Version of Record: https://www.sciencedirect.com/science/article/pii/S1364815222002651
Manuscript_5a57767d3375dd9c0cc036b66f1813c0

Comparative Analysis of Performance and Mechanisms of Flood Inundation Map Generation using Height Above Nearest Drainage

Zhouyayan Li [a,b*], Felipe Quintero Duque[b], Trevor Grout[c,d], Bradford Bates[c,d], Ibrahim Demir [a,b]

[a] Dept. of Civil and Environmental Engineering, University of Iowa, Iowa City, Iowa, USA

[b] IIHR Hydroscience and Engineering, University of Iowa, Iowa City, Iowa, USA

[c] National Water Center, Tuscaloosa, Alabama, USA

[d] Lynker, Leesburg, Virginia, USA

* Corresponding Author, Email: zhouyayan-li@uiowa.edu

Abstract

The National Water Center (NWC) implemented Height Above Nearest Drainage (HAND) for nationwide flood mapping in the continental United States. Although having a large coverage and high accuracy, the implementation (NWCH) relies heavily on the NHDPlus dataset which limits its potential to handle user defined datasets. Comparison of the NWCH model accuracy and computational performance against the original HAND is missing in the literature. This study evaluated the flood maps generated using NWCH and a web-based implementation of the original HAND (WBH). An in-depth sensitivity analysis was conducted for WBH. Results suggest that WBH can generate comparable inundation extent with few inputs in regions where the water depths from the synthetic and catchment rating curves are consistent. Multi-depth approaches help resolve underestimations of WBH. This study demonstrated the original

HAND's efficacy in flood mapping and its potential for applications for fast predictions with acceptable accuracy with limited computational resources.

Keyword: Flood inundation mapping; Height Above Nearest Drainage; model comparison; model configuration; performance analysis.

1 Introduction

Humans have been fighting against floods for centuries (Di Baldassarre *et al.* 2017, Ghosh and Kar 2018, de Lange 2019, Blöschl *et al.* 2020). Different from many other natural hazards, flooding is a result of combination of natural and anthropogenic causes (Munoz *et al.* 2018, Bentivenga *et al.* 2020, Nicholls *et al.* 2021). Paved roads and poorly designed urban pipeline networks can disrupt the drainage process and exacerbate urban flooding (Lancia *et al.* 2020, Sun *et al.* 2021). Levee and dam breach and inappropriate reservoir operation during heavy precipitation and flooding events can bring unexpected inundation to unprepared communities and cause massive direct (Tadesse and Fröhle 2020, Yildirim and Demir 2021) and indirect losses (Psomiadis *et al.* 2021, Alabbad *et al.* 2022). Rapid landscape transformation in both rural and urban areas, combined with climate change is quickly obsoleting our previous efforts of understanding, identifying (Haltas *et al.* 2021), and mapping flood events over the last few decades (Leitner *et al.* 2020, S Chegwidden *et al.* 2020, Abdrabo *et al.* 2022).

One possible solution to cope with the fast-changing pace of the physical world is to conduct flood modeling with forecasts of factors that affect the magnitude and pattern of floods as input, such as new precipitation dataset collected in real-time (Seo *et al.* 2019), evaporation predictions (Le and Bae, 2020) obtained from general circulation models (GCMs), and future projection of landcover schemes (Leitner *et al.* 2020, Janizadeh *et al.* 2021). Another viable option is to create and use fast modeling frameworks that are data-parsimonious, robust, and computationally efficient based on web technologies (Sit *et al.* 2019, 2021, Xu *et al.* 2019, Agliamzanov *et al.* 2020). Web applications are light-weighted and platform-independent compared to stand-alone software or plugins and thus are ideal for fast and flexible hydro-modeling and hydro-informatics applications.

Although being the most widely studied and used approach in the past few decades, hydrodynamic models are less likely to be the ideal future flood mapping framework to meet the rapidly growing need of being able to aid in fast response to and preparation for floods. In those usage scenarios, accuracy is no longer the only factor to consider as the weights of many other factors, especially speed, are on the rise. Currently, there are three main factors that prevent hydrodynamic models from fast applications to new locations or scenarios. First is the large data requirements of those models, such as channel profile, initial and side flows, and boundary conditions (Savage *et al.* 2016, Teng *et al.* 2017) versus the fact that regions with sufficient data that fulfill models' requirements are still minority worldwide (Ebert-Uphoff *et al.* 2017), as many of those data come from on-site measurements or land surveys and often cannot be easily obtained or updated in a timely manner (Musser *et al.* 2016, McGrath *et al.* 2018). Lacking robustness is another shortcoming of many existing physics-based modeling frameworks. For example, many hydrodynamic computations are extremely sensitive to Manning's roughness (Terezinha *et al.* 2017) which is an empirical coefficient for which the initial value is usually obtained from tables, field surveys, and empirical formulas (Papaioannou *et al.* 2017). Calibration is always required to obtain the ideal roughness value that yields the best simulation result for study regions (Papaioannou *et al.* 2017, Garrote *et al.* 2021). As Manning's n is mostly governed by the physical characteristics of the channel (Nohani 2019), it is sensitive to common channel alterations, such as vegetation growth and dredging. Lacking robustness means the model deployed to a new location needs substantial adjustments and calibrations before it can reflect the physical condition accurately and therefore will damage the efficiency of the entire model deployment. The final criterion to consider when selecting a model is computing costs, especially for flood forecasting and response applications. It is sometimes preferable to use models that produce results fast with acceptable accuracy rather than models that produce precise results but take days or even weeks to run. For instance, in October 2015, South Carolina was hit by record-breaking precipitation which further triggered a flooding event with an estimated 0.1 percent annual chance and tremendous economic loss and infrastructure damage (Mizzell *et al.* 2017, Brandt *et al.* 2019). However, the official flood inundation maps done by USGS were released four months later (Musser *et al.* 2016, Li *et al.* 2018). It is obvious that such time-consuming simulations will have very limited benefits to flood forecasts and quick response applications.

Over the last few decades, simplified-conceptual models have grown rapidly in flood inundation mapping. Due to their reduced complexity of model structures, data, and computational requirements, these models ensure a better balance between accuracy and speed. Many of these models are topography-based techniques that require a digital elevation model (DEM) or digital terrain model (DTM) as the primary input and only have a few parameters to adjust (Nardi *et al.* 2019, Baldassarre *et al.* 2020) as they do not generally solve hydraulic equations or require initial and boundary conditions for calculation. These models can potentially benefit from DEM products of hyper-resolution algorithms (Demiray *et al.* 2021) in the future.

Among all the simplified models, the Height Above Nearest Drainage (HAND) has been widely used for flood inundation extent prediction (Afshari *et al.* 2018, McGrath *et al.* 2018, Speckhann *et al.* 2018, Godbout *et al.* 2019, Jafarzadegan and Merwade 2019) because it produces comparable results to those produced by more complex modeling frameworks, such as the U.S. Army Corps of Engineers Hydrologic Engineering Center River Analysis System (HEC-RAS) (Afshari *et al.* 2018, Zheng *et al.* 2018, Li *et al.* 2022). A basic HAND-based inundation extent map is created by a pixel-by-pixel comparison between a particular water depth with the

HAND value, which is the elevation difference between the present pixel and the pixel in drainage networks to which it drains (Nobre *et al.* 2011). For a detailed introduction to the HAND model, see section 3.1. The HAND methodology has been adopted for many other research purposes such as uncertainty analysis (Jafarzadegan and Merwade 2019, Michael Johnson *et al.* 2019) and reach-averaged rating curve generation (Zheng *et al.* 2018). Moreover, rather than simply applying the framework for analysis and comparison, several studies have been conducted to improve the framework's accuracy (Zheng *et al.* 2018, Shastry *et al.* 2019) and computational efficiency (Liu *et al.* 2018).

