Continental Scale Assessment of Variation in Floodplain Roughness with Vegetation and Flow Characteristics

Gabriel Barinas[1], Stephen Paul Good[1], and Desiree Tullos[1]

[1]Oregon State University

December 14, 2023

Abstract

Quantifying floodplain flows is critical to multiple river management objectives, yet how vegetation within floodplains dissipates flow energy lacks comprehensive characterization. Utilizing over 3.4 million discharge measurements, in conjunction with aboveground biomass and canopy height measurements from NASA's Global Ecosystem Dynamics Investigation (GEDI), this study characterizes the floodplain roughness coefficient Manning's n and its determinates across the continental United States. Estimated values of n show that flow resistance in floodplains decreases as flow velocity increases but increases with the fraction of vegetation inundated. A new function (RMSE = 0.024, r^2 = 0.74) is proposed for predicting n based on GEDI vegetation characteristics and flow velocity, with GEDI derived n values improving predictions of discharge relative to those based only on land cover. This analysis provides evidence of key hydraulic patterns of energy dissipation in floodplains, and integration of the proposed function into flood and habitat models may reduce uncertainty.

Continental Scale Assessment of Variation in Floodplain Roughness with Vegetation and Flow Characteristics

Gabriel Barinas[1, 2, *], Stephen P. Good[1, 2], and Desiree Tullos[1, 2]

[1]Water Resources Graduate Program, Oregon State University

[2]Department of Biological & Ecological Engineering, Oregon State University

* Corresponding author: Gabriel Barinas (barinasg@oregonstate.edu)

Key Points:

- 4,927 estimates of floodplain roughness were calculated using flow observations and compared to LiDAR vegetation data.
- Floodplain roughness increases with increasing biomass and inundation depths and decreases with increasing flow velocity.
- Our model's Manning's n estimates yield lower errors in reach-scale floodplain flow predictions than n based solely on land cover.

Abstract

Quantifying floodplain flows is critical to multiple river management objectives, yet how vegetation within floodplains dissipates flow energy lacks comprehensive characterization. Utilizing over 3.4 million discharge measurements, in conjunction with aboveground biomass and canopy height measurements from NASA's Global Ecosystem Dynamics Investigation (GEDI), this study characterizes the floodplain roughness coefficient Manning's n and its determinates across the continental United States. Estimated values of n show that flow resistance in floodplains decreases as flow velocity increases but increases with the fraction of vegetation inundated. A new function (RMSE = 0.024, r^2 = 0.74) is proposed for predicting n based on GEDI vegetation characteristics and flow velocity, with GEDI derived n values improving predictions of discharge relative to those based only on land cover. This analysis provides evidence of key hydraulic patterns of energy dissipation in floodplains, and integration of the proposed function into flood and habitat models may reduce uncertainty.

Plain Language Summary

Quantifying the capacity of floodplains to dissipate energy from flowing water is important in managing rivers, restoring habitats, and reducing flood risks. By integrating overbank flood characteristics measured at USGS gauging stations with vegetation properties of floodplains measured by NASA, this study analyzed how energy dissipation in the floodplain, via a hydraulic roughness coefficient, varies with vegetation biomass and flood depths. Results indicate that floodplain roughness increases with the density of vegetation and decreases with flow velocity. A new mathematical function is presented to estimate floodplain roughness based on remotely sensed vegetation properties for various velocities.

1 Introduction

Floods are one of the most damaging natural disasters affecting society, costing billions of dollars in damages every year (Smith, 2020). Understanding these events is important for the protection of urban and agricultural development, risk management, and ecosystem restoration actions (Bulti & Abebe, 2020). Accordingly, a wide variety of hydraulic models have been developed for prediction and forecasting of river response to flood events and restoration actions, with the vast majority of these model predictions dependent on how a floodplain roughness attenuates flow (Hunter et al., 2007). Manning's equation (Manning, 1891) is the most widely used hydraulic formula relating roughness to discharge and velocities in river channels and floodplains (Yen, 1992). Its application requires knowledge of the geometric characteristics of the channel (area, hydraulic radius, and slope) as well as a key roughness coefficient, n. This empirical coefficient is used to account for energy dissipated due to friction losses, but it is rarely measured directly in the field (R. Ferguson, 2013) due to logistics and safety concerns, and it is difficult to predict for a future land use policy or engineering design. As a result, Manning's n is typically specified from simplified lookup tables (Chow, 1959; Cowan, 1956), and studies have demonstrated that uncertainties in n can lead to large errors in depth and discharge estimates (Durand et al., 2016; Lee & Mays, 1986).

Manning's equation in irrigation canals (Manning, 1891) has traditionally attributed energy losses in open channels primarily to vegetation. Lookup tables, such as those by (Chow, 1959), include specific n values for different land cover types, indicating the influence of vegetation on Manning's n. While most studies focus on flow resistance of vegetation in the main channel, limited attention has been given to variations in floodplain vegetation resistance during inundation events (R. Ferguson, 2013; Yen, 2002). Prior models (Fathi-Maghadam & Kouwen, 1997; Kouwen & Fathi-Moghadam, 2000; Petryk, 1975) of flow resistance for emergent vegetation, highlighted vegetation density as the most important factor contributing to Manning's *n*, and suggest *n* varies with the square root of the vegetation inundation fraction and inversely with flow velocity. However, these models were developed spanning limited conditions, e.g. only four individual trees of different types tested in (Kouwen & Fathi-Moghadam, 2000), and remain difficult to parametrize in practice. Furthermore, human modifications to floodplains, including the replacement of vegetation with agricultural fields, roads, and urban development, have altered floodplain roughness. Artificial structures like levees further decrease floodplain extent and disrupt land cover, reducing energy dissipation in the remaining floodplain (Knox et al., 2022). Consequently, the original vegetation classes developed for canals may no longer adequately explain floodplain roughness in overbank areas.

The main goal of this study was to characterize roughness in floodplains across the continental US and its relationship with flow and vegetation characteristics. Specially, we examined how floodplain roughness varied with flow velocity, vegetation inundation fraction, and floodplain biomass. Direct estimates of floodplain Manning's *n* were produced using field measurements collected by the United States Geological Survey (USGS) during overbank flows. Estimated *n* values were then related to remotely sensed vegetation height and biomass data to quantify their influence on energy dissipation in floodplains. Finally, an empirical function was developed to characterize interactions between floodplain roughness, velocity, and vegetation properties. Additionally, we conducted cross-validation analyses to validate our methodology and compared our results with existing approaches for estimating floodplain roughness.

2 Materials and Methods

In this study, Manning's equation is applied specifically to the floodplain, separate from the main river channel. The floodplain discharge is isolated by subtracting the discharge within the main channel from the total measured discharge (see Supporting Information Figure S1 for a schematic of the floodplain as defined in this study). Values of Manning’s *n* are then derived by inverting Manning's equation and solving for the floodplain roughness (see Supporting Information) during periods of overbank flow (Reclamation, 2001). The necessary parameters for calculation of *n* are obtained from field measurements datasets provided by the USGS (USGS, 2021a). The flood stage height is determined by the National Weather Service (NWS, 2021; Slater et al., 2015), and friction slope estimates are obtained from the National Hydrography Dataset (NHD) (USGS, 2021b). Estimates of *n* were constrained to those sites meeting strict quality control metrics including consistency with current USGS rating curves and observed channel geometries (Liu, 2011; Vinutha et al., 2018).

$$Q = \frac{k}{n} S^{1/2} R^{2/3} A \qquad \text{(eq. 1)}$$

where Q is discharge [L^3 t^{-1}], S is the friction slope, defining the energy loss along a reach [L L^{-1}], R the hydraulic radius [L], A is cross-sectional area [L^2], k is a unit conversion factor, and n is Manning’s roughness coefficient.

