most stable across the various flows and rivers. Bjerklie et al. (2020) adopted the line of organic correlation (LOC) regression to fit the best line. According to Kermack and Haldane (1950), LOC for the predicted values of the variable Y can be expressed as

$$y = \alpha x + \beta \tag{2}$$

where, α, and β are defined as

$$\alpha = \frac{\sigma_Y}{\sigma_X} \tag{3a}$$

$$\beta = \mu_Y - \frac{\sigma_Y}{\sigma_X}\mu_X \tag{3b}$$

The HYDRoSWOT subset was used to calibrate the slope, α, and y-intercept, β of Eq. (2) and evaluate the performance of LOC regression in estimating the max depth using mean depth.

3.3. Cross-section

For each site station, two different conceptual cross-section shapes (trapezoid, and parabola) were used to estimate channel geometry (Fig. 4).

The trapezoidal cross-section was chosen because, by adjusting the side slope and bottom width, it could also depict rectangular and triangular cross-sections. This means that while estimating the parameters of one type of cross-section, it could automatically transform into two other distinct cross-sectional shapes. For estimating the side slope (m), and bottom width (B_w), a half-channel conceptualization was adopted (see Fig. 4(a)). For a given site station, the field measurements of channel width and corresponding depth were used to fit a linear regression. The coefficient of this fit is the channel side slope, and the intercept can be considered half of the channel's bottom width. In order to avoid negative values for the channel side slope, a Non-Negative Linear

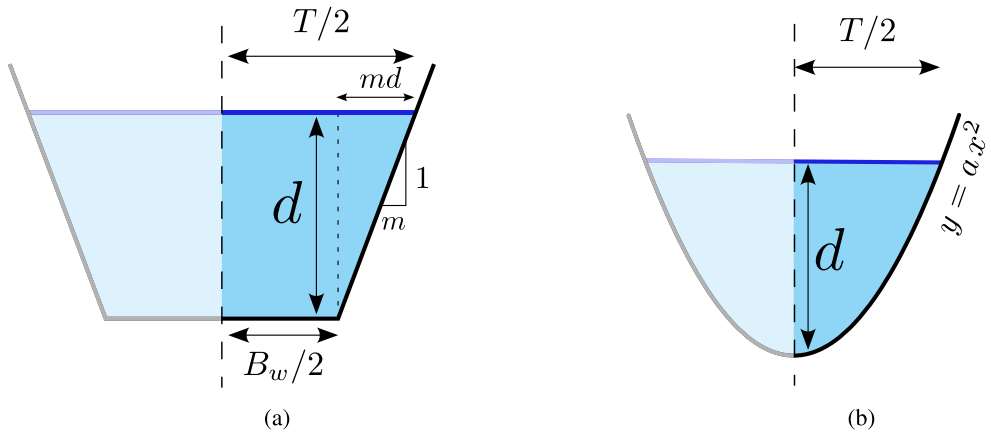


Fig. 4. Estimating parameters related to channel cross-section. A half-channel conceptualization of (a) a trapezoidal cross-section for estimating side slope (m) and bottom width (B_w), (b) a parabolic cross-section for estimating parabolic coefficient (a).

Regression (NNLR) method was used. As the estimation of the mean depth involved assuming an equivalent rectangular cross-section, the estimated maximum depth was incorporated in the following manner for the fitting process:

$$T/2 = m \times d_{max} + B_w/2 \tag{4}$$

where, T, and d_{max} are top width and maximum depth as well as m, and B_w are estimated side slope and bottom width, respectively.

The parabolic cross-section was selected based on the ratio of maximum-to-mean depth, which is investigated in Section 4.2. For a given site station, the field measurements of channel width and corresponding depth were used to fit a linear regression (with no intercept). Thus, the coefficient of this fit is the coefficient of the parabola defined by the equation $y = ax^2$ (Fig. 4(b)).

$$d_{max} = a \times (T/2)^2 \tag{5}$$

where, T, d_{max}, and a are top width, maximum depth, and parabolic coefficient, respectively.

3.4. Bankfull characteristics

In the estimation of bankfull characteristics, Heldmyer et al. (2022) proposes using the 99th percentile discharge for determining bankfull parameters. However, as per the definition provided by Parker (2006), for rivers with floodplains, bankfull is identified at the point of inflection in the curve fitting the river discharge versus stage (Fig. 5). This concept is rooted in the understanding that stage typically rises with increasing discharge while flow remains confined within the channel; however, this relationship shifts (decreasing) once the stream flow surpasses the bankfull state. Therefore, the inflection in the curve – a sudden change in slope of the fitting line – marks the stage at which flow exceeds bankfull conditions. While this study does not focus on the estimation of bankfull characteristics, the comprehensive data coverage provided by IFMHA allowed for testing and comparison of the two mentioned approaches across randomly selected site stations from the IFMHA dataset which is discussed in Section 4.4.

3.5. Roughness (Manning's n) coefficient

The determination of flow in an open channel requires an evaluation of the channel's resistance to flow, which is typically quantified by a roughness parameter, such as Manning's n coefficient. However, the characteristics of natural and constructed channels and the factors influencing channel roughness can exhibit considerable variability. In arid to semi-arid environments, Manning's equation parameters cannot be determined with sufficient accuracy by direct measurement of

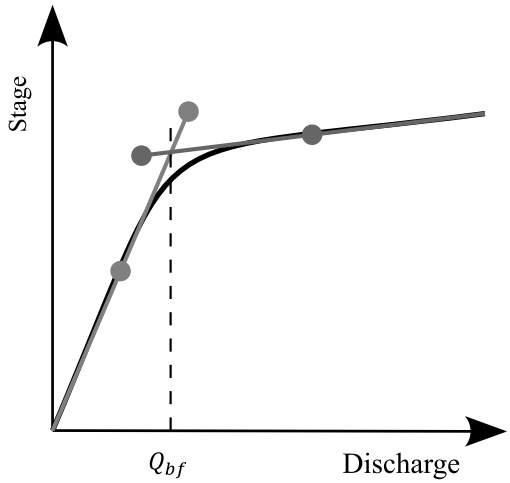


Fig. 5. In the case of rivers with floodplains, the stage tends to increase rapidly with increasing discharge when all the flow is confined to the channel, but much less rapidly when the flow spills significantly onto the floodplain. The rollover in the curve defines bankfull discharge (Q_{bf}) (Parker, 2006).