Currently, there is a substantial number of research studies that compare HAND with other flood modeling approaches, such as FLO-2D model (Komolafe *et al.* 2021), 1D/2D shallow water equations (Hocini *et al.* 2020), multivariant linear regression algorithm (Lababidi 2021), AutoRoute and HEC-RAS 2D (Afshari *et al.* 2018), and Planar plane and Inclined plane (McGrath *et al.* 2018). In the United States, the National Water Center (NWC) of the National Oceanic and Atmospheric Administration (NOAA) has developed a version of HAND to support national flood forecasting products. NOAA applies streamflow estimates from the National Water Model (NWM) to a nationwide HAND grid to generate national inundation maps by converting those stream flows to water depth to compare with the HAND values at the catchment level (Maidment, 2017), It is worth noting that the National Water Center’s HAND approach (herein referred to as NWCH) has some major implementation deviations from the original HAND methodology in order to utilize data from the NHDPlus dataset. For example, in NWCH, the stream network starts from pre-defined channel head sources and it is forced to align with NHDPlus streams (Zheng *et al.* 2018). NWCH also used D_∞ instead of D_8 flow model to compute the vertical distance (HAND value) of any hillslope pixels to the stream (Zheng *et al.*

2018). Last but not least, the flood extent of NWCH is generated in about 2.7 million reaches in the continental United States with the inundation condition of each pixel inside any given reach controlled by a distinct rating curve of that reach (Maidment 2017, Michael Johnson *et al.* 2019). Readers will find more details about the original HAND methodology and NWCH in sections 3.1 and 3.2. Although having some customized adaptations as discussed, many studies with a study area inside the US utilized the NWCH framework because of the accuracy and large coverage of the products for secondary analyses such as inundation mapping error assessment (Godbout *et al.* 2019), river channel geometry and rating curve estimation (Zheng *et al.* 2018), reach-level comparison against remotely sensed inundation maps (Michael Johnson *et al.* 2019). By contrast, the original HAND framework failed to receive as much research interest in the United States, which, therefore, necessitates a comparison between the NWCH and the original HAND approach. The reason for such a comparison is multi-folded. First, the NWCH approach does not transplant easily to other areas or countries due to the data availability issue of its dependencies. In addition, as the drainage network of NWCH is determined by pre-defined channel heads and stream networks, it does not adapt well to frequent changes in geo-morphic factors, such as elevation changes due to land cover change, urbanization, and dredging. Currently, the NWCH HAND layer is not designed to be updated frequently (Liu *et al.* 2016) and is thus not able to incorporate constant changes in the abovementioned physical factors. Furthermore, it is not easy for NWCH to keep up with the pace at which new data emerge, such as the crowdsourced water depth observations (McDougall and Temple-Watts 2012, Smith *et al.* 2017) and newly introduced high-resolution satellite-derived DEM products (Huber *et al.* 2021, Tapete *et al.* 2021). On the contrary, the original HAND is more flexible and adapts to new data much more easily. However, so far, the comparisons between different implementations of the HAND methodology are not well documented in the literature.

Therefore, the first objective of this study is to compare the flood maps generated using the NWCH and the original HAND method and to evaluate the NWCH flood maps using those generated with HEC-RAS and approved by FEMA. The NWCH is selected because it is, so far, the only implementation of HAND done by a national agency or institute and has been widely used and well-documented in the literature.

We then selected a client-side web-based inundation mapping system implemented by Hu and Demir (2021) (herein referred to as WBH) as the other comparison target representing the original HAND procedure.

It is reasonable that the original HAND may generate less accurate flood inundation extents in some cases because it is less demanding in data and computational resources. Therefore, the second objective of this study is to try to investigate the mechanism of the original HAND model through a parameter sensitivity analysis, analyze when this simpler method fails to bring satisfying results, and investigate how its performance can be improved without considerably increasing model complexity and data requirement.

Finally, the study summarizes the results with an in-depth analysis of the original HAND's limitations, a detailed discussion based on model sensitivity analysis, and the computational efficiency compared with NWCH implementation with the hope of further extending current understanding of the HAND model and providing results that could help local communities, stakeholders, and decision-makers (Ewing and Demir 2021) with implementing their own flood mapping applications based on HAND.

2 Study Area and Data Collection

The sub-watersheds that encompass the cities of Clarksville and Rock Valley, Iowa are selected as the area of study. The region for Clarksville study consists of two HUC 12 sub-watersheds, #070802020704 and #070802020701, that together cover 145 km^2 and have a total stream length of 112 km. The study region for Rock Valley is included in the HUC 12 sub-watershed #101720240804 with a drainage area of 105 km^2 and a total stream length of 66 km. The study area is depicted in Figure 1. These two study areas will be referred to as Clarksville and Rock Valley for the sake of simplicity. This simply indicates that the study areas are located near these two cities and does not imply that we are studying urban flooding in this paper. The sub-watersheds are chosen considering data availability and computational efficiency.

2.1 Reference Flood Inundation Maps and Comparison Scope

The Iowa Flood Center's (IFC) statewide floodplain mapping effort (Gilles et al. 2012) provided inundation extent and streamflow predictions using HEC-RAS modeling for areas of interest in this study and those flood maps will be used as reference flood maps for validation. The major reason for choosing the HEC-RAS modeling results as the reference is two-folded. First, those are relatively recent products that were generated after the 2008 Iowa floods compared to some other widely used reference sources such as Federal Emergency Management Agency (FEMA) flood risk maps. More importantly, those simulations provide flood maps with high and consistent quality that are generated with FEMA guidelines for the state of Iowa and therefore are more trustworthy. Given the above reasons, we find it unnecessary to create reference maps on our own or to consider other reference sources. The reference maps consist of a collection of inundation extent maps corresponding to a series of stage values separated by 0.5 foot.

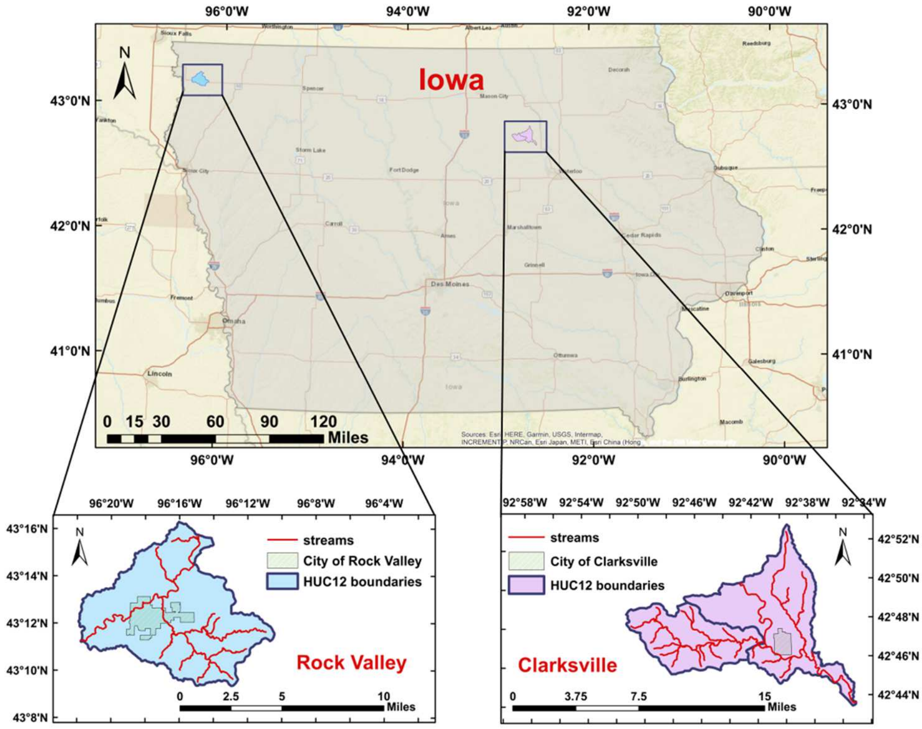


Figure 1. Location of the combined sub-watersheds #070802020704 and #070802020701, which contain the City of Clarksville, and the sub-watershed #101720240804, which contains the City of Rock Valley, Iowa

The NWCH inundation extent maps are the result of simulations run at the National Water Center with FIM 3.0.9.0 fr. Those maps are developed specifically for this study at the National Water Center. The WBH inundation maps are generated using the web-based flood inundation mapping system developed by Hu and Demir (2021) and enhanced by (Li and Demir 2022) (see subsection 2.3 for a brief introduction on the system). The NWCH inundation extent was provided within the two masks shown by the purple areas of Figure 2 and each of the two masks consists of several catchments pre-defined in NHDPlus dataset. The NHDPlus dataset contains about 2.7 million catchments for the continental United States. They each averages a surface area of 3 km^2 and a length of 2 km and is traversed by a single flow line (Maidment 2017). Those catchments are the smallest units that the simulation of NWCH runs on. Figure 2 shows the actual analysis (calculation) scope in green for WBH and evaluation (masks) areas in purple for Rock Valley and Clarksville.