At USGS gauging stations where *n* values are estimated, vegetation characteristics, such as aboveground biomass density and vegetation canopy height, are obtained from NASA's Global Ecosystem Dynamics Investigation (GEDI) (Potapov et al., 2021). GEDI is a LiDAR system mounted on the International Space Station that provides calibrated values of vegetation height and biomass globally at a 25m base resolution and gridded final products at 1km resolution (Dubayah et al., 2021, 2022; Milenković et al., 2022) . Previous research suggests that Manning's roughness coefficient is related to vegetation inundation fraction, flow velocity, and vegetation properties (Chow, 1959; Yen, 1992; Rob Ferguson, 2013). A semi-empirical function of n is formulated, based on prior models, that incorporates GEDI-derived vegetation properties. The function parameters are determined by fitting a linearized equation to values of Manning's roughness coefficient, flow velocity, and aboveground biomass at USGS sites. For a detailed explanation of the methodology please refer to the Extended Methodology section S1 in the supplementary information document.

To assess the performance of our newly developed function, we conducted a cross-validation analysis, which involved the application of Manning's equation to compute floodplain flow during observed overbank events. This process utilized the same measurements acquired by the USGS, along with Manning's n values estimated through a five-fold cross-validation approach (detailed in the Supplementary Information). Importantly, the Manning's n values used for fitting our function were distinct from those employed to validate discharge calculations at these sites.

To comprehensively evaluate our method, we compared the results not only against the directly measured discharge but also against discharges calculated using estimated roughness coefficients from other studies. These alternative approaches include the Geospatial Stream Flow Model (GeoSFM) proposed by (Asante et al., 2008), which parameterizes Manning's n values for different land cover classes for use in a distributed hydrologic model. This model integrates geospatial and time-series data in near-real time, generating daily forcing evapotranspiration and precipitation data from various remote sensing and ground-based sources. GeoSFM employs widely available terrain, soil, and land cover datasets for initial model setup and parameter estimation, making it adaptable for data-scarce environments. The model performs geospatial preprocessing and postprocessing tasks and hydrologic modeling within an ArcView GIS environment, offering seamless integration of GIS routines and time series processing. It identifies and maps wide-area streamflow anomalies, disseminating daily results, including streamflow and soil water maps, through various channels (Internet map servers, flood hazard bulletins, and more).

Additionally, Kalyanapu et al., (2009) determined Manning's n values by land cover class in a hydrologic modeling study focused on understanding the effects of land cover use on runoff and peak discharge. This research assesses the sensitivity of hydrologic models to Manning's *n* changes, a parameter crucial for representing surface roughness. Large watershed models often rely on land use/land cover datasets to assign Manning's n values based on land use or cover classes. While this approach is convenient, it introduces potential errors. Kalyanapu's study compared Manning's n values derived from manual inspection of aerial photos to those estimated using the National Land Cover Dataset (Homer et al., 2012). The results revealed significant differences in the magnitude and spatial distribution of Manning's n values, particularly at subcatchment levels. These differences, while not significantly altering runoff responses at the watershed outlet for large-scale models, became pronounced with increasing Manning's *n* deviation.

To ensure a fair and consistent comparison, we standardized our analysis using the International Geosphere-Biosphere Programme (IGBP) land cover classification (Loveland et al., 1999). Within this framework, we calculated the median velocity and median flow depth for each land cover class and subsequently derived the Manning's n value using our model. This approach allowed us to assess the performance of our function in relation to established methodologies and gain valuable insights into its efficacy in estimating floodplain roughness. We use medians instead of raw values to address the potential bias introduced by the inherent relationship between velocity and roughness, allowing for a fairer comparison against methodologies that do not consider velocity during the selection process.

3 Results

After data processing and quality control, a total of 4,927 estimates of floodplain Manning's *n* were calculated successfully at 804 sites, based on the analysis of 3,379,166 total measurements obtained from 31,142 unique gauge sites (Barinas et al., 2023). Included with this dataset of generated *n* values (see dataset in Supporting Information) are all the necessary variables measured by the USGS that were used when inverting Manning's equation to solve for *n*: measured discharge (Q), width (w), depth (z) obtained from USGS field measurements, and friction slope (S) from the NHD datasets. Intermediate variables are also included in this dataset: discharge, velocity, width, and depth, for both the main channel (Q_{mc}, V_{mc}, w_{mc}, z_{mc}) and the floodplain (Q_{fp}, V_{fp}, w_{fp}, z_{fp}). Complementary information included in the dataset are the USGS site ID, date of measurement, coordinates, and number of values of *n* calculated at that site.

Examining all floodplain roughness estimates over the continental United States, the national median of the estimated floodplain Manning's *n* values was 0.021, with a 5th and 95th percentile of 0.001 and 0.326, respectively. On average, a mean of 18 values of *n* were obtained per site, with an average of 155 values per state. Site-averaged *n* values revealed consistent spatial patterns across the continental United States (see Supporting Information Figure S2). These patterns are influenced by factors like vegetation biomass and velocities (Figure 1).

Vegetation biomass was shown to drive variability in floodplain roughness, with values of *n* for different vegetation classes and heights complied in Table S1 in the Supplementary Information). Areas dominated by Grasses, Shrubs, and Woodland, the most common vegetation classification in the GEDI dataset, tended to have a median Manning's n value of 0.017 for a median biomass on the analyzed sites of 18 Mg/Ha. Deciduous Broadleaf Trees, the second most common class, exhibited slightly higher roughness with a median Manning's n value of 0.025, having a median biomass of 77 Mg/Ha. Evergreen Broadleaf and Evergreen Needleleaf, despite having similar biomass densities (95 Mg/Ha and 106 Mg/Ha, respectively) contributed to different roughness values, with median Manning's n values of 0.030 and 0.010, respectively. Due to a limited number of samples, there were not enough observations to draw conclusions about the impact of Deciduous Needleleaf Trees on floodplain roughness (see Supplementary Information Table 1).

Even at a broad scale with the relatively low-resolution, remotely-sensed vegetation (GEDI) datasets used in this project, clear patterns were found between the floodplain Manning's *n* values and features (i.e. biomass, submergence) expected to predict *n* values at various velocity ranges (Figure 1). The values of *n* were inversely related to flow velocity and positively related to vegetation inundation fraction. Velocities were lowest at locations where Manning's *n* was highest. Within three velocity ranges, Manning's *n* varied with inundation fraction and vegetation biomass. Median Manning's *n* values ranged from 0.001-0.009 for the highest velocities ($V > 3$ m/s), whereas median *n* values ranged between 0.008 and 0.053 for mid-range velocity flows (1-3m/s). Under these mid to high velocities (V > 1m), Manning's *n* increased consistently with the inundation fraction and inconsistently with vegetative biomass. For low velocity flows (<1m/s), *n* ranged from 0.030 up to 0.388 and increases in roughness were associated inconsistently with both inundation fraction and vegetative biomass.