roughness characteristics, such as vegetation and variations in channel geometry. This is due to the dynamic nature of vegetation occupying the channel bed, banks, and overflow areas in intermittent and ephemeral streams, which can undergo significant changes over time or during flood events (Phillips and Tadayon, 2006). Field measurements of streamflow and the physical features of the channel can be used to back-calculate Manning's n, allowing for the development of empirical or physics-based models to estimate it indirectly.

The estimated channel geometry parameters (side slope and bottom width) were used to calculate the hydraulic radius, and then, using Manning's equation the roughness was calculated as follows,

$$n = \frac{A \times R^{2/3} \times S^{1/2}}{Q} \tag{6a}$$

$$R = \frac{A}{P} \tag{6b}$$

$$P = B_w + 2 \times d_{max} \times \sqrt{m^2 + 1} \tag{6c}$$

where, A, P, R, S, and Q represent cross-section area (m^2), wetted perimeter (m), hydraulic radius (m), channel gradient (m/m), and discharge (m^3/s), respectively.

In this study, the guidelines for the selection of roughness coefficients for natural and constructed channels were used to estimate

Table 6
The performance metrics of the proposed models for estimating Mean and Max depth of HYDRoSWOT subset.

Channel parameter	Number of data	Metrics			
		NSE	r	RMSE (m)	PBIAS
Mean depth	47,046	0.99	+0.99	0.17	−3.46
Max depth	47,046	0.94	+0.97	0.78	0.00

a maximum threshold for the roughness coefficients (Jarrett, 1985; Phillips and Tadayon, 2006). According to Cowan (1956), the selected base value of roughness n_0, can be adjusted by adding adjustment factors $n_1, \ldots, n_p$, and multiplying the meandering adjustment factor (m) as follows:

$$n = (n_0 + n_1 + n_2 + \cdots + n_p) \times m \tag{7}$$

This threshold was estimated for the most extreme conditions, i.e., the highest base n_0 value, and the highest adjustment factors for channel irregularity, variation in channel cross-section, obstructions, vegetation, and degree of meandering.

4. Results and discussion

The evaluation of the models is performed by comparing the estimated values with the measurements within a 99% confidence interval (i.e., data points falling outside the 0.5% lower and upper limits are considered outliers). This approach aims to mitigate the influence of outliers and measurement errors on the assessment of the models.

4.1. Mean depth

The performance metrics of the estimated mean depth are reported in Table 6. The results indicate a strong agreement between the estimated mean depth and ground truth data, with an NSE and r score of 0.99. These exceptional results indicate that further exploration of alternative models is unnecessary.

Fig. 6a visually compares the data distribution of ground truth (GT; HYDRoSWOT) with estimated (MODEL) mean depth using a box (whisker) plot. It shows that the model performs extremely well and can accurately capture the full range of values, including peaks. Fig. 6b shows the scatter plot that visually presents whether the model overestimates or underestimates the ground truth mean depths. It shows that almost all data points lie on the identity (1:1) line showing the agreement of estimated and measured values.

The mean depth value is missed for a significant part of the HydroSWOT dataset (about 75%). According to the performance metrics of the results, a good estimation of the mean depth could be made for the missing values using this approach.

4.2. Maximum depth

The validation performance of the LOC model using HydroSWOT is reported in Table 6. This model well performed on estimating maximum depth, with 0.94, and 0.97 for NSE and r scores, respectively. Fig. 7 shows the performance of the LOC model with ground truth data on the HYDRoSWOT subset. Fig. 7a shows model was able to cover a wide range of the upper bound values and outliers of data. This suggests that the model possesses the ability to estimate extreme values. Fig. 7b presents a scatter plot that visualizes the distribution of overestimated and underestimated values for the model. It is observed that the number of overestimated and underestimated values is roughly equal for the model.

4.3. Cross-section

As explained in Section 3.3, using a half-channel conceptualization a Non-Negative Linear Regression (NNLR) model was fit to estimate the side slope and bottom width. The obtained coefficient and intercept are representatives of the channel side slope and bottom width, respectively. For the side slope, since an NNLR model was used, no minimum threshold was necessary, but a maximum threshold of 12 (m/m) was set based on the maximum side slope used in NWM compound channel formulation (Gochis et al., 2020). In the case of channel bottom width, negative values were omitted. In total, the calculated parameters were within the chosen thresholds for 6169 sites (94.9% of all sites).

The fitted parameters of the LOC regression, including the slope (α) and intercept (β), are 1.48 and −0.01, respectively. By ignoring the intercept, the ratio of maximum-to-mean depth is close to the value of 1.50, implying that the cross sections tend toward a shape with a maximum-to-mean depth ratio consistent with a parabola (Bjerklie et al., 2020). As mentioned before, according to the 'Mean Value Theorem for Integrals', the mean depth is computed by dividing the area of the cross-section by the top width. Thus, for a parabolic cross-section, it is as follows:

$$A = \frac{2}{3} \times T \times d_{max} \tag{8a}$$

$$d_{mean} = \frac{A}{T} = \frac{2}{3} \times d_{max} \tag{8b}$$

where A is the area enclosed by the arc of the parabola (defined by the equation $y = ax^2$), equal to two-thirds of the maximum depth (d_{max}) multiplied by the top width (T); and d_{mean} represents the mean depth for the given parabola.