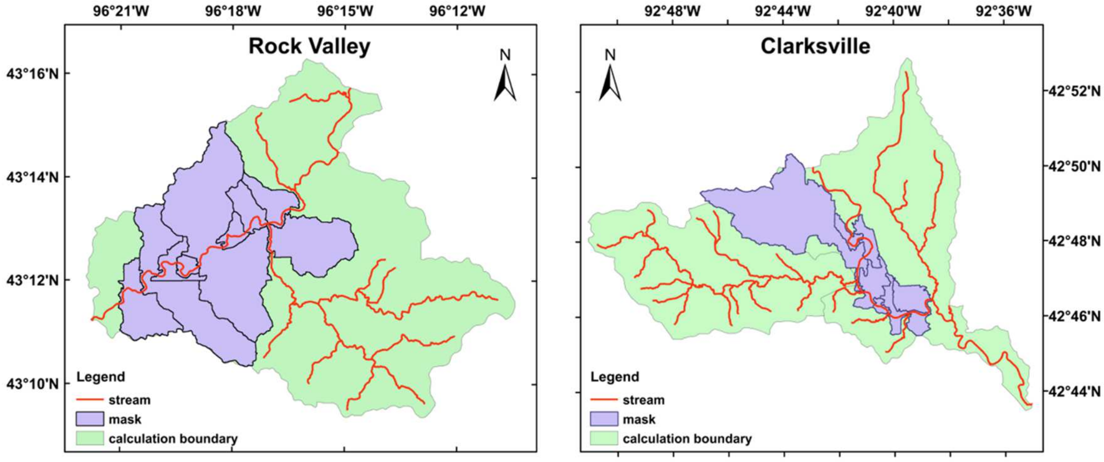


Figure 2. Location of catchments in which we compare flood inundation extent from NWCH, WBH, and the reference maps for Rock Valley (a) and for Clarksville (b)

There are two primary reasons for using a restricted area for comparison. The first is that the NWCH flood maps are obtained by putting together flood extents calculated in each of those small catchments. In this study, the catchments shown in Figure 2 were selected to represent the inundation condition of the study areas. Another reason is that detailed reference flood extent maps are only available in the vicinity of some Iowa cities. Therefore, using those two masks helps confine our analysis to areas surrounding the two study cities. We will investigate the 50-, 100-, and 500-year flooding scenarios in this study. The area and stream length of each catchment are listed in Table 1.

Table 1. Summary of terrain characteristics of the catchments in Rock Valley and Clarksville

Rock Valley				Clarksville			
HydroID	Area (km^2)	Median Thalweg Elevation (m)	length (m)	HydroID	Area (km^2)	Median Thalweg Elevation (m)	length (m)
21110025	7.75	370.86	1434	17450065	0.92	281.39	1461
21110023	0.52	372.28	1437	17450066	1.54	280.51	1465

21110021	0.76	373.78	1436	17450067	18.50	280.14	1414
21110022	15.29	372.87	1438	17450060	0.13	279.94	634
21110024	1.13	371.37	1437	17450058	2.44	279.73	1013
21110029	1.74	375.81	1350	17450059	0.41	279.50	996
21110026	3.41	369.97	1440	17450057	1.38	278.37	1126
21110020	1.46	374.77	1442	17450055	1.01	278.30	1073
21110866	5.10	377.15	1272	17450056	2.25	278.06	1046

2.2 NWCH Inundation Mapping and Data Requirements

The National Water Center leverages NHDPlus datasets to produce NWCH version 3.0 technique (NOAA 2021) to fulfill its directive to provide national inundation predictions. The datasets required for this task include HAND grids (full-resolution and mainstem configurations), synthetic rating curves (full-resolution and mainstem configurations), model network cross-walking information (full-resolution and mainstem configurations), a representation of the HAND-derived catchments (full resolution and mainstem configurations), and catchment-specific flow values. The flow values can be supplied to the model by a river discharge model, such as the National Water Model, or from historic observations or other sources (i.e., crowdsourcing).

NWCH 3.0 uses two configurations, full-resolution and mainstem. The full-resolution configuration’s stream network resembles that of the NWM. The mainstem configuration resembles only the stream segments that are downstream of an official Advanced Hydrologic Forecasting Service (AHPS) forecast site. NWCH 3.0 uses a mainstem configuration to better represent inundation in higher-order streams, whereas the full-resolution configuration is subject to underprediction of inundation extent in higher-order streams, primarily because of artificial restriction of inundation by catchment boundaries. The NWCH implementation can be found at the GitHub repository of Flood Inundation Mapping for U.S. National Water Model (https://github.com/NOAA-OWP/inundation-mapping).

2.3 WBH Inundation Mapping and Data Requirements

The WBH system can provide on-demand inundation predictions and hydro-spatial analysis products utilizing both pre-stored and user-supplied datasets. With the necessary dataset, the system can work with a variety of different calculation methods, perform result comparison and hydro-spatial analysis, and perform flood mitigation analysis in any study region provided by the user (Li and Demir, 2022). Therefore, it enables the performance testing of models with varied configurations for this study.

The amount of data required for the system to generate inundation maps depends on the products desired by users and the corresponding calculation procedures. Data and information required in this study include NED DEM and river networks from NHDPlus dataset, LiDAR-based DEM, synthetic rating curves derived from HEC-RAS simulations for the two study areas, reach-averaged rating curves and reach information (such as location, area, and stream length as specified in Table 1), and location and relevant information of the closest USGS gauges for the two study areas (#06483500 for Rock Valley and #05462000 for Clarksville).

To compare the performance of NWCH with the WBH, the DEM raster from the NHDPlus is clipped, translated to meters, and resampled to 10m from the original 30m resolution. For sensitivity analysis with various model setups, the 1-m LiDAR-based DEM was resampled to 5-m resolution. The 10m resolution was chosen to maintain consistency with the NWCH data, and the 5m resolution was chosen to balance computing efficiency and accuracy for sensitivity analysis.

It is worth noting that some of the items stated above are required simply because we're comparing our results to those from NWCH and therefore need to maintain data consistency.

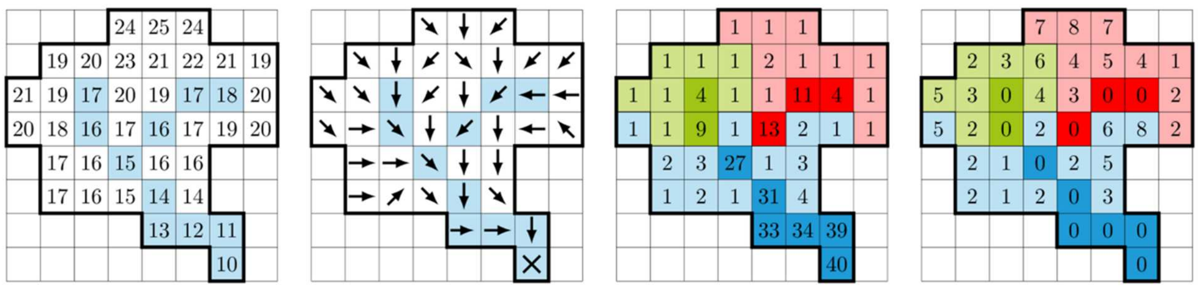
Indeed, the WBH system is far more adaptable in terms of data requirements. For example, in this study, we need the rating curves to determine the stage for a specific discharge, but we can also feed the system with water depth measurements or observations. Similarly, while the river network and information about USGS gauges aid in deciding the placement of outlet pixels in this study, the location of outlet pixels can be completely custom without restrictions.

3 Methodology

3.1 HAND Model

Height Above Nearest Drainage (HAND) model is a quantitative terrain descriptor initially introduced by (Rennó *et al.* 2008). HAND values are the differences in elevation between each pixel on hillslope and the nearest point in the river network that drains it. Numerous studies have established that the HAND model accurately represents the soil water environment (Nobre *et al.* 2011). Computing the HAND value starts from removing spurious depressions and flats from the raw DEM to make it hydrologically coherent (Rennó *et al.* 2008, Nobre *et al.* 2011). Then, the flow direction of each pixel on the DEM is calculated using one of the widely adopted algorithms, e.g., D_8 (Mark 1984), or D_∞ (Tarboton 1997).

Next, we calculate the accumulating area of each pixel by taking the total number of upstream pixels the current one drains. By comparing the accumulating values with a predefined drainage threshold, we separate drainage points (dark-color grid cells in Figure 3c) from non-drainage ones. Then, we divide all non-drainage pixels into sub-drainage areas based on which drainage point they drain to as shown in Figure 3c. For instance, all light green and light red pixels in Figure 3c are all non-drainage pixels, but they are colored differently because they flow to different stream pixels (dark red and dark green) as indicated by the flow direction of each pixel in Figure 3b. Finally, the HAND value of each non-drainage pixel is obtained by subtracting the elevation of the nearest drainage pixel from its original elevation, which is also called elevation normalization. The color division shown in Figures 3c and 3d just determines to which drainage pixel the elevation of non-drainage pixels should be normalized and serves no other purpose. For example, the HAND value of the left-most light green pixel is 5 because its HAND value is determined by subtracting the elevation of the darker green pixel (elevation 16) one row under it from its original elevation (21). That dark green pixel is the nearest drainage pixel of the left-most light green according to flow direction. After each pixel got its HAND value, it is no longer necessary to distinguish between each other. For instance, the bottom-most blue pixel drains all pixels above it, but it has no impact on HAND values of the green and red pixels as those are controlled by upstream drainage points and are determined before the flow converges to any blue pixels. The HAND values for drainage pixels are set to zero meaning they do not have drainage potential as they are the lowest points within the drainage network. The final product or HAND model is a matrix of HAND values of the same numbers of column and row as the DEM processed. Figure 3 shows a graphic representation of the HAND procedure.