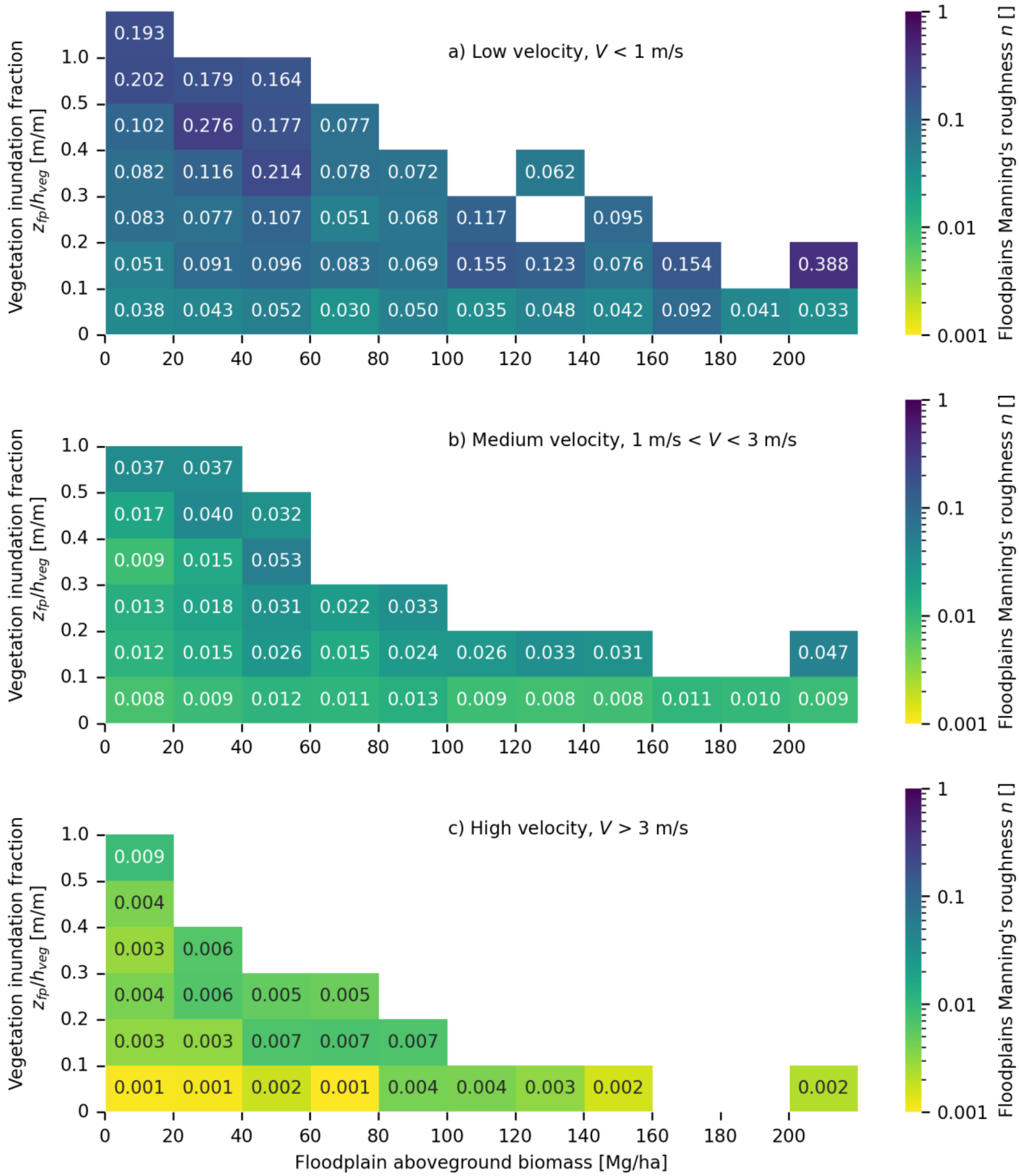

Figure 1. Median floodplain Manning's n values for different levels of floodplain aboveground biomass and vegetation inundation fraction. Numerical values within each box represent the median n value for the corresponding range of vegetation inundation fraction and aboveground biomass and results shown only when at least five values are available.

Based on calculated n values, observed flow velocities (V) and depths within the floodplain (z_{fp}), as well as GEDI estimated vegetation height (h_{veg}) and biomass (B), an empirical function relating Manning's n (See Extended Methodology S1) provided a reasonable fit to observed data ($r^2 = 0.74$):

$$n = 0.0321 \frac{B^{0.20}}{V^{0.99}} \left(\frac{z_{fp}}{h_{veg}}\right)^{0.5} \quad \text{(eq. 2)}$$

This function, visualized across observed conditions in Figure 2, predicted n with a root mean squared error (RMSE) of 0.024 (see scripts in Supporting Information). It further illustrated how Manning's n varies with flow and vegetation properties, with an inverse proportionality between Manning's n and flow velocity. A difference in roughness of nearly one order of magnitude was found between low velocities (<1m/s) and very high velocities (up to 5m/s) (Figure 2). Within specific velocity ranges, the values of n are notably influenced by vegetation inundation fraction, with greater roughness associated with higher levels of inundated vegetation. Furthermore, the data and function demonstrated that biomass tended to increase roughness more at low biomass levels (visually inspecting tangent lines revealed the inflection point to be approximately 30 Mg/Ha), whereas its influence decreased at higher biomass levels. This could explain why the function had less predictive power with biomass at higher levels of vegetation inundation fractions; High inundation fractions were not frequently observed at high biomass levels.

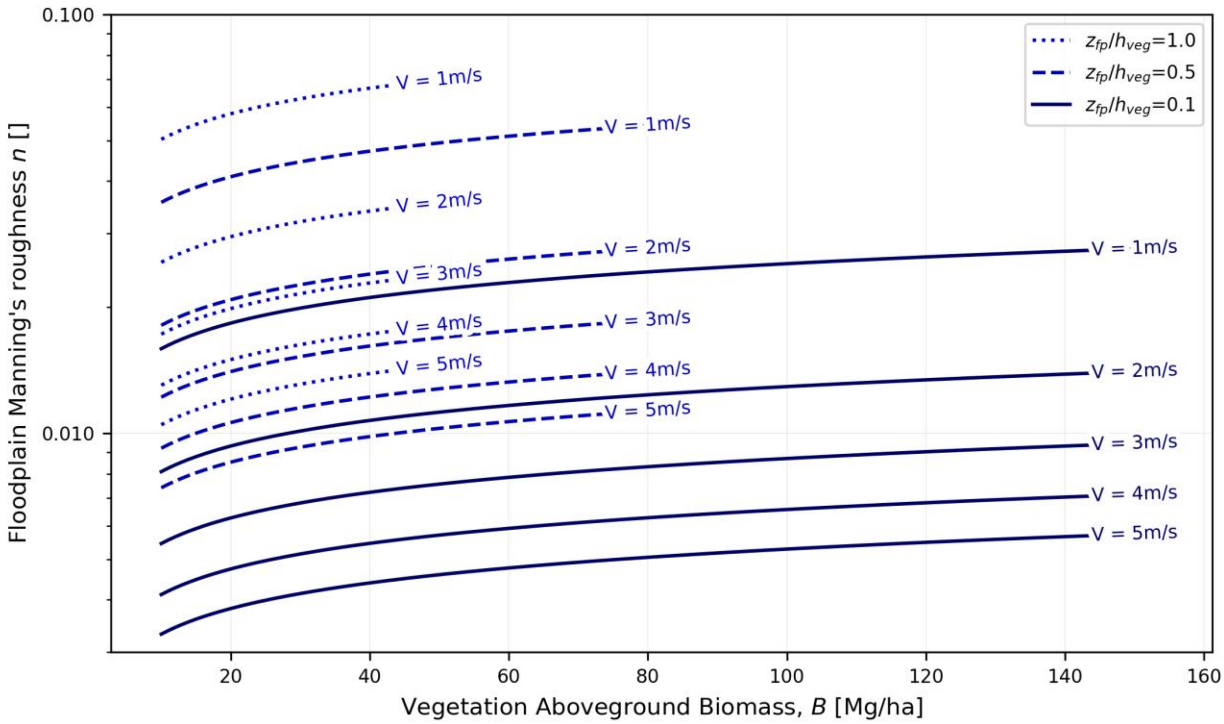

Figure 2 – Manning's *n* modelled as a function of aboveground biomass, *B*, and flow velocity, *V*, modeled for different levels of vegetation inundation fraction (z_{fp}/h_{veg}). Lines extend up to biomass levels of 50, 80 and 150 Mg/ha for fractions of inundation of 1.0, 0.5 and 0.1, respectively based on the total number of values within each range as depicted in Figure 1.

The cross-validation analysis conducted in this study reveals the performance of the proposed function in estimating USGS measured flows (Figure 3). Our findings indicate that this function offers higher accuracy and less dispersion, as evidenced by a Kling-Gupta efficiency (KGE) of 0.38 and a percent bias (PBIAS) of -16%. In comparison, alternative methods for determining roughness coefficients yielded less accurate results, with KGE values of 0.33 and 0.10, and PBIAS values of -43% and -85% for GeoSFM and Kalyanapu et al. (2009), respectively.