Fig. 8 compares the wetted perimeter estimated by trapezoidal and parabolic cross-sections for 12 randomly selected sites. The trapezoidal cross-section is a function of maximum depth, while the parabolic cross-section is estimated based on top width. In the cases of 05428500, 01591400, 10068500, 06486000, and 02481880, the results deviate from the diagonal line, indicating differences in estimating the wetted perimeter between these two approaches. In these instances, the values form a vertical line on the plot. This shape implies that the wetted perimeter is dominated by the bottom width of the trapezoidal cross-section, suggesting that these sites have wide channels where the bottom width has a more significant impact on the wetted perimeter compared to the side slopes. For cases 03415000 and 02298123, the wetted perimeter estimated by both cross-sections aligns more with the diagonal line. This suggests that the side slope in this trapezoidal cross-sections plays a more important role in determining the wetted perimeters.

4.4. Bankfull characteristics

According to Fig. 9, except for certain cases requiring special investigation (such as 05278930, 11482500, and 06307740), the 99th percentile discharge is significantly higher than the inflection point. For some locations (06670500, 01095505, 01166500, and 06610732), the median (50th percentile discharge) might better represent the inflection point on the hypothetical fitted line, while for others, it could still result in overestimation (13215000) or underestimation (05453100). This analysis implies that identifying bankfull characteristics given river discharge is not a straightforward task, as it depends on the nonlinear relationship between river discharge and stage. Additionally, this relationship is influenced by other factors such as the heterogeneity of watersheds, the stochastic nature of turbulent flow, and the characteristics of erosion processes, which contribute to the non-deterministic nature of estimating bankfull characteristics based on discharge. Given that the relationship between observed inputs and system response is stochastic, the modeling approach should also be stochastic (Shojaeezadeh et al., 2018). To employ a reliable probabilistic approach,

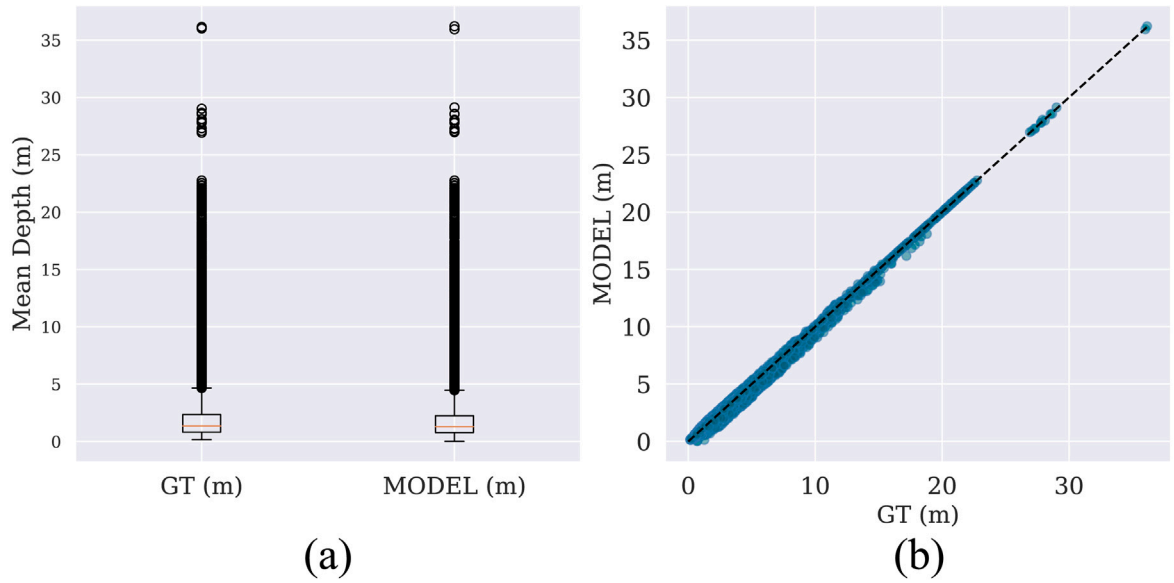


Fig. 6. Visual comparison of the measured (GT; HYDRoSWOT) with the estimated (MODEL) mean depth. (a) Visually compares the data distribution (the outliers showed by black circles represent peak values), and (b) shows the scatter plot.

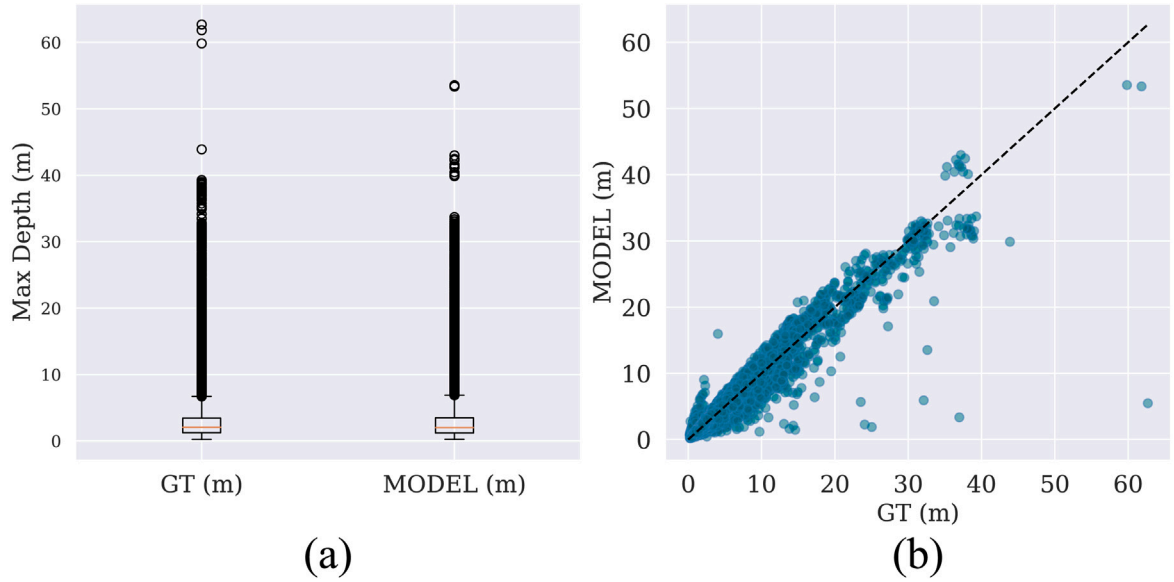


Fig. 7. The performance of LOC model on the HYDRoSWOT subset. (a) Visually compares the data distribution (the outliers showed by black circles represent peak values), and (b) shows the scatter plot.

it is essential to have a sufficient number of samples for statistical significance. By providing statistically significant data for each site, the IFMHA dataset enables stochastic approaches to estimate channel geometry parameters given flow discharge.