(a) DEM (b) Drainage direction (c) Drainage area and sub-catchments (d) HAND value

Figure 3. Major calculation steps of the HAND model. Figure reproduced from Rebolho et al. (2018)

3.2 NWCH Inundation Mapping

In the NWCH, each river segment is encoded with a “feature_id” and a discharge value. Because the NWCH-derived hydrologic network differs from the National Water Model network, model crosswalk information is needed to associate NWM discharge values with the NHDPlus catchments. After crosswalking the NWM discharge values to the NWCH catchments, catchment-specific synthetic rating curves are used to interpolate stage heights from the discharge values on a catchment-by-catchment basis. This interpolation results in a spatial array where values are encoded according to the catchment-specific interpolated stages. The HAND grid is then subtracted from this spatial array to derive inundation depths.

This process is performed for both the NWCH full-resolution and the mainstem configurations to generate two depth grids for the same area. A follow-up procedure is performed to mosaic the full-resolution and mainstem grids, prioritizing the maximum pixel value, i.e., maximum depth, when the same pixel location has a value provided by both configurations. Depending on the use-cases for the inundation information, the final mosaicked depth grid may be reclassified to a binary wet/dry inundation map and converted to a polygon.

For the purpose of this analysis, only the NWCH full-resolution configuration was used, i.e., not the full-resolution and mainstem composite inundation map.

3.3 WBH Inundation Mapping

WBH compares the HAND value of each pixel directly with depth values to decide the inundation extent. The stages may come from measurements at hydrometric stations, estimates obtained from rating curves, and crowd-sourced observations collected during flooding events on social media. In this study, the stage estimates for the entire study area and each catchment are obtained from the synthetic rating curve by HEC-RAS and the reach-averaged rating curves produced along with NWCH's generation of the HAND layer (Michael Johnson *et al.* 2019). .

We developed four methods for calculating the inundation extents of the WBH including a single depth (D_S), and three multi-depth approaches, namely area-weighted depth (D_A), stream-length-weighted depth (D_L), and the local depth (D_{Local}) approaches. As implied by the name, D_S applies a single water-depth value calculated from the stage estimate using HEC-RAS synthetic rating curve at the USGS gauge location to the entire study area. Whereas D_A, D_L, and D_{Local}

make use of a group of water depths derived from reach-averaged rating curves for each of the previously specified catchments summarized in Figure 2 and Table 1. The difference between these three multi-depth approaches lies in how we calculate these depths. For D_{Local}, water depths corresponding to stream flows of return periods of 50-, 100-, and 500-year in each catchment will be used without further processing, which thus allows each catchment in Figure 2 to have its own water depth to compare with HAND values. Whereas D_A and D_L take the average of those water depths weighted by catchment areas and stream lengths and then the average values will be shared among all pixels without differentiating which catchment they belong to. In this study, we use Eq. 1 to calculate the water depth at a specific location.

$$D = S + E_{ref} \pm F_{adj} - E_{target} \qquad \text{(Eq. 1)}$$

D and S refer to water depth and stage height, which are measured against the bottom of the river channel and a certain datum, respectively. The bottom elevation of the river channel can be different from the elevation of the datum for stage measurements. In those cases, the stage and water depth can be different. E_{ref} is the elevation of datum against which stage values are measured. E_{target} is the elevation of the location for which we calculate the water depths. F_{adj} is the elevation converting variable when the elevation datum of stages (normally NGVD29 for USGS gauges) is different from the DEM we used (NAVD88 in our case).

The two USGS gauges serve as the reference points for D_S. Because there are no USGS gauges in each small catchment, we adopted the median elevation of thalweg of each reach as the elevation of references for the multi-depth ones (D_A, D_L, and D_{Local}). These elevation values have the same vertical datum as other DEM pixels and thus do not need datum conversion.

3.4 Performance Comparison between NWCH and WBH

The WBH inundation extent maps for 50-, 100-, and 500-year flooding scenarios were created and compared to corresponding reference extents and the NWCH extents. Here, we only use D_S

for comparison to see if WBH can produce comparable results with the minimum data, namely, a DEM, a drainage threshold value without calibration, and a single water depth. The same 10-m DEM used by the NWCH is fed into the WBH, it then performs a series of automated processes to remove artificial pits and flats from the raw DEM and create a hydrologically coherent surface. The depth data was derived using Eq. 1 and the stage estimates from the synthetic rating curve. For the drainage threshold, we assumed that no additional information or guidelines were available and thus chose 4.0 km^2 as previous studies had shown its efficacy (Nobre *et al.* 2016).

3.5 Sensitivity Analysis of the Performance of WBH

Model configurations with various drainage thresholds and depth values are computed and evaluated. As stated in subsection 3.1, drainage pixels have a HAND value of 0, and they are the points to which the elevation of non-drainage pixels is normalized. By changing the threshold value, we can modify the numbers of drainage pixels and thus modify the simulated network. In this study, the threshold being tested starts from 1 percent of the study area and increases by 1% each time until the model performance stabilizes. For each flooding scenario investigated in this study, D_S, D_A, D_L, and D_{Local} are computed and applied along with each threshold value, resulting in $e \times n \times (t_1 + t_2)$ different model configurations, where *e* is the numbers of flooding scenarios, n is the number of depth calculating approaches, and t_1 and t_2 and the number of tested threshold values in Rock Valley and Clarksville, respectively.

3.6 Evaluating Model Performance

A two-by-two contingency Matrix (Provost 1998) was used to categorize any pixel on a map's simulated inundation conditions into one of four classes: True-Positive (TP) means the pixel is predicted inundation by the model and indicated inundation on the reference map; True-Negative (TN) means the pixel is predicted dry by both the model and the reference; False-Positive (FP) means the pixel is predicted inundated by the model but is dry on the reference map; and False-

Negative (FN) the pixel is predicted dry by the model but is actually inundated by the reference. The contingency matrix is depicted in Figure 4.

		Real Values	
		Positive	Negative
Predicted Values	Positive	TP	FP
	Negative	FN	TN

Figure 4. The contingency matrix to indicate the inundation condition of any pixel on predicted maps and the reference map. Figure reproduced from Li et al. (2022)

To further facilitate interpretation, we will compare the predicted extents with the reference visually and mathematically with the following indexes. Numerous indexes are available in the literature that can be used to evaluate model performance (Wilks 2011). To assess the agreement between the two maps, we used Proportion Correct, Bias, Hit Rate, Kappa value, and Fitness-statistic.

Proportion Correct (PC) has a value between 0 and 1, with 1 being the best. PC is a widely used index with the limitation of being unable to distinguish between FP and FN because they are treated equally in Eq. 2. It is calculated as follows:

$$PC = \frac{TP + TN}{TP + FN + FP + TN} \quad \text{(Eq. 2)}$$

Bias (B) is a positive value that with the best possible value of 1. B is not an accuracy measure (Wilks 2011) but indicates whether the scene is overestimated (B > 1) or underestimated (B < 1) in general. It is worth noting that B is not a measurement for model performance but an indicator of how many overestimations the model is made versus underestimations. In other words, B equals one does not necessarily mean the model achieves a high accuracy but just means the model made about the same amount of over- and underestimations. B is calculated as:

$$B = \frac{TP + FP}{TP + FN} \quad \text{(Eq. 3)}$$

Hit Rate (H) ranges between 0 and 1 with the best possible value of 1. H represents the ratio of inundated pixels on the reference maps that are captured by the predictions. H is also referred to as the Probability of Detection (POD), the true-positive fraction, and the sensitivity. (Wilks 2011). It is calculated as:

$$H = \frac{TP}{TP + FN} \quad \text{(Eq. 4)}$$

Kappa Value (K) can be negative, indicating that the prediction is worse than a random guess (Juurlink and Detsky 2005). The best value for K is 1. It is calculated as follows:

$$\mathrm{K} = \frac{N(TP + TN) - \big((TP + FP) \times (TP + FN) + (FP + TN) \times (FN + TN)\big)}{N^2 - \big((TP + FP) \times (TP + FN) + (FP + TN) \times (FN + TN)\big)} \quad \text{(Eq. 5)}$$

Fitness Statistics (F), also known as Critical Success Index (CSI) (Wilks 2011), ranges between 0 to 1 with the best possible value of 1. It is calculated as:

$$F = \frac{TP}{TP + FN + FP} \quad \text{(Eq. 6)}$$

K and F complement each other. K focuses more on the dry pixels and are prone to bias when there are much more correctly predicted dry pixels than correctly predicted flooded pixels (Afshari *et al.* 2018). Whereas F stresses more on the consistency of the flooded pixels on both maps.