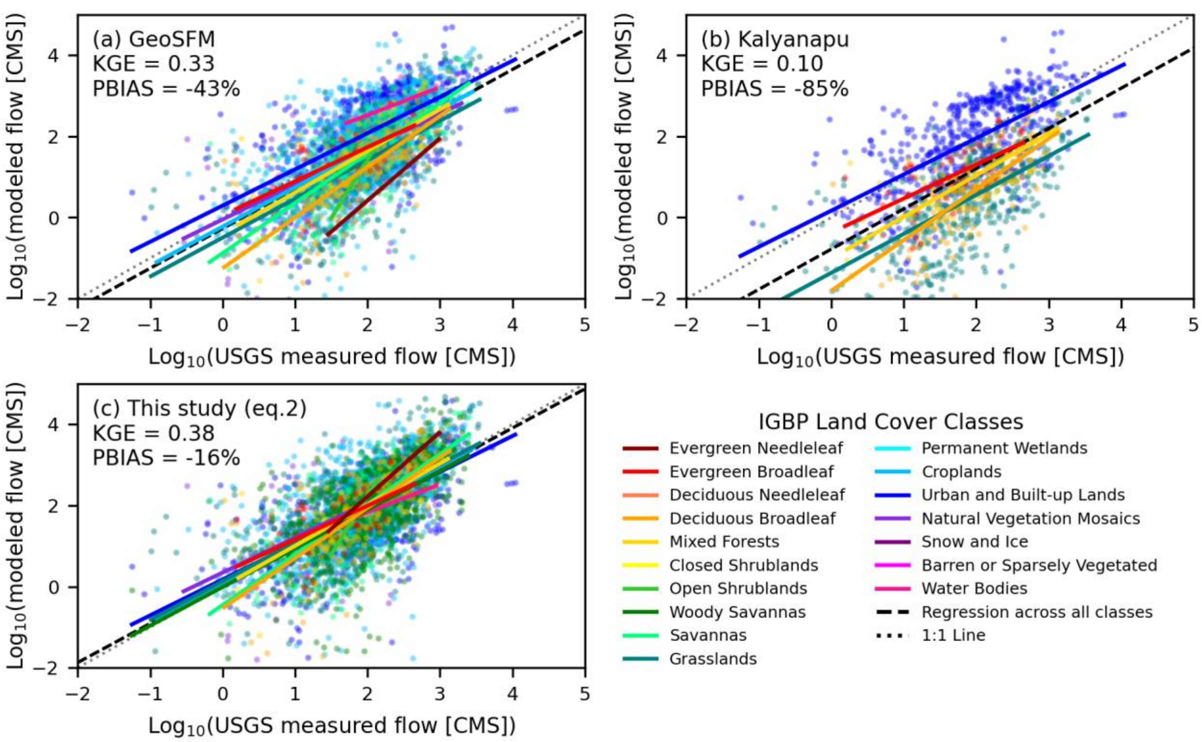

Figure 3 - Measured vs. estimated discharge based on three approaches to estimating floodplain roughness: (a) the Geospatial Stream Flow Model (GeoSFM), (b) Kalyanapu et al's study (2009) on land-use effects on model outputs, and (c) from the function developed in this study (eq. 2). These were calculated with median velocities and median flow depths per land cover class. Kling-Gupta efficiency (KGE) and percent bias (PBIAS) are reported across all vegetation classes.

4 Discussion

Floodplains serve critical functions for society through dissipation of flood energy, among other functions, but understanding of floodplain hydraulics contains large uncertainties due in part to limited field observations of floodplain roughness. This study produced new estimates of floodplain roughness coefficients that span the range of the continental United States. The average estimates of floodplain Manning's *n* calculated in this study were similar to *n* values modeled from field measurements of vegetation features (Kouwen & Fathi-Moghadam, 2000) and the values in Chow's look-up table (Chow, 1959). Chow identified the range of average *n* values for floodplains as being from 0.040 (in cleared land with stumps) up to 0.150 (for dense willows). In comparison, the average *n* estimates in this dataset were 0.060 for low canopy height and low levels of biomass, and 0.090 for high canopy height and biomass. Kouwen and Fathi-Moghadam's (2000) study also presented mean values of *n* for four tree types that range between 0.100 for high velocity flows (2 m/s) and 0.200 for very low velocity flows (0.1 m/s) for submerged conditions (z_{pf}/h_{veg}=1), dropping down to a range of 0.030 to 0.070 for low inundation (z_{pf}/h_{veg}=0.1). A similar pattern was observed in mean values of floodplain *n* in this study (Figure 2), ranging from 0.030 for low inundation and comparable velocity (V = 1-3 m/s), up to an average of 0.250 for low velocity (V < 1 m/s) and high inundation fraction. The approach presented here has the advantage of applying global, remotely sensed biomass datasets, compared with Kouwen and Fathi-Moghadam's vegetation index, which requires local measurements of frequency, mass, and height of the trees.

Field observations revealed that Manning's *n* in floodplains was generally lower at higher velocities than at lower velocities. Even though in practice Manning's *n* is often assumed to be a constant value solely determined based on the characteristics of the surface, in reality it has been demonstrated that *n* varies with discharge (Box et al., 2021; Chow, 1959; R. Ferguson, 2013). In most river channels, Manning's *n* decreases as discharge and stage increase due to lower roughness along the banks and the submergence of bed forms with increasing flow depths (USGS, 2012). This phenomenon is also consistent with the long history of roughness in pipe flow studies (Rouse, 1943). Like river channels, where previous research has shown that flow and velocity tend to have an inverse relationship with flow resistance, our calculations demonstrate a similar pattern in floodplains. This alignment with existing research suggests that flow and velocity in both river channels and floodplains exhibit an inverse relationship with flow resistance (Chow, 1959; R. Ferguson, 2013). Mechanistically, the inverse relationship could be a result of higher roughness reducing velocities, or the bending of flexible vegetation that reduces roughness at higher velocities. Datasets presented herein are inadequate for determining the source of the relationship.

This work demonstrated that GEDI's vegetation characteristics can be used to estimate floodplain roughness. Vegetation inundation fraction was an important predictor of Manning's *n*, as demonstrated in other settings (Nepf, 2012). In addition, this national Manning's *n* database reflects how floodplain roughness increases with aboveground biomass, though relative inundation demonstrated a stronger influence on roughness than biomass. This makes sense given that a key factor influencing Manning's *n* is the total vegetation cross section obstructing flow, not just the height of the canopy (Chow, 1959). Furthermore, previous studies have found that the density of vegetation in channels was a dominant parameter for Mannings's *n* in emergent conditions (Fathi-Maghadam & Kouwen, 1997) and the analysis here demonstrated that this finding translated to the floodplain as well. Since GEDI measures these vegetation properties globally, estimations of floodplain roughness can be extended worldwide with this method, with some caveats discussed below.

As previously outlined in the methodology section, our assessment involved a cross-validation analysis of the function defined in Equation 2. This process included the application of Manning's equation (eq. 1) to calculate floodplain flow during observed overbank events, using measurements from the US Geological Survey (USGS) and Manning's n coefficients estimated by our function. We also compared our findings with discharge estimates obtained from previous studies by Asante et al. (2008) and Kalyanapu et al. (2009), offering valuable insights into the robustness of our approach.

Our cross-validation analysis reveals notable advantages of the proposed function, which is rooted in US Geological Survey (USGS) gage data. This function demonstrated superior performance with a Kling-Gupta efficiency (KGE) of 0.38 and a percent bias (PBIAS) of -16% in estimating USGS measured flows. In comparison, alternative methods for determining roughness coefficients, such as GeoSFM (KGE = 0.33, PBIAS = -43%) and Kalyanapu's approach (KGE = 0.10, PBIAS = -85%), yielded less accurate results. Importantly, the other methods consistently underestimated flow rates across various land cover types when relying on constant roughness coefficients. This artifact is due to land cover –roughness coefficient classifications being defined based on steady and uniform flow conditions in channels (Chow, 1959) and not accounting for variation of resistance with changing flow, especially during flood events with higher flow rates. This is evident in the fact that the hydrologic models analyzed in these works utilized hydrographs, which involve unsteady flow characterized by changing flow over time. As a consequence, the roughness coefficient becomes variable in reality but not in the models. By incorporating a vegetation- and submergence-dependent Manning's *n* coefficient, the proposed function captured varying hydraulic conditions, leading to improved flow estimates when compared to methods that rely on a roughness coefficient that is independent of hydraulic conditions. Supporting this interpretation, both the GeoSFM and Kalyanapu et al. (2009) methods demonstrated relatively accurate estimates for short vegetation classes such as urban areas, built-up lands, and croplands, although they still lacked the precision displayed by our function in this study.