4.5. Roughness (Manning's n) coefficient

Various values of Manning's n may be assigned to a particular reach based on different flow conditions. However, the value assigned should reflect the combined impact of the factors that impede the flow (Phillips and Tadayon, 2006). Therefore, to allocate a specific roughness coefficient for each site, a process was undertaken in which the streamflow records within the 99% confidence intervals were identified (i.e., data points falling outside the 1% upper limit are considered outliers), and subsequently, the median of the calculated roughness coefficients was selected as the representative coefficient for that location.

Fig. 10 compares Manning's roughness coefficients estimated for the two different channel geometries (trapezoidal and parabolic). Manning's roughness coefficient was calculated based on the two different estimations of wetted perimeters for each of the 6169 sites in the previous section. Among these sites, 64 had roughness coefficients calculated using Manning's equation that exceeded the maximum Manning's n threshold (<0.4745) estimated by Eq. (7). Therefore, these 64 sites were excluded from the plots presented in Fig. 10.

Fig. 10a indicates that the values of Manning's roughness coefficient on the scatter plot align closely with the identity line (R^2 = 0.98). In this figure, the parabolic cross-section yields slightly higher values for Manning's roughness coefficient, as it generates lower wetted perimeter values compared to the trapezoidal cross-section, which maintains a constant bottom width. This result implies that although Manning's roughness coefficient depends on several factors related to the characteristics of the channel, including channel geometry, the effect of the channel cross-section should be considered relatively trivial. According

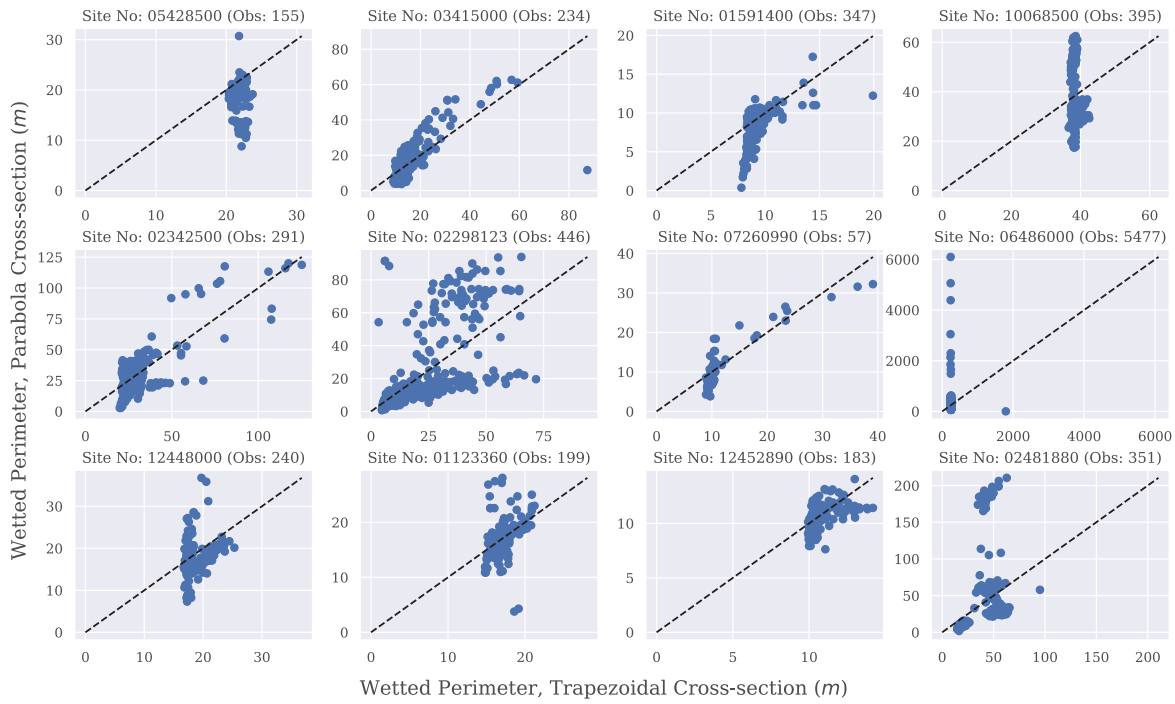


Fig. 8. The range of wetted perimeter values estimated by trapezoidal versus parabolic cross-sections for 12 USGS site stations. The USGS site number and the number of observations are reported as a title for each figure.