4 Results

4.1 Comparison of NWCH and WBH Flood Inundation Predictions

Table 2 summarizes the comparison between NWCH and WBH for Clarksville and Rock Valley. The 4.0 km^2 drainage threshold for WBH is selected based on previous findings of related research (Nobre *et al.* 2016).

Table 2. Comparison summary between NWCH and WBH for Clarksville and Rock Valley

Study Area	DEM	WBH Drainage Threshold	Return Periods Involved and Corresponding Results	Performance Metrics
Rock Valley	10m NED DEM	4.0 km^2	100- and 500-year (Fig. 5 and Table 3)	PC[1], B[2], H[1], K[1], F[1]
Clarksville		4.0 km^2	100- and 500-year (Fig. 6 and Table 3)	

1: The higher the metrics value, the better the performance is.
2: The closer the value is to one, the more balanced the results are in terms of overestimation and underestimation.

The results of the NWCH and WBH compared to the reference inundation extent for Rock Valley and Clarksville in 100- and 500-year flooding scenarios are shown in Figures 5 and 6. Table 3 displays the evaluating indexes in comparison to the reference.

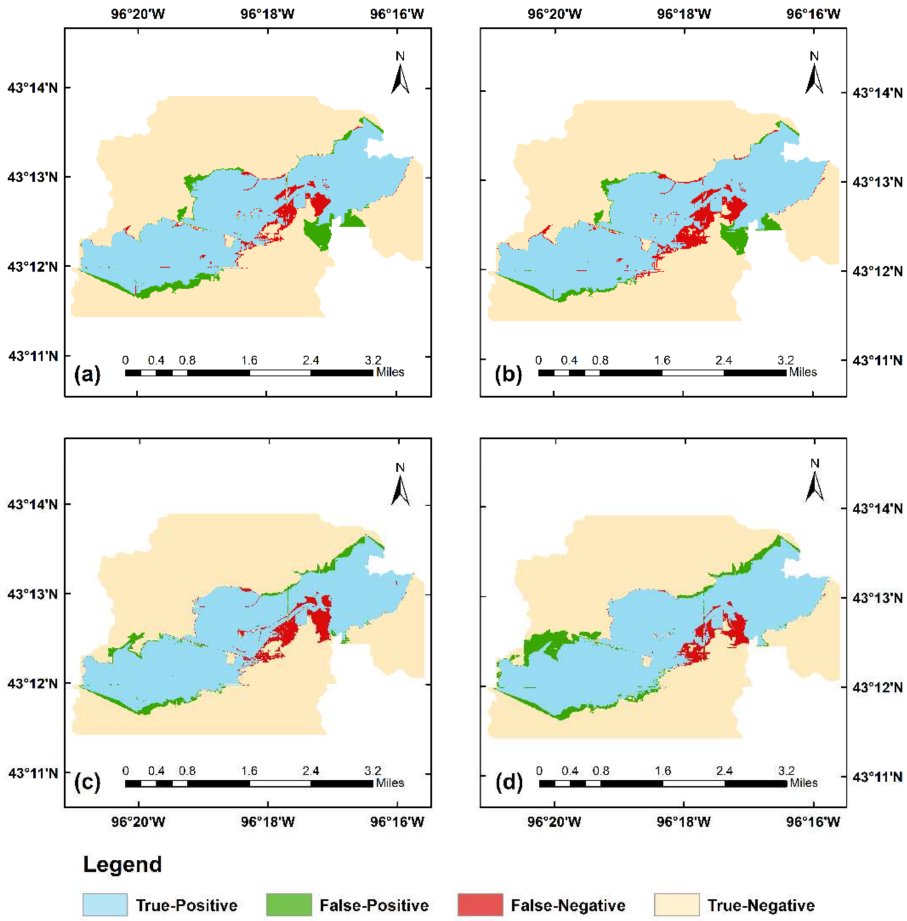


Figure 5. Predictions of inundation extent in Rock Valley compared to reference maps. (a) Evaluation of WBH in 100-year flood scenario, (b) evaluation of WBH in 500-year flood scenario, (c) evaluation of NWCH in 100-year flood scenario, (d) evaluation of NWCH in 500-year flood scenario

Comparing the False-Positive areas (in green) on 100-year predictions with those on 500-year predictions in Figure 5, the WBH generates slightly less overestimation for the 500-year flooding scenario around the lower-left and upper-right corners but more underestimation (in red) in the middle of the map. When compared to the WBH, the NWCH predicts slightly more overestimation along the upper border of the inundation extent in both flooding scenarios while producing less underestimation in the middle of the image in the 500-year one. According to the B index in Table 3, the predictions of the NWCH and WBH approaches for the 100-year flood are slightly underpredicted and overpredicted, respectively, and it is the opposite for the 500-year flood. Other indices show no significant difference in performance between the two modeling frameworks for both flooding events.

Table 3. Numbers of pixels classified as True-Positive, False-Negative, False-Positive, and True-Negative when compared to reference maps and the corresponding evaluating indexes.

	flooding event	Threshold (km^2)		PC	B	H	K	F
Rock Valley	100-yr event	-	NWCH	0.95	0.99	0.92	0.89	0.86
		$4km^2$	WBH	0.94	1.04	0.93	0.87	0.84
	500-yr event	-	NWCH	0.94	1.06	0.94	0.87	0.84
		$4km^2$	WBH	0.94	0.98	0.91	0.87	0.84
Clarksville	100-yr event	-	NWCH	0.93	1.06	0.97	0.87	0.88
		$4km^2$	WBH	0.83	0.68	0.67	0.66	0.66
	500-yr event	-	NWCH	0.95	1.02	0.96	0.89	0.91
		$4km^2$	WBH	0.87	0.77	0.76	0.74	0.75

In Clarksville, the WBH with D_S failed to capture as many inundated pixels on the reference map as the NWCH does for both flooding events, as shown in Figure 6 by comparing (a) with (c) and comparing (b) with (d). On both banks of the main channel's central portion, there is significant underestimation on the WBH map (a and b). NWCH extents, on the other hand, are generally more accurate while being slightly overestimated for both events. The performance of NWCH in the 500-year scenario is more balanced in terms of the amount of over and underestimation than in the 100-year scenario, as shown in Table 3.

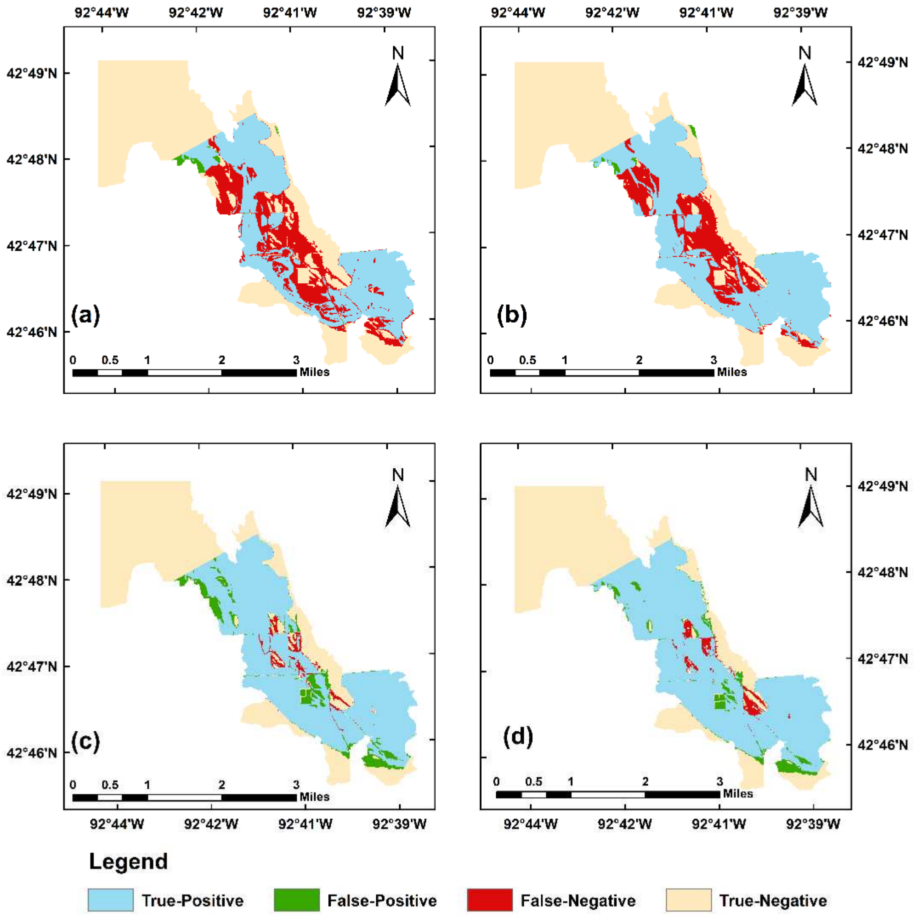


Figure 6. Predictions of inundation extent in Clarksville compared to reference maps. (a) Evaluation of WBH in 100-year flood scenario, (b) evaluation of WBH in 500-year flood scenario, (c) evaluation of NWCH in 100-year flood scenario, (d) evaluation of NWCH in 500-year flood scenario

4.2 Performance of WBH with Different Model Configurations

Table 4 summarizes the comparisons aimed at demonstrating the sensitivity of WBH toward its model parameters. We did not include the drainage threshold exceeding 17% of the study area in Clarksville, because the result became quite stable once the drainage area exceeds about 14% of the study area, which is also confirmed by performance curves of Rock Valley depicted in Figure 7.

