Deep Learning-based Models for Estimating River Channel Width

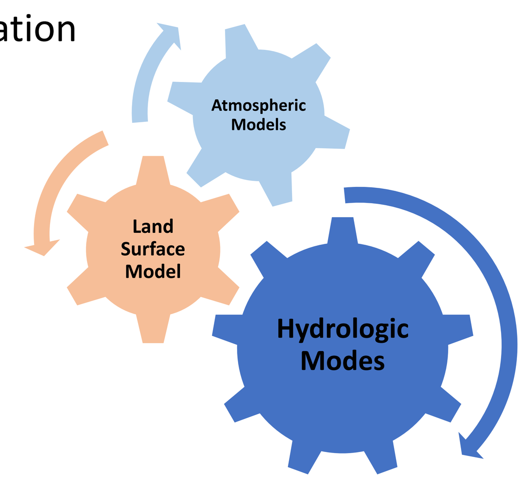
Mohammad Erfani, Zahra Ghahremani, Laura Manuel, Shuyu Chang
Erfan Goharian, Jennifer Pierce, Ehab Meselhe, Kimberly Van Meter, Chaopeng Shen, and Sagy Cohen

Background

National Water Model

NWM has been developed to facilitate improved representations of terrestrial hydrologic processes related to the spatial redistribution of surface, subsurface and channel waters across the land surface and to facilitate coupling of hydrologic models with atmospheric models. NWM includes:

- 1D Land Surface Parameterization
- Surface Overland Flow
- Saturated Subsurface Flow
- **Channel Routing**
- Reservoir Routing
- Conceptual Baseflow



Channel Routing in NWM

In order to represent a simplification of the behavior of a flood wave when it exceeds the channel bank, a compound channel formulation is added to the NWM.

- Trapezoidal geometry
- Side slope (z),
- Bottom width (Bw)
- Roughness (n)

Compound Channel in NWM

In order to represent a simplification of the behavior of a flood wave when it exceeds the channel bank, a compound channel formulation is added to the NWM.

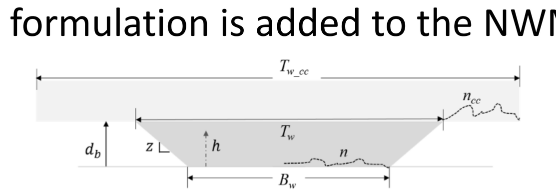


Fig 1: Cross-sectional schematic of trapezoidal channel and compound channel in NWM

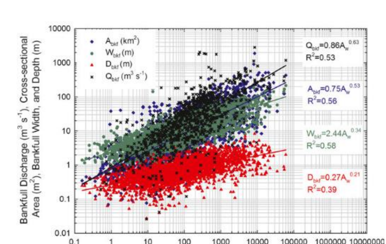
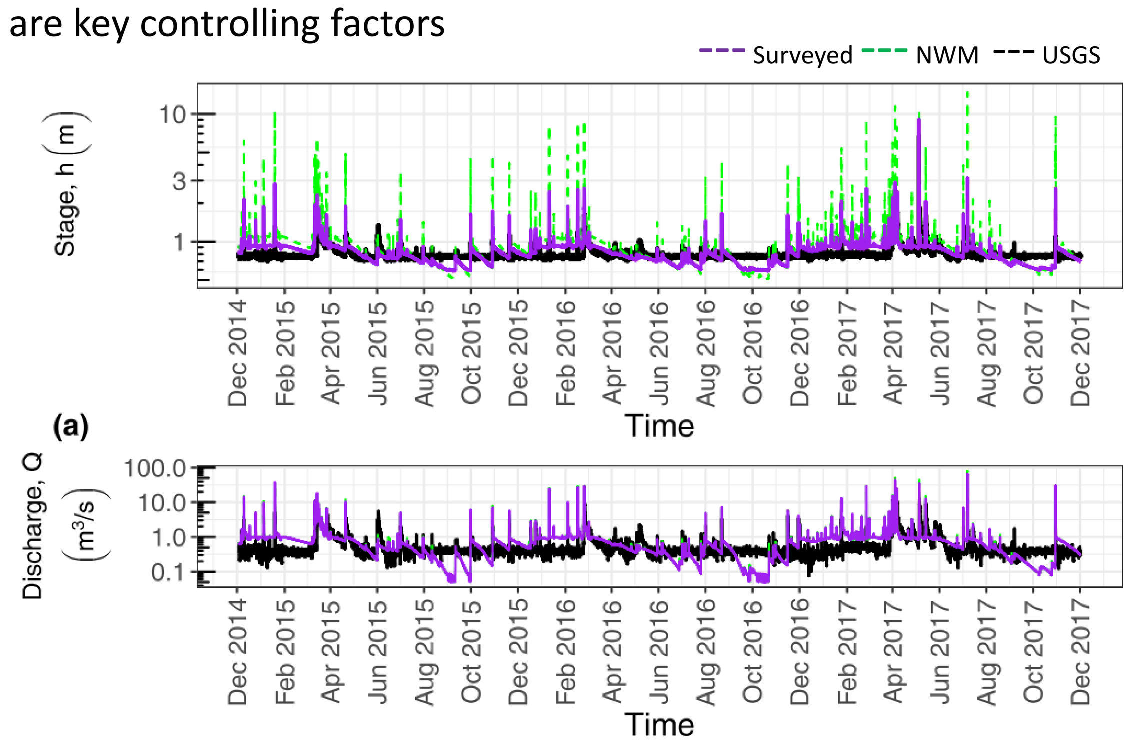