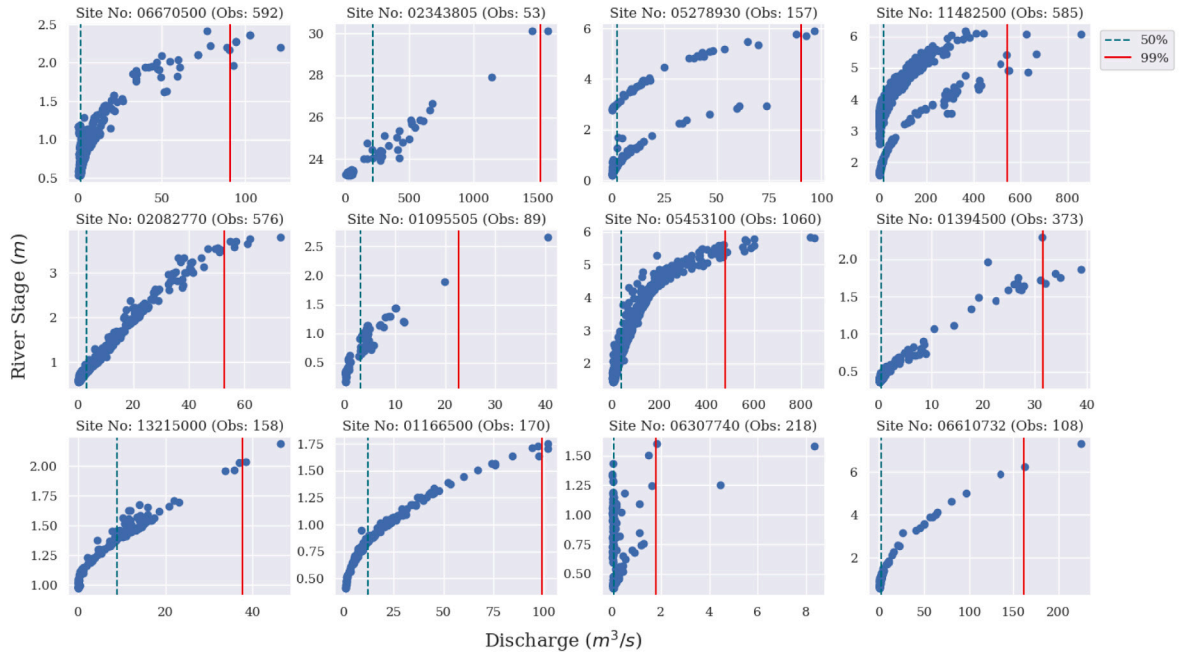


Fig. 9. The trend of increasing water stage with increasing discharge and the location of the 50th and 99th percentile discharge for twelve randomly selected USGS site stations from the IFMHA dataset.

to Fig. 10b, the distribution and other statistics of calculated values for Manning's roughness coefficients are similar for the two different channel geometries. This emphasizes that channel geometry does not have a significant impact on the values of Manning's roughness coefficient.

Moreover, while the results of this study show the shape of the channel cross-section may not significantly impact Manning's roughness coefficient, it remains a crucial parameter for accurately routing flow and estimating flood inundation areas.

5. Summary and conclusion

An accurate representation of channel geometry is vital for effectively routing flow and estimating flood inundation in hydrologic and hydrodynamic models. However, current approaches relying on rating curve methods or remote sensing measurements have limitations that compromise precision. Furthermore, regional calibration of hydraulic geometry equations necessitates extensive and diverse datasets, while the existing datasets are constrained by limited records and

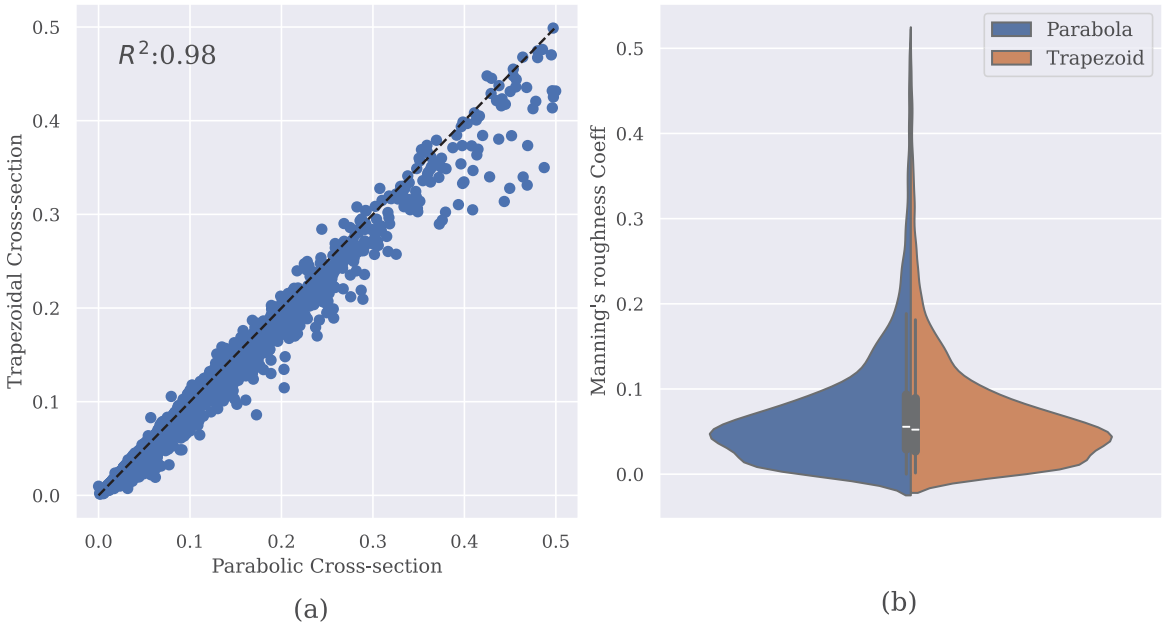


Fig. 10. Manning's roughness coefficients estimated by two different channel geometries: (a) scatter plot, (b) violin plot of the obtained results for 6169 sites.

missing values. To enhance the accuracy of flow routing, flood forecasting, and water resources management, it is imperative to integrate different available data sources for developing and enhancing hydrological models and comprehensive frameworks. Additionally, robust data-driven models should be further researched and developed to effectively harness large datasets, thereby improving the estimation of channel geometry and other hydrological processes.