Table 4. Comparison summary for model parameter sensitivity analysis

Study Area	DEM	Drainage Threshold Range (%)	Return Periods Involved and Corresponding Result Figures	Performance Metrics

Rock Valley	5m LiDAR DEM	1-31	50- (Fig. 7), 100-, 500-year (Fig. 8)	PC[1], B[2], H[1], K[1], F[1]
Clarksville		1-17	50- (Fig. 9), 100-, 500-year (Fig. 10)	

1: The higher the metrics value, the better the performance is.
2: The closer the value is to one, the more balanced the results are in terms of overestimation and underestimation.

For the 50-year flooding scenario in Rock Valley, Figure 7 depicts the WBH model performance among 31 threshold values and four water depth calculation approaches. As illustrated in Figure 7, the pattern of performance variation does not differ significantly between D_S, D_A, D_L, and D_{Local}. The B index is high when the threshold of 1% of total study area is used, indicating that the scene is overestimated. As the threshold increases, B and H decrease while PC, K, and F rise, indicating that the performance becomes balanced in terms of the number of overestimations and underestimations. Some indexes show abrupt changes at 2% and 4%. After exceeding the 8% threshold for the total calculation area, the performance becomes stable. For scenarios with D_S, D_A, and D_L approaches, the stable performance results are slightly overestimated (B value greater than 1) with PC and H values close to each other. Whereas the inundation extent is moderately underestimated for scenarios with D_{Local}.

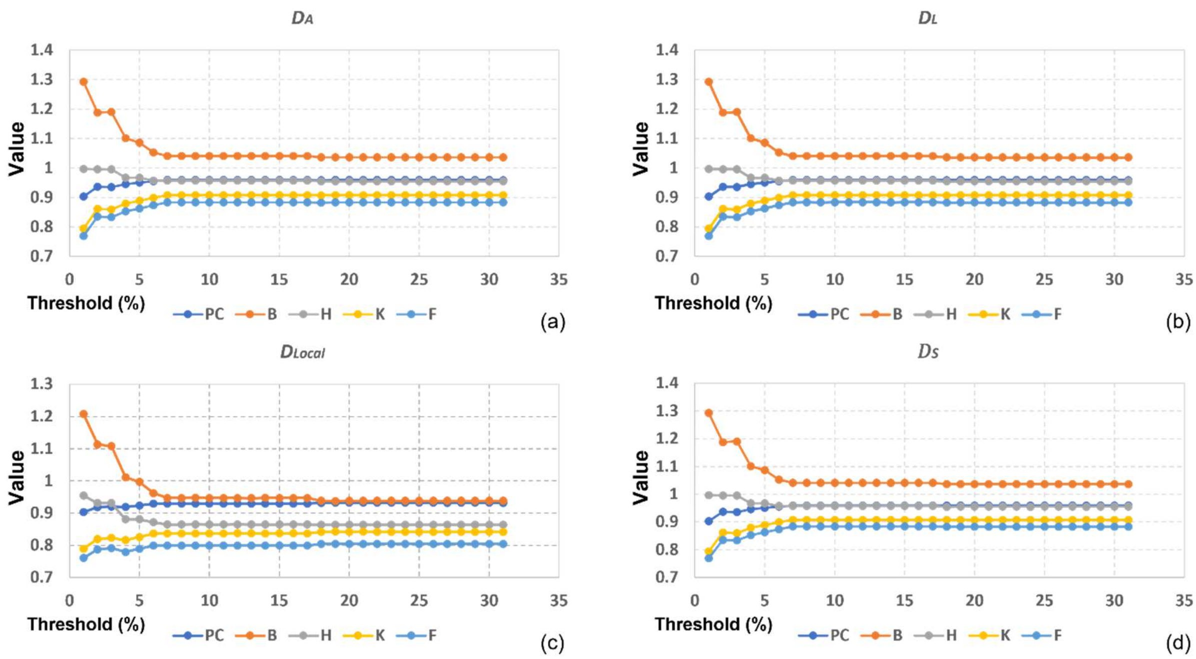


Figure 7. Comparing the model performance in Rock Valley in the 50-year flood scenario using D_A (a), D_L(b), D_{Local} (c), and D_S (d) approaches

Figure 8 depicts the bar charts with grouped indexes for the 8% of the study area after which the model performance becomes steady. As shown in Figure 8 (a), all the indexes are similar among the configurations utilizing D_A, D_L, and D_S except for H and B, for which D_L has slightly larger values The model configuration with D_{Local} produces the lowest performance for the 100-year flood among the three multi-depth approaches, with lower values for all indexes. The patterns for F, K, and PC are the same for the 500-year scenario presented in Figure 8 (b), as the value using D_A is the largest, followed by D_S and then D_{Local}, and finally the value using D_L. For B and H, the greatest value comes from the D_L configuration, followed by comparative values from D_A and D_{Local} cases, and finally the one using D_S. As B and H increase with the increase of positive predictions, a high value of B and H indicates the case using D_L generates more overestimation compared to the other three approaches.

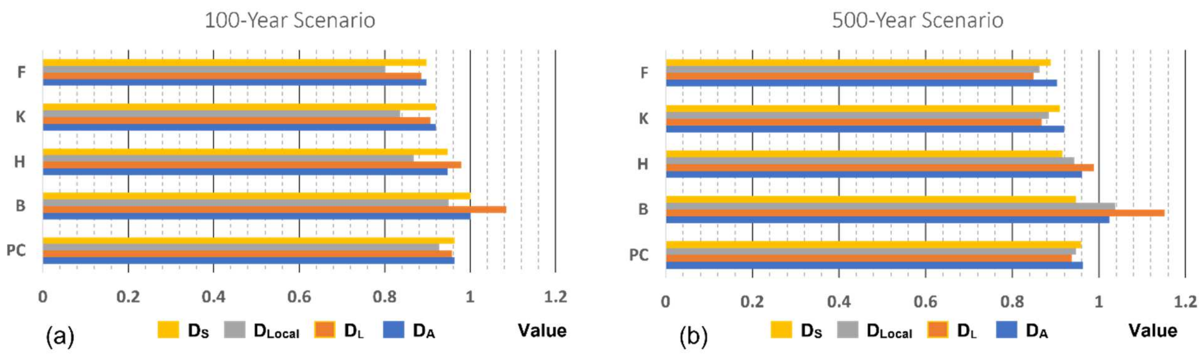


Figure 8. Comparing the stable performance of WBH at a fixed threshold of 8% for Rock Valley in 100-year (a) and 500-year (b) flood scenarios

Figure 9 shows model performance variation for Clarksville in the 50-year flooding event. There are two major changing points the model performance line among all thresholds tested. The two changes occur when the thresholds are about 6 and 12% of the study area, corresponding to the amount of 8.7 and 17.41 km^2. The 6 and 12% thresholds divide the performance curves into three stair-like ranges, namely below 6%, between 6 and 12%, and above 12%, which the performances are similar to each other.