The Manning's *n* dataset and the function proposed in Eq. 2 have the potential to improve the performance of large-scale models such as the National Water Model (NWM). Many attempts are currently being made to reduce uncertainty in nationwide models (Johnson et al., 2019; Rojas et al., 2020), but have been focused on improving its performance by updating the geometry and roughness parameters of the main channel, without extending improvements to the floodplain (Heldmyer et al., 2022). Integrating the results from this work on floodplain roughness at USGS gauge locations into the NWM could be a logical next step.

Our study introduces a novel approach to enhance the NWM, especially during flood events, by incorporating dynamic floodplain roughness values. These values account for variations in flow velocity and vegetation properties, essential factors that are traditionally treated as constants in large-scale models. This integration offers the potential for more accurate flood predictions, improved flood risk assessments, and enhanced river management strategies. It's important to acknowledge the possibility of adjustments to other key parameters, such as channel roughness. While our study doesn't prescribe a specific approach for these adjustments, it opens an intriguing avenue for future research and collaboration.

The study datasets were subject to some limitations, including those inherent to the USGS monitoring network (Kiang et al., 2013; Tu et al., 2023), as discussed in the SI. Gaging limitations may narrow the generalizability of the results to LULCs (Land Use Land Cover) and geographic regions included in this gaging network. Further, assumptions about the geometry of a river's cross-section were made that could be inconsistent in some channels, such as where the local slope is too high or width too narrow to maintain that a hydraulic radius that is approximately equal to the hydraulic depth. Furthermore, the vegetation phenology is a snapshot in time, though it has been established that considerable differences exist in vegetation characteristics between seasons that can impact flow (Bond et al., 2020). To provide high-quality biomass and height estimates, the GEDI averages measurements. The resulting derived products do not represent a specific time of year, in contrast with USGS field measurements that were made on a specific date. Finally, vegetation data were sampled from GEDI's 1 km^2 gridded product for the area around each USGS gauge site, which leads to questions regarding what area influences Manning's roughness. Energy dissipation occurs via multiple processes during a flood (R. Ferguson, 2013), but the area of influence that has a direct effect on flow is poorly understood and is worthy of further study. The assumption made for these calculations is that the 1 km^2 average for the vegetation characteristics taken from GEDI measurements is representative of the actual area influencing energy dissipation during a flood. This assumption may not be valid at sites where there is a large variation in land cover within a 1 km^2 grid.

5 Conclusions

Floodplain roughness is a critical aspect of managing floodplains, and its societal relevance will rise with rising floodwaters under climate change, expanding floodplain development, aging flood infrastructure, and rising emphasis on floodplain reconnection for nature-based flood infrastructure and ecological restoration. While Manning's n is typically assumed to be a constant value in floodplain analysis and engineering applications, this study demonstrated that accurate estimation of current and modified floodplain roughness should rely on vegetation submergence and velocities, with biomass playing a smaller role.

The dataset of floodplain Manning's *n* generated in this work, and its correlation with flow and vegetation characteristics, further supported prior findings that flow resistance during a flood increases with submergence depth and biomass, and that resistance is inversely related to flow velocity. This work utilized a unique coupling of existing datasets, considering tall vegetation biomes, and demonstrated how flow and vegetation properties influence roughness across a wide range of regions and climates in the continental United States, rather than limited to a specific site or sites. Results should be generalizable across scales and landscapes that align with the input datasets and should support the management and restoration community in establishing sustainable floodplains.

Acknowledgments

We would like to acknowledge the support of the Fulbright Program and Oregon State University through the Graduate Fellowship program. This project has been partially funded by the NASA grant 80NSSC21K0198 award to Stephen Good. We would also like to thank Cara Walter and Harrison Lee Kutz for their support with the processing and retrieval of NHD and GEDI data.

Open Research

The scripts to reproduce the dataset, figures and tables in this study are openly available at Barinas et al. (2023) under a Creative Commons Attribution License format (free registration required).

6 References

Asante, K. O., Artan, G. A., Pervez, S., Bandaragoda, C., & Verdin, J. P. (2008). Technical manual for the geospatial stream flow model (GeoSFM). *World Wide Web*, *605*, 594–6151.

Barinas, G., Good, S., & Tullos, D. (2023). Dataset and Supporting Information for Barinas et. al. Floodplain Roughness Study [Data set]. CUAHSI HydroShare. Retrieved from https://doi.org/10.4211/hs.8c5b71cf3a7449b0b94c447a3e87e63f

Bond, S., Kirkby, M. J., Johnston, J., Crowle, A., & Holden, J. (2020). Seasonal vegetation and management influence overland flow velocity and roughness in upland grasslands. *Hydrological Processes*, *34*(18), 3777–3791. https://doi.org/10.1002/hyp.13842

Box, W., Järvelä, J., & Västilä, K. (2021). Flow resistance of floodplain vegetation mixtures for modelling river flows. *Journal of Hydrology*, *601*, 126593. https://doi.org/10.1016/j.jhydrol.2021.126593

Bulti, D. T., & Abebe, B. G. (2020). A review of flood modeling methods for urban pluvial flood application. *Modeling Earth Systems and Environment*, *6*(3), 1293–1302. https://doi.org/10.1007/s40808-020-00803-z

Chow, V. T. (1959). Chow, V. T. (1959). Open-channel hydraulics. McGraw-Hill, New York - Google Search. Retrieved August 28, 2022, from

Cowan, W. L. (1956). Estimating hydraulic roughness coefficients. *Agricultural Engineering*, *37*(7), 473–475.

Dubayah, R. O., Luthcke, S. B., Sabaka, T. J., Nicholas, J. B., Preaux, S., & Hofton, M. A. (2021). GEDI L3 Gridded Land Surface Metrics, Version 2. *ORNL DAAC*. https://doi.org/10.3334/ORNLDAAC/1952

Dubayah, R. O., Armston, J., Healey, S. P., Yang, Z., Patterson, P. L., Saarela, S., et al. (2022). GEDI L4B Gridded Aboveground Biomass Density, Version 2. *ORNL DAAC*. https://doi.org/10.3334/ORNLDAAC/2017

Durand, M., Gleason, C. J., Garambois, P. A., Bjerklie, D., Smith, L. C., Roux, H., et al. (2016). An intercomparison of remote sensing river discharge estimation algorithms from measurements of river height, width, and slope: SWOT DISCHARGE ALGORITHM INTERCOMPARISON. *Water Resources Research*, *52*(6), 4527–4549. https://doi.org/10.1002/2015WR018434

Fathi-Maghadam, M., & Kouwen, N. (1997). Nonrigid, Nonsubmerged, Vegetative Roughness on Floodplains. *Journal of Hydraulic Engineering*, *123*(1), 51–57. https://doi.org/10.1061/(ASCE)0733-9429(1997)123:1(51)

Ferguson, R. (2013). 9.5 Reach-Scale Flow Resistance. In *Treatise on Geomorphology* (pp. 50–68). Elsevier. https://doi.org/10.1016/B978-0-12-374739-6.00230-X

Ferguson, Rob. (2013). 9.5 Reach-Scale Flow Resistance. In *Treatise on Geomorphology* (pp. 50–68). Elsevier. https://doi.org/10.1016/B978-0-12-374739-6.00230-X

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. https://doi.org/10.5194/hess-26-6121-2022

Homer, C. H., Fry, J. A., & Barnes, C. A. (2012). The national land cover database. *US Geological Survey Fact Sheet*, *3020*(4), 1–4.

Hunter, N. M., Bates, P. D., Horritt, M. S., & Wilson, M. D. (2007). Simple spatially-distributed models for predicting flood inundation: A review. *Geomorphology*, *90*(3), 208–225. https://doi.org/10.1016/j.geomorph.2006.10.021

Johnson, J. M., 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. https://doi.org/10.5194/nhess-19-2405-2019

Kalyanapu, A., Burian, S., & Mcpherson, T. (2009). Effect of land use-based surface roughness on hydrologic model output. *Journal of Spatial Hydrology*, *9*, 51–71.