This paper aims to advance data-driven discovery and modeling of fluvial dynamics by introducing IFMHA, a large-scale dataset for channel geometry and stream flow characteristics. It is built upon the HYDRoSWOT dataset and includes 2,802,532 records from 10,050 sites (NWIS gage stations) collected from NWIS site inventory for surface water field measurements and stream attributes of NHDPlusV2 dataset. After filtering out records with missing and zero values from both the HydroSWOT and IFMHA datasets, IFMHA retains 8431 records (with an average of 248 records per site), while HydroSWOT has 5914 records (with an average of 8 records per site), restoring access to 42% of the USGS site stations that were excluded in the HydroSWOT dataset during the filtering process. In IFMHA, even after excluding site stations with fewer than 50 observations, 6498 sites remained (with an average of 317 records per site), whereas this number diminishes to just 26 sites (with an average of 87 records per site) in HydroSWOT. The IFMHA dataset is invaluable for estimating river geometry across the Contiguous United States using stochastic approaches and ensuring a sufficient number of samples for achieving statistical significance.

Furthermore, a series of analyses for several channel geometry parameters (including channel mean depth, maximum depth, wetted perimeter, and roughness) was conducted using conceptual models to demonstrate the potential of utilizing large datasets. Results were compared with the channel parameters measured and reported in HYDRoSWOT dataset. In general, IFMHA provides hydraulic geometry data for 10,050 streamgaging stations across the US. Hydraulic geometry data can be used to explain the dynamics of river cross-section characteristics, which directly effect flood routing and flood inundation mapping. In addition, IFMHA encompasses a broad range of field attributes, such as channel material, channel type, and contributing drainage area. Due to its extensive coverage of hydraulic and hydrologic attributes, IFMHA offers opportunities for estimating a myriad of fluvial attributes, which could provide novel insights for the water resources community.

CRediT authorship contribution statement

Seyed Mohammad Hassan Erfani: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. **Mahdi Erfani:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. **Sagy Cohen:** Funding acquisition, Methodology, Supervision, Validation, Writing – original draft. **Austin R.J. Downey:** Resources. **Erfan Goharian:** Resources, Supervision, Writing – original draft.

Declaration of competing interest

1. Financial Conflicts: I declare that I (or my co-authors) have no financial interests, either directly or indirectly, in any organization, company, or entity that could potentially gain or lose financially from the publication of this work. There are no sources of financial support or sponsorship that have influenced the research presented in the manuscript.

2. Personal or Academic Conflicts: I confirm that there are no personal, academic, or professional relationships or conflicts that could influence the impartiality or objectivity of the review and editorial process. This includes any associations with individuals who are part of the editorial board or who may be reviewing the paper.

3. Intellectual Property and Patents: I attest that there are no patents or intellectual property rights related to the content of the manuscript that could be affected by its publication. There are no patent applications or intellectual property issues to disclose.

4. Non-Financial Conflicts: I acknowledge that there are no non-financial conflicts of interest, such as personal or ideological beliefs, that might influence the paper's content or the review process.

Data availability statement

IFMHA dataset and the source codes used in this study are publicly available in the GitHub repository https://github.com/smhassanerfani/ifmha.

Acknowledgments

This research was funded, in part, by the National Oceanic and Atmospheric Administration (NOAA), United States through the Cooperative Institute for Research to Operations in Hydrology (CIROH), FIM award number A22-0305. This study stems from a project during the 2022 NOAA National Water Center Summer Institute. We would like to thank the Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUASHI) and the University of Alabama which organized and hosted the 2022 Summer Institute.

References

Allen, George H., Pavelsky, Tamlin M., 2015. Patterns of river width and surface area revealed by the satellite-derived North American river width data set. Geophys. Res. Lett. 42 (2), 395–402.

Altenau, Elizabeth H, Pavelsky, Tamlin M, Durand, Michael T, Yang, Xiao, Frasson, Renato Prata de Moraes, Bendezu, Liam, 2021. The surface water and ocean topography (SWOT) mission river database (SWORD): A global river network for satellite data products. Water Resour. Res. 57 (7), e2021WR030054.

Ames, Daniel P, Rafn, Eric B, Van Kirk, Robert, Crosby, Benjamin, 2009. Estimation of stream channel geometry in idaho using GIS-derived watershed characteristics. Environ. Model. Softw. 24 (3), 444–448.

Bárdossy, A., Singh, S.K., 2008. Robust estimation of hydrological model parameters. Hydrol. Earth Syst. Sci. 12 (6), 1273–1283.

Beven, Keith, 2024. A brief history of information and disinformation in hydrological data and the impact on the evaluation of hydrological models. Hydrol. Sci. J. (just-accepted).

Bieger, Katrin, Rathjens, Hendrik, Allen, Peter M, Arnold, Jeffrey G, 2015. Development and evaluation of bankfull hydraulic geometry relationships for the physiographic regions of the United States. JAWRA J. Am. Water Resour. Assoc. 51 (3), 842–858.

Bieger, Katrin, Rathjens, Hendrik, Arnold, Jeffrey G, Chaubey, Indrajeet, Allen, Peter M, 2016. Development and comparison of multiple regression models to predict bankfull channel dimensions for use in hydrologic models. JAWRA J. Am. Water Resour. Assoc. 52 (6), 1385–1400.

Bjerklie, David M, Fulton, John W, Dingman, S Lawrence, Canova, Michael G, Minear, Justin T, Moramarco, Tommaso, 2020. Fundamental hydraulics of cross sections in natural rivers: Preliminary analysis of a large data set of acoustic Doppler flow measurements. Water Resour. Res. 56 (3), e2019WR025986.

Blackburn-Lynch, Whitney, Agouridis, Carmen T, Barton, Christopher D, 2017. Development of regional curves for hydrologic landscape regions (HLR) in the contiguous United States. JAWRA J. Am. Water Resour. Assoc. 53 (4), 903–928.