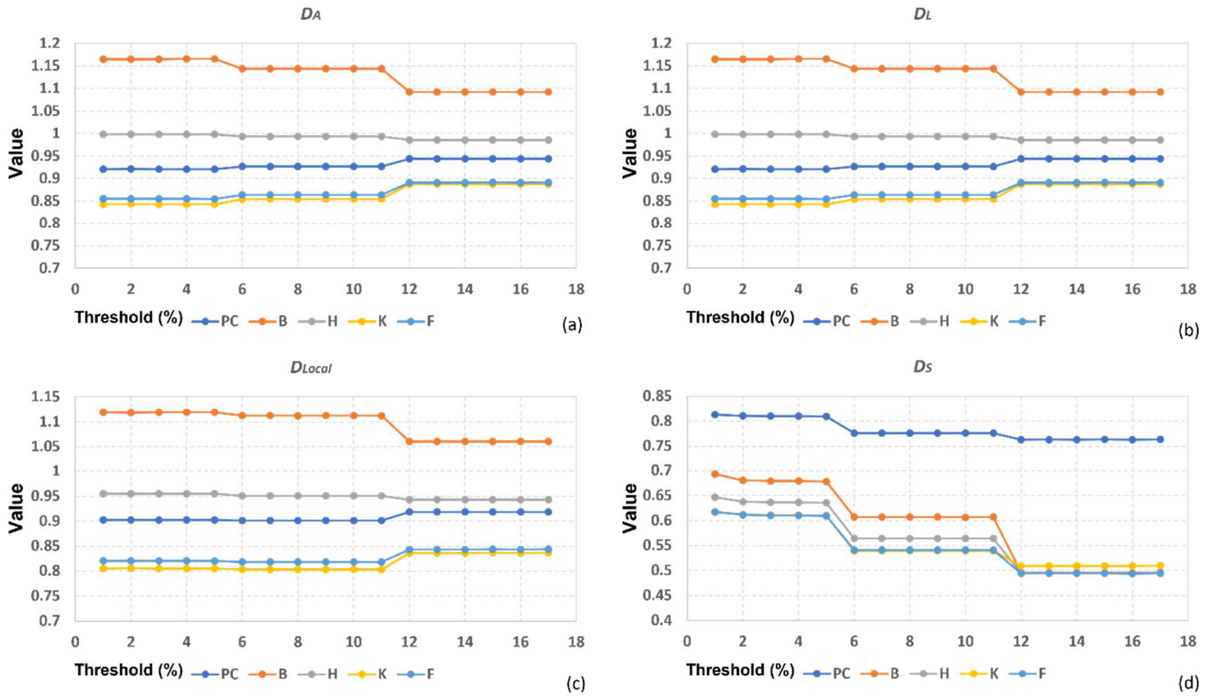


Figure 9. Model performance in Clarksville in the 50-year flood scenario using D_A (a), D_L(b), D_{Local} (c), and D_S (d) approaches

Figure 10 shows the summarizing bar charts of grouped indexes with the four water depths and a drainage threshold of 12% of the study area for the 100- and 500-year flooding scenarios in Clarksville. As shown by Figure 10 that the performance with D_A and D_L approaches are similar, followed by the configuration with D_{Local}, while D_S leads to the worst matching case for Clarksville in both flooding scenarios. Also, all configurations except for the ones with D_S show overestimated inundation extent in general, whereas the B index for the case with D_S is about 0.7 in both flooding scenarios indicating there are major underestimations in the prediction.

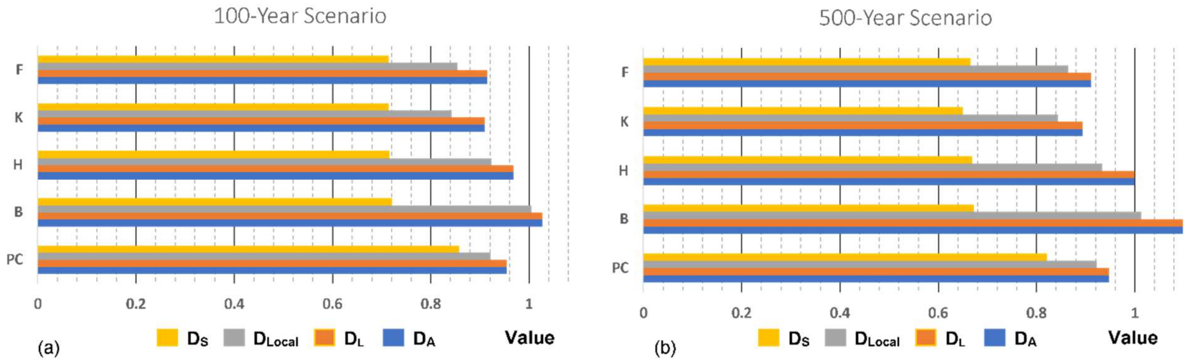


Figure 10. Comparing the stable performance of WBH at a fixed drainage threshold of 12% for Clarksville in 100-year (a) and 500-year (b) flood scenarios

Comparing Figure 7 results with Figure 9 we see that the significant performance changes for Rock Valley occur between the drainage thresholds of 1% to 5 %, whereas it occurs between 5% to 12 % for Clarksville. For both areas, the results of three multi-depth approaches are quite comparable, while D_{Local} is slightly worse but better balanced in the number of overestimations versus underestimations. The overall performance increases as the drainage threshold increases until it reaches the threshold of about 8 %. For Clarksville, D_S failed to catch as many inundated pixels indicated on the reference map compared to the other three techniques. The extent predicted by D_S is underestimated even with a 1% drainage threshold which corresponds to 1.45 km^2 which is the smallest threshold tested. Therefore, as the threshold increases, it brings more underestimations and lowers the model performance. The performance with the multi-depth approaches, by contrast, is more accurate and resistant to underestimation.

Figure 11 depicted the influence of changes in thresholds on flood extent for thresholds 1%, 4%, 8%, and 12% of the area of Rock Valley and Clarksville with D_A in the 50-year scenario.

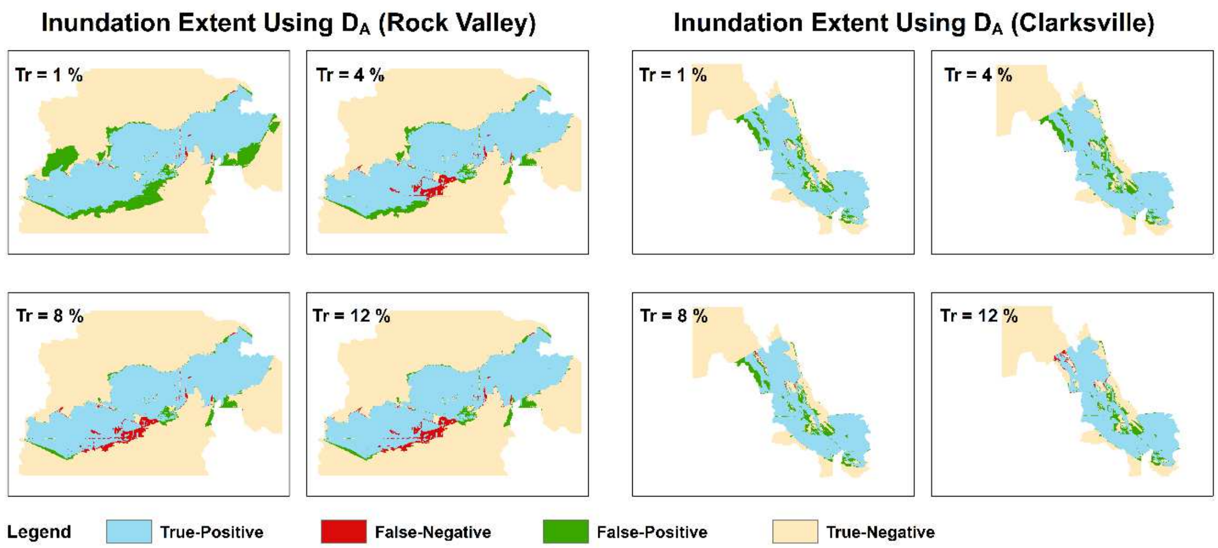


Figure 11. Inundation extent in 50-year scenario in Rock valley and Clarksville with drainage thresholds equal to 1%, 4%, 8%, and 12% of the study areas

As mentioned in previous sections, increasing drainage thresholds will remove some drainage pixels from the previous stream network. A shrinking stream network means the nearest drainage points of some hillslope pixels will move downward along the river channel due to the cancelation of their previous drainage pixels. A more downstream drainage pixel usually leads to an increase in the HAND value of those hillslope pixels as the elevation of drainage pixels tends to decrease along the river channel and HAND is the elevation difference between any hillslope pixel and its nearest drainage pixel. Eventually, some hillslope pixels could be no longer inundated if their new HAND values exceed the water depth. Figure 11 well depicted the shrinking of the predicted flood extents as we increase drainage thresholds in both study areas. However, given the fact the locations of those HAND value changes are determined by flow direction rather than chosen manually, this may reduce overestimation and improve predictions, such as the case in Clarksville, but it may also introduce more underestimation, such as the case in Rock Valley, and therefore and does not always indicate improvements to results.

5 Discussion

5.1 Analysis of the Underestimation Factors for WBH

While WBH provided comparable predictions for Rock Valley, it has some underestimation in Clarksville caused by three possible factors: 1) input data resolution loss during data format conversion; 2) limitations in algorithm used for resolving flats, which results in elevation increase in specific locations; and 3) low values for depth derived from the synthetic rating curve.

There are two major factors that could introduce input data resolution loss during the calculation of the WBH system's HAND matrix including data format and conversion. Because TIFF is not a well-supported format for web applications, the DEM data is converted to an RGB PNG file from the original format of TIFF for use in the web-based system. The system must then convert the RGB values back to elevation values in order to perform a pixel-level computation. This two-way conversion might introduce uncertainty into the HAND matrix calculation as pixels on RGB images range from 0 to 255 but have a much wider range in their original format Fitting a range of values to a narrower new range could compress the data and result in precision loss. Furthermore, the system stores each pixel's HAND value in integer format. This is because the system is designed and optimized to run efficiently on standard personal computers. As a result, it makes use of the efficient built-in data structure of the programming language that is used to create the system. The system's core language, JavaScript, handles arrays of integers more efficiently than arrays of float numbers.