Fig 2: Drainage Area vs. Bankfull Discharge, Cross-sectional Area, Width, and Mean Depth for the Contiguous U.S.

Motivation

For Streamflow Predictions and Flood Inundation Modeling (FIM), channel dimensions and riverbed roughness parameters are key controlling factors



Brackins, John, Nishani Moragoda, Azbina Rahman, Sagy Cohen, and Christopher Lowry. "The Role of Realistic Channel Geometry Representation in Hydrological Model Predictions." *JAWRA Journal of the American Water Resources Association* 57, no. 2 (2021): 222-240.

AGU FALL MEETING
Chicago, IL 12-16 December 2022

Dataset

TAURAAT: The Attribute table of USGS (Surface-water: Field measurements) for Rivers And Associated Tributaries

TAURAAT can be considered as an updated version of **HYDRoSWOT**-HYDRoacoustic dataset in support of Surface Water Oceanographic Topography. TAURAAT includes **10,050** site stations (out of 10,081 sites represented by HYDRoSWOT) and represents five important channel geometry and characteristics of streamflow (i.e., streamflow, stage, channel width, channel area, and channel velocity) collected from the USGS stream gaging station (Surface-water: Field measurements) network and includes **2,802,532** records of all different types of USGS field measurements methods(Table 1). The time span of the records starts from 1845-05-05 to 2022-10-24.