Brackins, John, Moragoda, Nishani, Rahman, Azbina, Cohen, Sagy, Lowry, Christopher, 2021. The role of realistic channel geometry representation in hydrological model predictions. JAWRA J. Am. Water Resour. Assoc. 57 (2), 222–240.

Canova, M.G., Fulton, J.W., Bjerklie, D.M., 2016. USGS Hydroacoustic Dataset in Support of the Surface Water Oceanographic Topography Satellite Mission. hydroswot, U.S. Geological Survey, URL https://doi.org/10.5066/F7D798H6.

Clerc-Schwarzenbach, Franziska Maria, Selleri, Giovanni, Neri, Mattia, Toth, Elena, van Meerveld, Ilja, Seibert, Jan, 2024. HESS Opinions: A Few Camels or a Whole Caravan? Vol. 2024, EGUsphere, pp. 1–29, [preprint].

Cowan, Woody L., 1956. Estimating hydraulic roughness coefficients. Agric. Eng. 37 (7), 473–475.

Coxon, Gemma, Addor, Nans, Bloomfield, John P, Freer, Jim, Fry, Matt, Hannaford, Jamie, Howden, Nicholas JK, Lane, Rosanna, Lewis, Melinda, Robinson, Emma L, et al., 2020. CAMELS-GB: hydrometeorological time series and landscape attributes for 671 catchments in great Britain. Earth Syst. Sci. Data 12 (4), 2459–2483.

Dingman, S. Lawrence, 2007. Analytical derivation of at-a-station hydraulic–geometry relations. J. Hydrol. 334 (1–2), 17–27.

Dunne, Thomas, Leopold, Luna B., 1978. Water in Environmental Planning. Macmillan.

Feng, Dapeng, Liu, Jiangtao, Lawson, Kathryn, Shen, Chaopeng, 2022. Differentiable, learnable, regionalized process-based models with multiphysical outputs can approach state-of-the-art hydrologic prediction accuracy. Water Resour. Res. 58 (10), e2022WR032404.

Ferguson, Robert I., 1986. Hydraulics and hydraulic geometry. Prog. Phys. Geogr. 10 (1), 1–31.

Gochis, DJ, Barlage, M, Cabell, R, Casali, M, A, Dugger, FitzGerald, K, McAllister, M, McCreight, J, RafieeiNasab, A, Read, L, Sampson, K, Yates, D, Zhang, Y, 2020. The WRF-Hydro Modeling System Technical Description, (Version 5.1.1). NCAR Technical Note, p. 107, https://ral.ucar.edu/sites/default/files/public/projects/wrf_hydro/technical-description-user-guide/wrf-hydro-v5.1.1-technical-description.pdf.

Gudmundsson, Lukas, Leonard, Michael, Do, Hong X, Westra, Seth, Seneviratne, Sonia I, 2019. Observed trends in global indicators of mean and extreme streamflow. Geophys. Res. Lett. 46 (2), 756–766.

Heldmyer, Aaron, Livneh, Ben, McCreight, James, Read, Laura, Kasprzyk, Joseph, Minear, Toby, 2022. Evaluation of a new observationally based channel parameterization for the national water model. Hydrol. Earth Syst. Sci. 26 (23), 6121–6136.

Hutton, Christopher J, Brazier, Richard E, Nicholas, Andrew P, Nearing, Mark, 2012. On the effects of improved cross-section representation in one-dimensional flow routing models applied to ephemeral rivers. Water Resour. Res. 48 (4).

Jarrett, Robert D., 1985. Determination of Roughness Coefficients for Streams in Colorado. Water-Resources Investigations Report 85-4004, US Department of the Interior, Geological Survey.

Kauffeldt, Anna, Halldin, Sven, Rodhe, Allan, Xu, C-Y, Westerberg, Ida K, 2013. Disinformative data in large-scale hydrological modelling. Hydrol. Earth Syst. Sci. 17 (7), 2845–2857.

Kermack, K.A., Haldane, John B.S., 1950. Organic correlation and allometry. Biometrika 37 (1/2), 30–41.

Kratzert, Frederik, Klotz, Daniel, Herrnegger, Mathew, Sampson, Alden K, Hochreiter, Sepp, Nearing, Grey S, 2019. Toward improved predictions in ungauged basins: Exploiting the power of machine learning. Water Resour. Res. 55 (12), 11344–11354.

Kratzert, Frederik, Nearing, Grey, Addor, Nans, Erickson, Tyler, Gauch, Martin, Gilon, Oren, Gudmundsson, Lukas, Hassidim, Avinatan, Klotz, Daniel, Nevo, Sella, et al., 2023. Caravan-a global community dataset for large-sample hydrology. Sci. Data 10 (1), 61.

Krysanova, Valentina, Donnelly, Chantal, Gelfan, Alexander, Gerten, Dieter, Arheimer, Berit, Hattermann, Fred, Kundzewicz, Zbigniew W, 2018. How the performance of hydrological models relates to credibility of projections under climate change. Hydrol. Sci. J. 63 (5), 696–720.

Lane, Rosanna A, Coxon, Gemma, Freer, Jim E, Wagener, Thorsten, Johnes, Penny J, Bloomfield, John P, Greene, Sheila, Macleod, Christopher JA, Reaney, Sim M, 2019. Benchmarking the predictive capability of hydrological models for river flow and flood peak predictions across over 1000 catchments in great Britain. Hydrol. Earth Syst. Sci. 23 (10), 4011–4032.

Lees, Thomas, Buechel, Marcus, Anderson, Bailey, Slater, Louise, Reece, Steven, Coxon, Gemma, Dadson, Simon J, 2021. Benchmarking data-driven rainfall–runoff models in great britain: a comparison of long short-term memory (LSTM)-based models with four lumped conceptual models. Hydrol. Earth Syst. Sci. 25 (10), 5517–5534.

Leopold, Luna Bergere, Maddock, Thomas, 1953. The Hydraulic Geometry of Stream Channels and Some Physiographic Implications. Vol. 252, US Government Printing Office.