Kiang, J. E., Stewart, D. W., Archfield, S. A., Osborne, E. B., & Eng, K. (2013). *A national streamflow network gap analysis* (USGS Numbered Series No. 2013–5013). *A national streamflow network gap analysis* (Vol. 2013–5013). Reston, VA: U.S. Geological Survey. https://doi.org/10.3133/sir20135013

Knox, R. L., Morrison, R. R., & Wohl, E. E. (2022). A river ran through it: Floodplains as America's newest relict landform. *Science Advances*, *8*(25), eabo1082. https://doi.org/10.1126/sciadv.abo1082

Kouwen, N., & Fathi-Moghadam, M. (2000). Friction Factors for Coniferous Trees along Rivers. *Journal of Hydraulic Engineering*, *126*(10), 732–740. https://doi.org/10.1061/(ASCE)0733-9429(2000)126:10(732)

Lee, H.-L., & Mays, L. W. (1986). Hydraulic Uncertainties in Flood Levee Capacity. *Journal of Hydraulic Engineering*, *112*(10), 928–934. https://doi.org/10.1061/(ASCE)0733-9429(1986)112:10(928)

Liu, X. (2011). Accuracy Assessment of Lidar Elevation Data Using Survey Marks. *Survey Review*, *43*(319), 80–93. https://doi.org/10.1179/003962611X12894696204704

Loveland, T. R., Zhu, Z., Ohlen, D. O., Brown, J. F., Reed, B. C., & Yang, L. (1999). An analysis of the IGBP global land-cover characterization process. *Photogrammetric Engineering and Remote Sensing*, *65*, 1021–1032.

Manning, R. E. (1891). On the flow of water in open channels and pipes.

Milenković, M., Reiche, J., Armston, J., Neuenschwander, A., De Keersmaecker, W., Herold, M., & Verbesselt, J. (2022). Assessing Amazon rainforest regrowth with GEDI and ICESat-2 data. *Science of Remote Sensing*, *5*, 100051. https://doi.org/10.1016/j.srs.2022.100051

Nepf, H. M. (2012). Flow and Transport in Regions with Aquatic Vegetation. *Annual Review of Fluid Mechanics*, *44*(1), 123–142. https://doi.org/10.1146/annurev-fluid-120710-101048

NWS. (2021). USGS WaterWatch - Streamflow conditions. Retrieved November 3, 2022, from https://waterwatch.usgs.gov/

Petryk. (1975). Analysis of flow through vegetation.

Potapov, P., Li, X., Hernandez-Serna, A., Tyukavina, A., Hansen, M. C., Kommareddy, A., et al. (2021). Mapping global forest canopy height through integration of GEDI and Landsat data. *Remote Sensing of Environment*, *253*, 112165. https://doi.org/10.1016/j.rse.2020.112165

Reclamation, U. S. B. of. (2001). *Water measurement manual*. The Bureau. Retrieved from https://books.google.com/books?hl=es&lr=&id=7PdRAAAAMAAJ&oi=fnd&pg=PA11&dq=water+measurements+manual+usbr&ots=ajRs6q1ds-&sig=atpXC2u8__bFWTy9ndNikoeB7a8

Rojas, M., Quintero, F., & Krajewski, W. F. (2020). Performance of the National Water Model in Iowa Using Independent Observations. *JAWRA Journal of the American Water Resources Association*, *56*(4), 568–585. https://doi.org/10.1111/1752-1688.12820

Rouse, H. (1943). Evaluation of Boundary Roughness. *Proceedings Second Hydraulic Conference, University of Iowa Bulletin 27*.

Slater, L. J., Singer, M. B., & Kirchner, J. W. (2015). Hydrologic versus geomorphic drivers of trends in flood hazard. *Geophysical Research Letters*, *42*(2), 370–376. https://doi.org/10.1002/2014GL062482

Smith, A. B. (2020). U.S. Billion-dollar Weather and Climate Disasters, 1980 - present (NCEI Accession 0209268) [Data set]. NOAA National Centers for Environmental Information. https://doi.org/10.25921/STKW-7W73

Tu, T., Comte, L., & Ruhi, A. (2023). The color of environmental noise in river networks. *Nature Communications*, *14*(1), 1728. https://doi.org/10.1038/s41467-023-37062-2

USGS. (2012). USGS Data Series 668: Manning's Roughness Coefficients for Illinois Streams. Retrieved September 29, 2022, from https://pubs.usgs.gov/ds/668/pdf/DataSeries_668_2.pdf

USGS. (2021a). Measurements - Surface Water for USA: Streamflow Measurements. Retrieved August 25, 2022, from https://waterdata.usgs.gov/nwis/measurements

USGS. (2021b). National Hydrography Dataset | U.S. Geological Survey. Retrieved August 25, 2022, from https://www.usgs.gov/national-hydrography/national-hydrography-dataset

Vinutha, H. P., Poornima, B., & Sagar, B. M. (2018). Detection of Outliers Using Interquartile Range Technique from Intrusion Dataset. In S. C. Satapathy, J. M. R. S. Tavares, V.

Bhateja, & J. R. Mohanty (Eds.), *Information and Decision Sciences* (pp. 511–518). Singapore: Springer. https://doi.org/10.1007/978-981-10-7563-6_53

Yen, B. C. (1992). *Channel Flow Resistance: Centennial of Manning's Formula*. Water Resources Publication.

Yen, B. C. (2002). Open Channel Flow Resistance. *Journal of Hydraulic Engineering*, *128*(1), 20–39. https://doi.org/10.1061/(ASCE)0733-9429(2002)128:1(20)

Geophysical Research Letters

Supporting Information for

Continental Scale Assessment of Variation in Floodplain Roughness with Vegetation and Flow Characteristics

Gabriel Barinas[1, 2], Stephen P. Good[1, 2], and Desiree Tullos[1, 2]

[1]Water Resources Graduate Program, Oregon State University
[2]Department of Biological & Ecological Engineering, Oregon State University

Contents of this file

Additional Supporting Information (Files uploaded separately at Barinas et al., 2023)

Dataset file *fp_mannings.csv*
Script files *dataset.py* and *figures.py*
Required files: *NHDFlowline.csv*, *measurements.csv, gedi_L3L4.csv, ModisLC, geosfm.csv and kalyanapu.csv.*

Extended Methodology S1

1. Floodplain roughness definition

Manning's equation (Manning, 1891) is extensively applied in hydraulic modeling and is written as:

$$Q = \frac{k}{n} S^{1/2} R^{2/3} A, \quad \text{(S1)}$$

where Q is discharge [L3 t-1], S is the friction slope, defining the energy loss along a reach [L L-1], R the hydraulic radius [L], A is cross-sectional area [L2], k is a unit conversion factor, and n is Manning's roughness coefficient. The coefficient n is a representation of the roughness of the surface over which water is flowing and incorporates surface characteristics such as smoothness, grain size, vegetation and/or obstructions (Chow, 1959).

Here we conceptualize the floodplain as a wide, rectangular, cross-sectional area (see Figure S1) and apply Manning's equation explicitly to the floodplain alone, separate from the main river channel. The floodplain discharge, Q_{fp}, is isolated by taking the total measured discharge, Q_t, and subtracting the discharge within the main channel, Q_{mc}. The width of water in the floodplain, w_{fp} = w_t- w_{mc}, is assumed to be much greater than flow depth, z_{fp}, and therefore the hydraulic radius of flow in the floodplain is approximately equal to the floodplain flow depth (Reclamation, 2001). Rearranging Manning's equation (eq. S1) for the floodplain and solving for floodplain roughness, n_{fp}, yields the following relationship:

$$n_{fp} = \frac{k\, w_{fp}\, {z_{fp}}^{5/3}\, S^{1/2}}{Q_t - Q_{mc}} \quad \text{(S2)}$$

2. Bankfull width estimation

Parameters corresponding to the total flow (Q_t, w_t) are collected during overbank discharge measurements made by the United States Geological Survey (USGS) at gauging stations. These total parameters were used to derive floodplain specific parameters necessary to solve for the floodplain roughness. At bankfull depth, z_{bf}, the width of the main channel, w_{mc}, is not specified or measured, and must be estimated by determining the cross-sectional geometry of the main channel and floodplain. A piecewise function based on measurements of w and z was used to determine w_{mc} , with the main channel depth assumed a power function of the width, i.e. $z \propto w$ (Durand et al., 2016), and the floodplain as sloping linearly away from bankfull stage. The piecewise form of w as a function of z was then expressed as:

$$w = \frac{(z - z_0)^{1/s}}{u} \qquad \text{when } z < z_{bf} \qquad \text{(S3a)}$$

$$w = m\left(z - z_{bf}\right) + \frac{\left(z_{bf} - z_0\right)^{1/s}}{u}, \qquad \text{when } z \geq z_{bf} \qquad \text{(S3b)}$$

where m is the cross-sectional up-slope of the floodplain, u and s are parameters that define the shape of the main channel curvature, and z_0 defines its starting point.