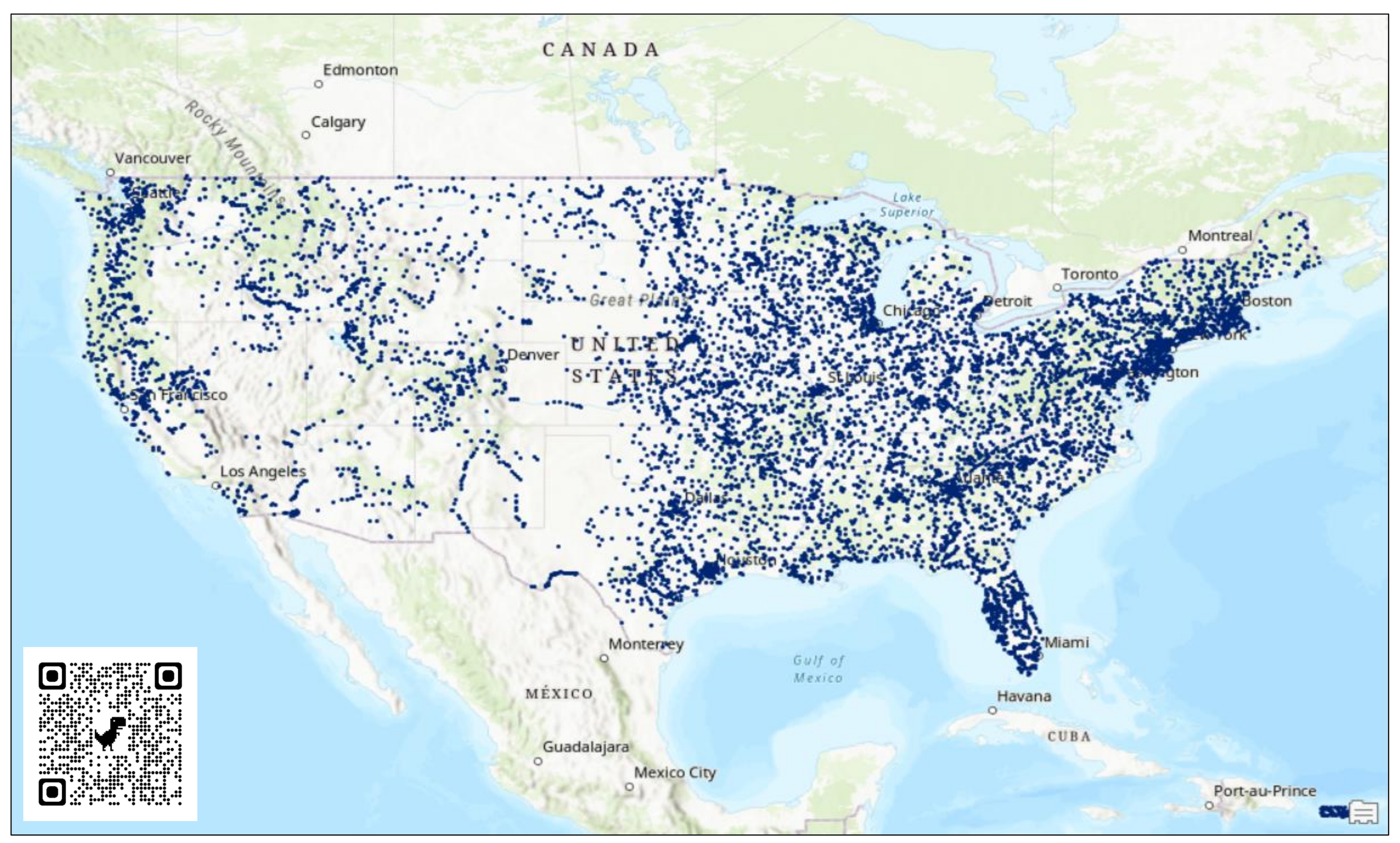


Fig 3: Map of the distribution of stream gages across the United States that comprise the HYDRoSWOT dataset

Dataset Analysis

After removing zero and missing values of three important feature columns including streamflow, river stage and channel width, **2,312,896** records remain. Moreover, to conduct significant statistics, **sites with more than 50 observations** are selected from valid records (out of no zero/missing records). Eventually, **2,279,530** observations (represented by **7,446** sites) are considered for following analyses.

Sites with Positive Discharge

2,199,163 observations (represented by **7,098** sites) include only positive values for discharge. The river stage verses discharge plot for some sample sites are as follows (blue, green and red lines represents 2.5, 50 and 97.5 percentile of data):