Lettenmaier, Dennis P, Alsdorf, Doug, Dozier, Jeff, Huffman, George J, Pan, Ming, Wood, Eric F, 2015. Inroads of remote sensing into hydrologic science during the WRR era. Water Resour. Res. 51 (9), 7309–7342.

Li, Dan, Xue, Yuan, Qin, Chao, Wu, Baosheng, Chen, Bowei, Wang, Ge, 2022. A bankfull geometry dataset for major exorheic rivers on the Qinghai-Tibet plateau. Sci. Data 9 (1), 498.

Lin, Peirong, Pan, Ming, Allen, George H, de Frasson, Renato Prata, Zeng, Zhenzhong, Yamazaki, Dai, Wood, Eric F, 2020. Global estimates of reach-level bankfull river width leveraging big data geospatial analysis. Geophys. Res. Lett. 47 (7), e2019GL086405.

McKay, Lucinda, Bondelid, Timothy, Dewald, Tommy, Johnston, J, Moore, Richard, Rea, Alan, 2012. NHDPlus Version 2: User Guide. Vol. 745, US Environmental Protection Agency.

Mendoza, Pablo A, Clark, Martyn P, Mizukami, Naoki, Gutmann, Ethan D, Arnold, Jeffrey R, Brekke, Levi D, Rajagopalan, Balaji, 2016. How do hydrologic modeling decisions affect the portrayal of climate change impacts? Hydrol. Process. 30 (7), 1071–1095.

Mersel, Matthew K, Smith, Laurence C, Andreadis, Konstantinos M, Durand, Michael T, 2013. Estimation of river depth from remotely sensed hydraulic relationships. Water Resour. Res. 49 (6), 3165–3179.

Newman, AJ, Clark, MP, Sampson, Kevin, Wood, Andrew, Hay, LE, Bock, A, Viger, RJ, Blodgett, D, Brekke, L, Arnold, JR, et al., 2015. Development of a large-sample watershed-scale hydrometeorological data set for the contiguous USA: data set characteristics and assessment of regional variability in hydrologic model performance. Hydrol. Earth Syst. Sci. 19 (1), 209–223.

Nguyen-Quang, Trung, Polcher, Jan, Ducharne, Agnès, Arsouze, Thomas, Zhou, Xudong, Schneider, Ana, Fita, Lluís, 2018. ORCHIDEE-ROUTING: revising the river routing scheme using a high-resolution hydrological database. Geosci. Model Dev. 11 (12), 4965–4985.

Nickles, Cassandra, Beighley, Edward, Durand, Michael, Prata de Moraes Frasson, Renato, 2020. Integrating lateral inflows into a SWOT mission river discharge algorithm. Water Resour. Res. 56 (10), e2019WR026589.

Orlandini, Stefano, Rosso, Renzo, 1998. Parameterization of stream channel geometry in the distributed modeling of catchment dynamics. Water Resour. Res. 34 (8), 1971–1985.

Parker, Gary, 2006. 1D sediment transport morphodynamics with applications to rivers and turbidity currents. URL http://hydrolab.illinois.edu/people/parkerg/morphodynamics_e-book.htm.

Parker, Gary, Wilcock, Peter R, Paola, Chris, Dietrich, William E, Pitlick, John, 2007. Physical basis for quasi-universal relations describing bankfull hydraulic geometry of single-thread gravel bed rivers. J. Geophys. Res.: Earth Surf. 112 (F4).

Phillips, Jeff V., Tadayon, Saeid, 2006. Selection of Manning's Roughness Coefficient for Natural and Constructed Vegetated and Non-Vegetated Channels, and Vegetation Maintenance Plan Guidelines for Vegetated Channels in Central Arizona. US Department of the Interior, Geological Survey, p. 49. http://dx.doi.org/10.3133/sir20065108, - edition. ISBN: 2006-5108. URL http://pubs.er.usgs.gov/publication/sir20065108.

Sawicz, Keith, Wagener, Thorsten, Sivapalan, Murugesu, Troch, Peter A, Carrillo, Gigi, 2011. Catchment classification: empirical analysis of hydrologic similarity based on catchment function in the eastern USA. Hydrol. Earth Syst. Sci. 15 (9), 2895–2911.

Shojaeezadeh, Shahab Aldin, Nikoo, Mohammad Reza, McNamara, James P, AghaKouchak, Amir, Sadegh, Mojtaba, 2018. Stochastic modeling of suspended sediment load in alluvial rivers. Adv. Water Resour. 119, 188–196.

Stewardson, Michael, 2005. Hydraulic geometry of stream reaches. J. Hydrol. 306 (1–4), 97–111.

Turnipseed, D. Phil, Sauer, Vernon B., 2010. Discharge Measurements at Gaging Stations. Technical Report, U.S. Geological Survey, Reston, VA, http://dx.doi.org/10.3133/tm3A8, URL http://pubs.er.usgs.gov/publication/tm3A8.

Westerberg, Ida K, Wagener, Thorsten, Coxon, Gemma, McMillan, Hilary K, Castellarin, Attilio, Montanari, Alberto, Freer, Jim, 2016. Uncertainty in hydrological signatures for gauged and ungauged catchments. Water Resour. Res. 52 (3), 1847–1865.

Yoo, Chulsang, Lee, Jinwook, Lee, Myungseob, 2017. Parameter estimation of the muskingum channel flood-routing model in ungauged channel reaches. J. Hydrol. Eng. 22 (7), 05017005.

Zhao, Fang, Veldkamp, Ted IE, Frieler, Katja, Schewe, Jacob, Ostberg, Sebastian, Willner, Sven, Schauberger, Bernhard, Gosling, Simon N, Schmied, Hannes Müller, Portmann, Felix T, et al., 2017. The critical role of the routing scheme in simulating peak river discharge in global hydrological models. Environ. Res. Lett. 12 (7), 075003.