Discharge in the main channel above flood stage was represented as a rectangular cross-section, as such flow in the main channel Q_{mc} should assumed to be proportional to $(z)^{5/3}$, following Manning's function (eq S1). When the flow reaches bankfull condition i.e. z = z_{bf}, $Q_{mc} = Q_{bf}$, and thus the constant of proportionality is equal to $Q_{bf}/{z_{bf}}^{5/3}$, and Manning's function (eq. S1) for the main channel flow above bankfull was simplified as:

$$Q_{mc} = \left(\frac{Q_{bf}}{{z_{bf}}^{5/3}}\right) z^{5/3} \qquad \text{(S4)}$$

where Q_{bf} is the flow at flood stage given that Q_{mc}= Q_t when $z = z_{bf}$.

3. Data sources and quality control

Most of the parameters (Q_t, w_t, z) required for the calculation of Manning's n with Eq. S2-S4, were obtained from the field measurements datasets available from the USGS WaterData platform (USGS, 2021a). The WaterData platform is part of the USGS efforts to monitor, assess, and deliver information about streamflow quality, use and availability. The platform provides access to field measurements at nearly 73,000 sites under USGS management. Consistent with Slater et al., (2015), the flood stage height (z_{bf}) was obtained from the WaterWatch platform (NWS, 2021). These values were determined by the National Weather Service by defining the flood stage as the lowest bank at which inundation of the surrounding area begins to cause damage. Friction slope estimates were obtained from the National Hydrography Dataset (USGS, 2021b), a database of features that includes a drainage network of US waterbodies.

Quality control measures on the floodplain data involved multiple steps. The rating curves at USGS sites are regularly adjusted to adapt the relationship to geometry changes associated with erosion or deposition at a gauging location. To account for this effect, only river discharge and geometry measurements where the measured values were within 10% of the respective rating curve value were included in the analysis. Additionally, sites with a low number of measurements over the flood stage (< 3) were also removed. As a way of avoiding the calculation of n with measurements with a high level of uncertainty in the width-depth relationship (eq. 3), as evidenced by regression curves with a high root mean squared error (RMSE), only samples with width measurements higher than 1.96 times the RMSE of the fit in eq. 3 were considered (Liu, 2011). Furthermore, a large percentage of sites from the resulting dataset (27%) had a slope set at a value of 0.0001 within the National Hydrography Dataset, representing a minimum fixed value within the database. Due to the high uncertainty and potential error from including the fixed minimum slopes, these 298 sites were also excluded from analysis.

In addition, the results of this work are subject to limitations of the USGS gaging network and to uncertainties inherent in gaging stochastic and modified systems. The lack of representation of certain geographic areas within the USGS gaging network have been reported elsewhere (Kiang et al., 2013), as have some of the drivers of temporal noise and uncertainty in streamflow over time (Tu et al., 2023). Application of the results outside the geographic areas and LULC (Land Use Land Cover) conditions from which these data were derived may generate uncertainties that we were unable to quantify with this analysis.

4. Remote sensed vegetation datasets

Flow in floodplains is expected to be strongly influenced by the vegetation in the floodplain (Box et al., 2021) but vegetation characteristics (density, height, etc.) are not typically measured in the field during flood conditions. Here we used aboveground biomass density, B [M L-2], and vegetation canopy height, h_{veg}, as characterized by the NASA Global Ecosystem Dynamics Investigation (GEDI). GEDI utilizes a full waveform Light Detection and Ranging (LiDAR) system to make measurements of vegetation structure at 25m resolution (Potapov et al., 2021), which are then aggregated to a 1km spatial resolution grid. In this work, for each USGS site, we obtained the canopy height estimates from the L3B version 2 gridded product (Dubayah et al., 2021) and the aboveground biomass estimates from the L4B version 2 gridded product (Dubayah et al., 2022).

Within GEDI's L4B dataset, there is a Prediction Stratum (PS) classification, determined by plant functional types described as: Deciduous Broadleaf Trees, Evergreen Broadleaf Trees, Evergreen Needleleaf Trees, Deciduous Needleleaf Trees, and Grasses, Shrubs, and Woodlands grouped as one class. This classification was used to categorize our dataset based on the level of biomass and canopy height by extracting the GEDI data from the pixel where each gauge location fell within (See Table S1). It is important to note that Gridded GEDI datasets, while providing unique information about vegetation height and biomass, is limited by its 1km resolution, capable of measuring only vegetation above a certain height.

5. Theoretical modeling

Prior research suggests that *n* is proportional to the square-root of the vegetation inundation fraction, i.e. n ∝ $(z_{fp}/h_{veg})^{1/2}$, and that it is also related to flow velocity and vegetation properties (Kouwen & Fathi-Moghadam, 2000). A mathematical model by

Kouwen was based on data from four tree species:

$$n = 0.228\left(\frac{V}{\sqrt{\frac{\xi E}{\rho}}}\right)^{-0.23}\left(\frac{y_n}{h}\right)^{0.5} \qquad \text{(eq. 5)}$$

where V is flow velocity, ξE is a vegetation index, ρ is the density of the fluid, and yn/h is the depth of submergence (z_{fp}/h_{veg}), and 0.228 and -0.23 empirically fit. Based on this approach, we formulate an analogous expression incorporating GEDI derived vegetation properties as $c = a_1 V^{a_2} B^{-a_3}$, where a_1, a_2 and a_3 are model parameters, B is aboveground biomass, V is flow velocity, and c is Manning’s n normalized by the square root of vegetation inundation fraction, i.e. $c = n/(z_{fp}/h_{veg})^{1/2}$. To ensure positive c (and n) values, the linearized equation:

$$\ln(c) = a_1 + a_2 \ln(V) + a_3 \ln(B) \qquad \text{(eq. 6)}$$

was fit to values of c, V, and B at USGS sites in our dataset to determine a_1, a_2 and a_3.

To limit the uncertainty caused by outliers in the dataset during the development of the model, the range of c values was restricted with the use of the interquartile range (IQR) (Vinutha et al., 2018). The minimum c value included in the analysis was the first quartile minus 1.5 times the IQR and the maximum value was the third quartile plus 1.5 times the IQR, where the IQR is equal to the difference between the third and the first quartile.

The developed function underwent cross-validation by splitting the USGS dataset, after quality control, into five randomized equal subsets, with each subset serving as validation during separate simulations. Subsequently, we combined all five validation subsets to create a comprehensive validation dataset that includes all of the original USGS gauge locations. This approach enables us to thoroughly assess the applicability and representativeness of our empirical function across the entire set of gauge locations. This new validation set was then used to compare the performance of our model against other works.

Figures S1 and S2

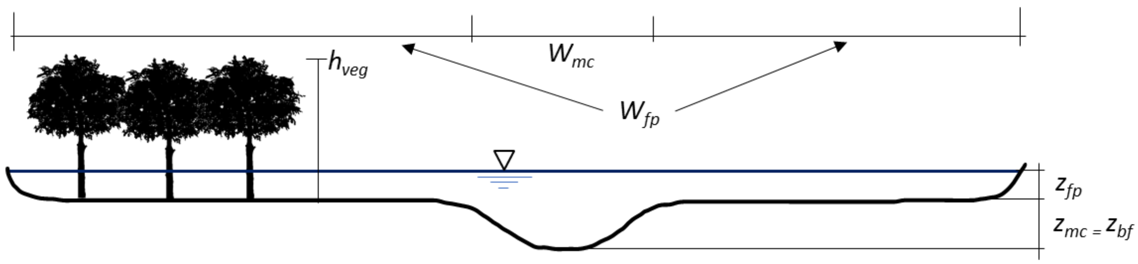

Figure S1 – Cross section diagram showing the variables used in the analysis. W_{mc} is the width of the main channel, W_{fp} is the width of the floodplain, Z_{mc} and Z_{bf} are the depth of the main channel during bankfull conditions, and h_{veg} is the height of the vegetation.