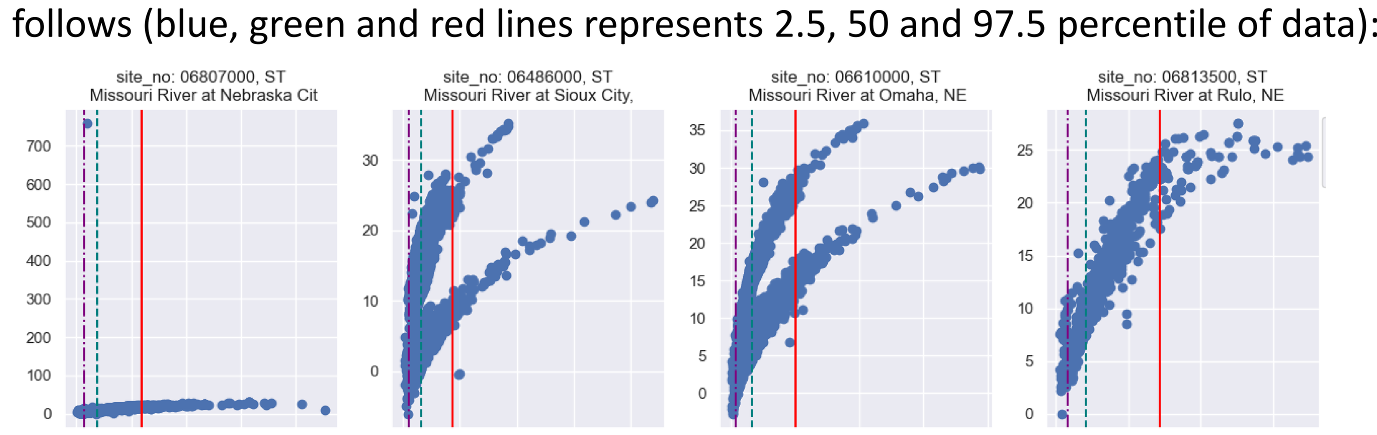


Fig 4: The stage-discharge relation for some randomly selected site stations with positive discharge values.

Sites with Negative Discharge

80,367 observations (represented by **348** sites) include negative for discharge.

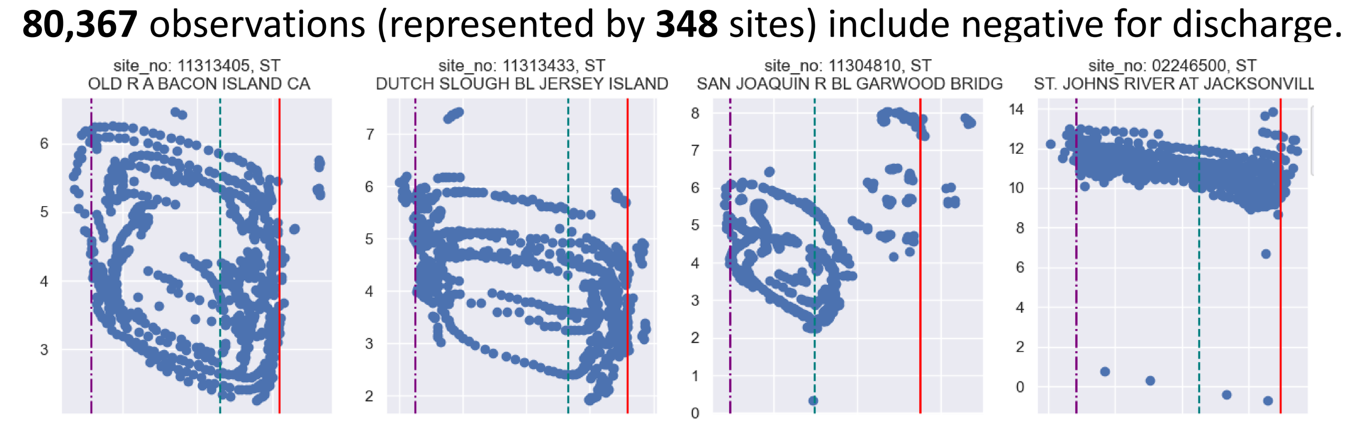


Fig 5: The stage-discharge relation for some randomly selected site stations with negative discharge values.

Methodology

Channel Geometry Analysis

In this project, instead of selecting one observation to represent the channel in a specific stage (e.g., active-channel, bankfull and overbank flow), the best fitted distribution to the important properties of the channel (discharge, stage and width) are estimated, and each site station is represented by the best fitted distribution parameters associated with channel geometry characteristics of that site. For instance, if it is assumed that normal distribution is the best fit for the observations of discharge in a specific USGS site station, two values (i.e., mean and std) will represent those observations of discharge. Through this way, each site station can be represented by one record for all observations that are measured for that site.

Dataset Analysis

Ten most common distributions (i.e., cauchy, chi2, expon, exponpow, gamma, lognorm, norm, powerlaw, rayleigh, uniform) were fitted to channel discharge/ width observations of 7,098 sites. The frequency of best fitted distribution is shown that "lognorm" is the best fit for channel discharge (for more than 50% of the site stations) and "cauchy" is the best fit for channel width (for about 4,000 site stations).