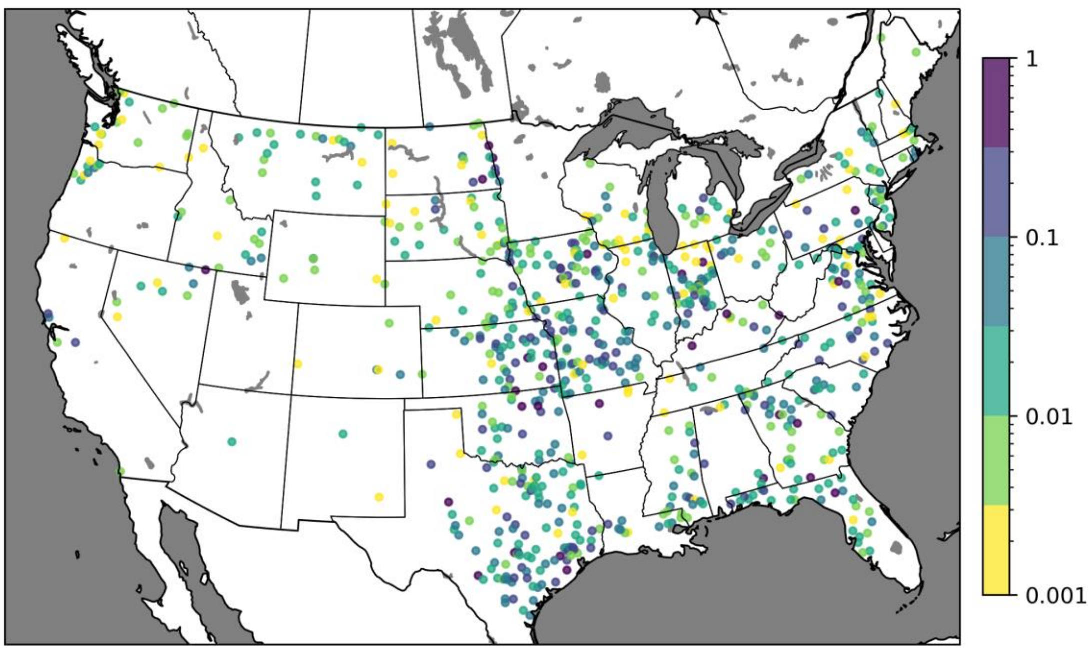

Figure S2 - Floodplain roughness (Manning's n) estimates for USGS sites.

Table S1

Mean floodplain Manning's n and Aboveground Biomass (B) [Mg/Ha] classified by tree structure and ranges of vegetation height. Values given are the median ± one median absolute deviation and the (samples count). The median vegetation height for these sites was 10m, while the 33rd and 66th percentiles were 7.4m and 13.5m. Values of 7.5m and 14m for vegetation height were selected for the ranges in order to have roughly the same number of total samples in each range.

<table>
<tr><th rowspan="3">Land Cover</th><th colspan="8">Floodplain Vegetation Biomass, B [Mg/Ha],
floodplain n, and (sample count)</th></tr>
<tr><th colspan="2">h_{Veg} <7.5m</th><th colspan="2">h_{Veg} 7.5-14m</th><th colspan="2">h_{Veg} >14m</th><th colspan="2">All heights</th></tr>
<tr><th>B</th><th>n</th><th>B</th><th>n</th><th>B</th><th>n</th><th>B</th><th>n</th></tr>
<tr><td rowspan="2">Deciduous Broadleaf Trees</td><td>25
±6</td><td>0.023
±0.022</td><td>63
±20</td><td>0.026
±0.022</td><td>130
±29</td><td>0.026
±0.020</td><td>77
±45</td><td>0.025
±0.022</td></tr>
<tr><td colspan="2">(340)</td><td colspan="2">(518)</td><td colspan="2">(652)</td><td colspan="2">(1514)</td></tr>
<tr><td rowspan="2">Evergreen Broadleaf Trees</td><td>37
±0</td><td>0.005
±0.004</td><td>50
±0</td><td>0.025
±0.022</td><td>98
±28</td><td>0.032
±0.022</td><td>95
±25</td><td>0.030
±0.022</td></tr>
<tr><td colspan="2">(5)</td><td colspan="2">(15)</td><td colspan="2">(150)</td><td colspan="2">(170)</td></tr>
<tr><td rowspan="2">Evergreen Needleleaf Trees</td><td>62
±32</td><td>0.007
±0.006</td><td>44
±5</td><td>0.011
±0.009</td><td>108
±2</td><td>0.011
±0.009</td><td>106
±19</td><td>0.010
±0.008</td></tr>
<tr><td colspan="2">(10)</td><td colspan="2">(37)</td><td colspan="2">(79)</td><td colspan="2">(126)</td></tr>
<tr><td rowspan="2">Grasses, Shrubs and Woodlands</td><td>12
±8</td><td>0.012
±0.010</td><td>28
±16</td><td>0.038
±0.034</td><td>51
±2</td><td>0.022
±0.013</td><td>18
±12</td><td>0.017
±0.014</td></tr>
<tr><td colspan="2">(976)</td><td colspan="2">(540)</td><td colspan="2">(196)</td><td colspan="2">(1737)</td></tr>
<tr><td rowspan="2">Unclassified</td><td>6
±2</td><td>0.021
±0.017</td><td>44
±15</td><td>0.025
±0.021</td><td>106
±36</td><td>0.020
±0.016</td><td>44
±36</td><td>0.023
±0.018</td></tr>
<tr><td colspan="2">(141)</td><td colspan="2">(264)</td><td colspan="2">(279)</td><td colspan="2">(787)</td></tr>
<tr><td rowspan="2">All GEDI land cover classes</td><td>15
±10</td><td>0.014
±0.012</td><td>44
±20</td><td>0.028
±0.024</td><td>110
±38</td><td>0.023
±0.017</td><td>38
±27</td><td>0.021
±0.018</td></tr>
<tr><td colspan="2">(1472)</td><td colspan="2">(1374)</td><td colspan="2">(1356)</td><td colspan="2">(4927)</td></tr>
</table>

Dataset

The dataset included as part of the Supplementary Information document is the result of the analysis that took place during this study. The dataset file named '*fp_mannings.csv*', consists of 4,927 calculations of Manning's n at each of the 804 USGS gauge sites that remained after quality control. The file also includes all variables collected and derived from USGS field measurements: discharge, width, depth (total, main channel, and floodplain), channel slope, site ID, coordinates, and number of estimates on that site.

Scripts

Included as supplementary information there are two scripts: *dataset.py* and *figures.py*.

The *dataset.py* script automates the process of calculating the Manning's *n* *fp_mannings.csv* file. It is divided into 3 sections: a SETUP section for module and file imports, a RUN section for defining the main function and running it for each state, and a MERGE section that puts together the results of each state into a single file. This script requires the *NHDFlowline.csv* and *measurements.csv* files which are included in Barinas et al., (2023). Other necessary files are downloaded automatically from USGS websites for each site: https://nwis.waterdata.usgs.gov/ for site coordinates; https://waterdata.usgs.gov/ for rating curves; and https://waterwatch.usgs.gov/ for flood stages.

The *figures.py* script creates the figures included in the main paper and in this document. This script is divided into 4 sections: IMPORTS loads the necessary modules and files required; MAP corresponds to Figure S2 in this document; MODEL corresponds to Figure 2 in the main manuscript; HEATMAP corresponds to figure 1 in the main manuscript, and VALIDATION corresponds to figure 3 in the main manuscript. This script requires remote sensed data by the Ecosystem Dynamics Investigation Mission GEDI. All GEDI data included in the *gedi_L3L4.csv* file refers to the pixel value where all USGS sites in the *fp_mannings.csv* file fell within and the data collected corresponds to the L3 and L4B version 2 gridded products (Dubayah et al., 2022). Finally, for the validation section of the script, the file *ModisLC.csv,* which contains the MODIS land cover classification corresponding to each gauge location, and the files *geosfm.csv* and *kalyanapu.csv*, which contains the values of *n* for each land cover type as presented in the original papers (Asante et al., 2008; Kalyanapu et al., 2009).