USGS 03599500 DUCK RIVER AT COLUMBIA, TN

USGS site number 03599500 is considered as an instance. The histogram, and Marginal (best fitted distribution) for channel discharge and width are shown in Figure 1. The Experimental Cumulative Distribution Function (ECDF) is used to show there is no "Binning Bias," for the histograms.

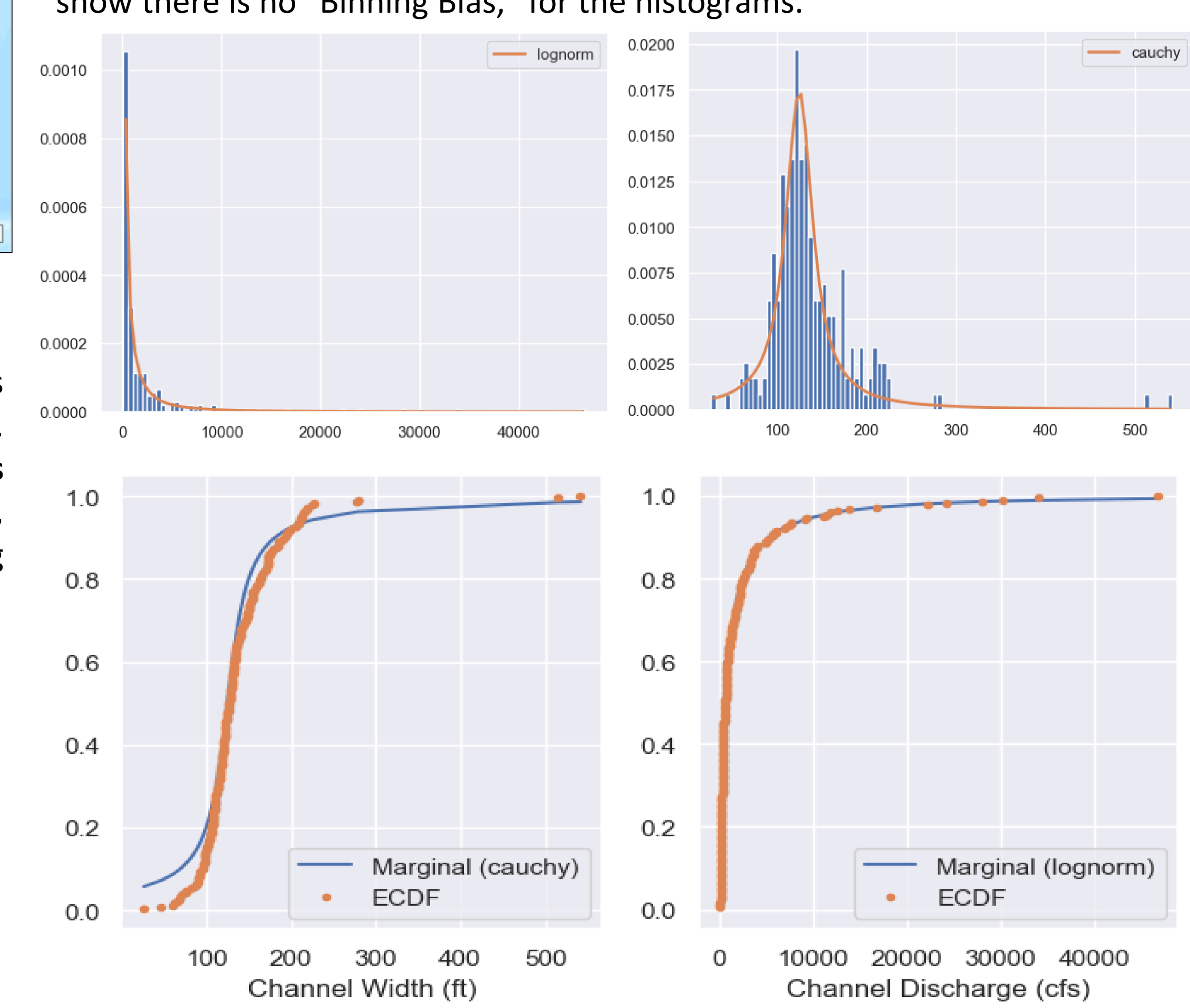


Fig 6: Histogram, ECDF and marginals (best fitted distribution) for channel discharge and width for site 03599500

Joint Probability Distribution

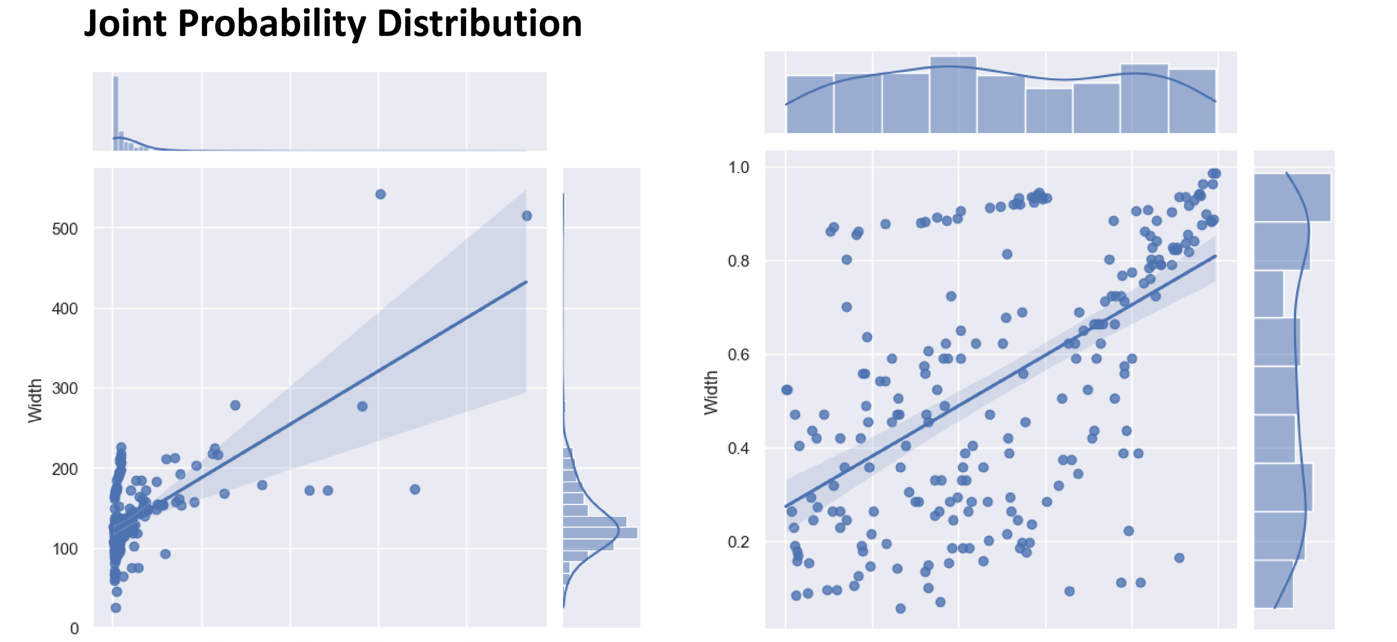


Fig 7: Bi-variate graph for real values and marginal CDF (channel discharge vs. width) for site 03599500

Methodology

Deep Learning

Adding aerial imageries to models will improve the performance of the models for predicting river width as airborne images have useful information about channel geometry.

- Geometry of bankfull and floodplain
- NAIP imagery includes information about vegetation
- NAIPs have high spatial resolution
- DEM bare ground topographic surface

CNNs encode certain properties of images into features which enhance predictive power for estimating the geometry of the river.

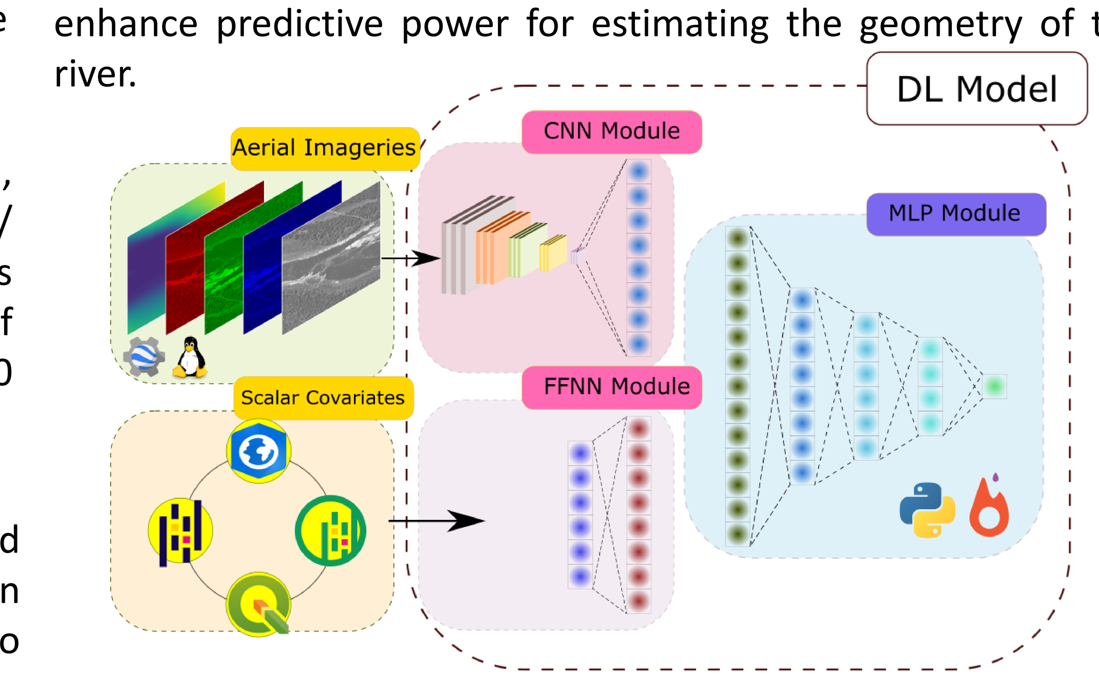


Fig 8: Overview of the model architecture

Results

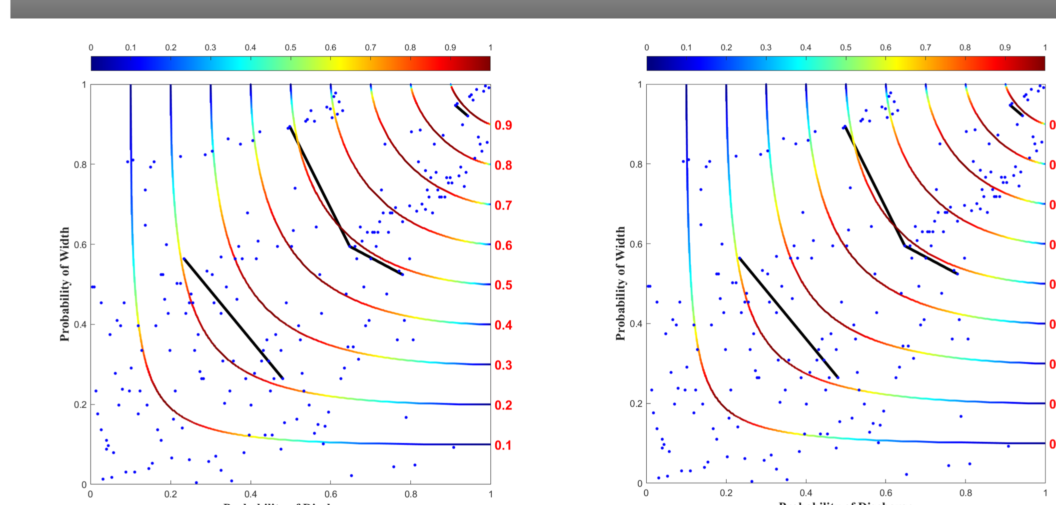


Fig 9: Copula probability space for Gaussian and T families

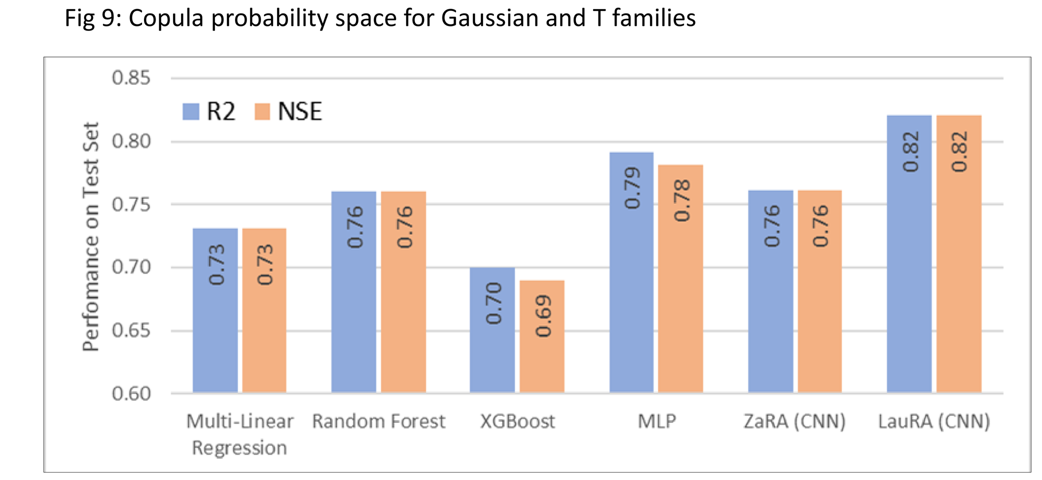


Fig 10: The summary the performance of different ML models.

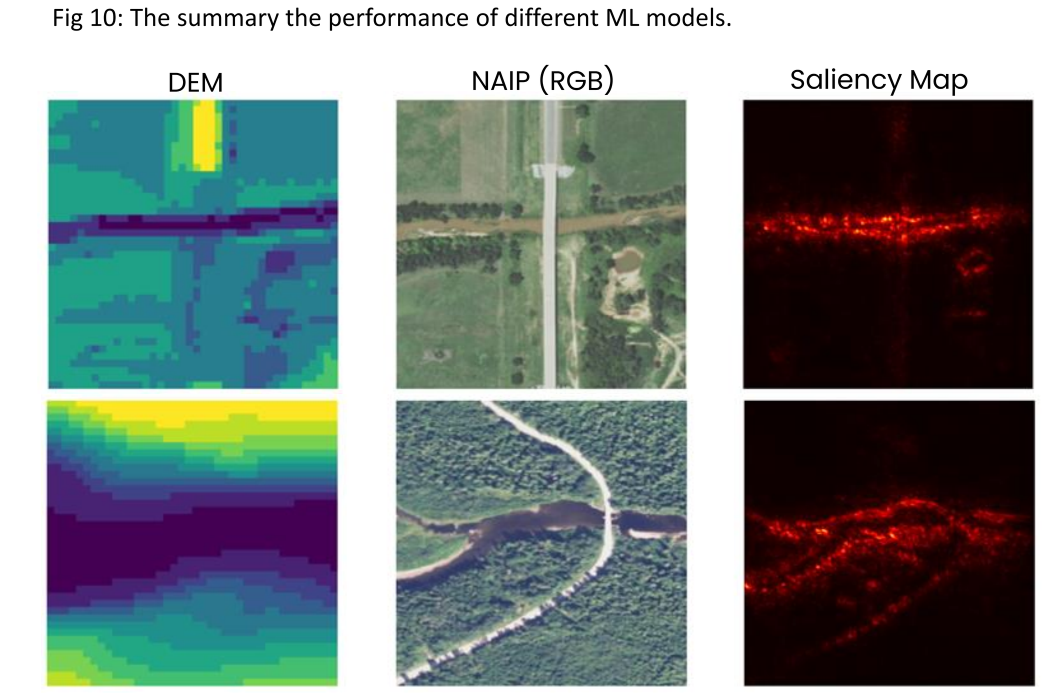


Fig 11: Visualization of trained model using guided backpropagation saliency