As mentioned in the previous section, HAND matrix generation is based on flow directions of pixels and is derived from a hydrologically coherent DEM after flats (due to both natural flats and pit-resolving algorithms) are removed. To resolve the imperfections on the DEM, the WBH system employs the algorithm proposed by (Barnes *et al.* 2014). This algorithm first detects flats, which consist of a cluster of nearby pixels with equal elevation values, and then raises the elevation of those pixels based on their distance from the surrounding non-flat pixels (referred to as *the gradient away from higher terrain*) and the outlet of the entire flat area (referred to as *the gradient towards lower terrain*). The algorithm increases the elevation of a pixel more if that pixel is closer to the non-flat surrounding pixels. Similarly, the further a pixel is away from the flat area's outlet, the greater the elevation increase will be. As a result, it ensures the generation of flow direction for each pixel at the expense of changing directions for some pixels for which the elevation values increase. In the upper box of Figure 12, the stream initially flows downward before merging with the mainstream to the right. However, as the segment's elevation rises, it flows upward, merges with the main channel, and disconnects the stream in between. Same reason also applies to the situation in the lower box but is less obvious. This is unavoidable for flat resolving algorithms because they require elevation changes to force flows to drain from previously flat regions. As a result, the inundation condition and localized flow directions may differ from what the raw DEM indicates.

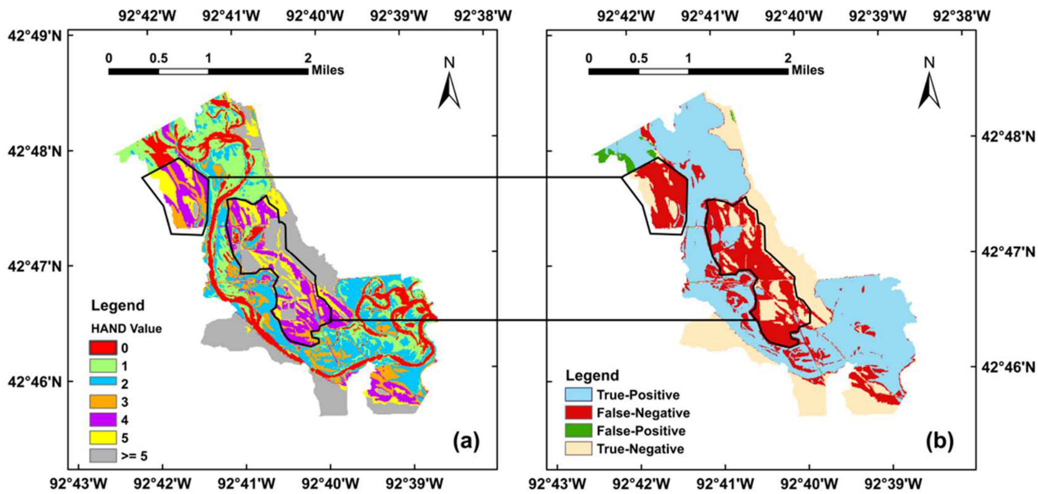


Figure 12. Predictions of flood inundation in a 100-year scenario (b) and visualization of the corresponding HAND matrix (a). The underestimation in the upper and lower black boxes is primarily due to elevation changes caused by the pit-removal algorithm and a small depth estimate derived from the synthetic rating curve.

Although the first two factors can cause some localized changes in the inundation condition, we believe that it is the last reason the primary contributor to the underestimation in Clarksville because elevation increase caused by the DEM conditioning occurred to only 4.8 percent of the pixels in the comparison. Among those 4.8 percent pixels, approximately 77 percent only had a one-meter elevation increase. Furthermore, the results in Rock Valley were generated using the same computation framework and algorithm but did not show significant underestimation compared to those from NWCH. The depth values derived from the synthetic rating curve and the corresponding depths in each catchment are shown in Figure 13.

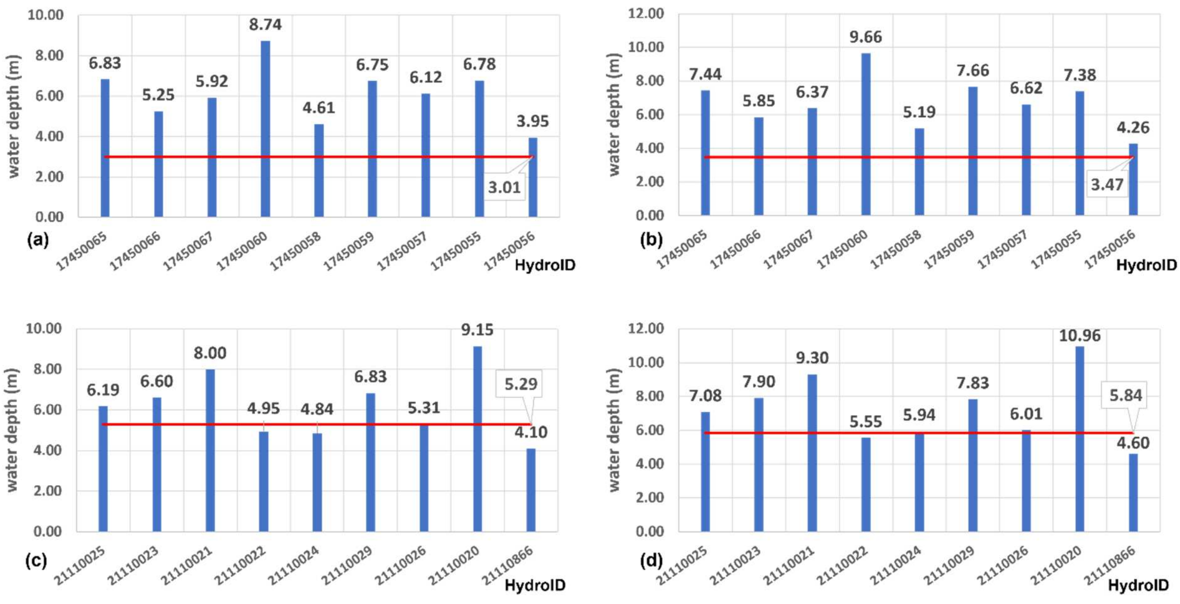


Figure 13. The depth estimates in meter from the synthetic rating curve (red horizontal line) and from each catchment in 100-year flood scenarios in Clarksville (a) and Rock Valley (c)and 500-year flood scenarios in Clarksville (b) and Rock Valley (d)

The HAND model is a ‘static’ inundation mapping technique, as opposed to models that rely on hydrodynamic simulations, such as a HEC-RAS model. As a result, it may fail to provide as reliable inundation extent predictions in areas where a single depth is not enough to reflect the situations in sub-areas. As illustrated in Figure 13 (a) and (b), the synthetic depths are too small compared to the depths in each catchment, therefore making the single value less representative of the inundation condition in general. However, D_S can produce favorable results in areas where synthetic situation is relatively consistent with conditions in catchments, such as Rock Valley. Furthermore, compared to multi-depth approaches, D_S requires significantly less data and computational efforts compared to multi-depth approaches even the multi-depth ones are already quite data parsimonious compared to many traditional flood modeling approaches and the NWCH implementation.

We believe that an efficient flood response strategy could first benefit from a fast model, such as the single-depth HAND framework, that requires the least data but can accurately show where major inundation will happen in order to support mitigation and planning decisions (Carson *et al.* 2018, Teague *et al.* 2021). Then it can follow a refined model, such as the multi-depth HAND, to ensure the inundation extent prediction free from major mismatches to benefit the accurate evacuation (Alabbad *et al.* 2021) and protection of people and property. Traditional flood inundation models can still be used for long-term planning, damage assessment, and documentation for the flood characteristics (such as inundation extent, localized maximum volume & stage), and serve as a reference to validate and improve data-driven flood models.

Although the running time of NWCH and WBH is hardly comparable as the former runs on the Resourcing Open Geospatial Education and Research (ROGER) supercomputer housed at NSF cyberGIS Facility whereas the latter runs on a personal desktop, we still include some rough estimations of both implementations here to give readers a sense of their computation efficiency. Liu *et al.* (2016) reported that for NWCH, the 10-m HAND raster for each HUC 6 computation unit tested required an estimated 65.26 CPU cores and took about 0.54 CPU hours to generate on ROGER HPC. Whereas the WBH requires about 15 seconds to finish computing the 5m HAND raster for a typical HUC 12 sub-watershed on a standard desktop with i7-2600 CPU (Li and Demir 2022). Given the fact there are about 335 HUC 6 watersheds and 90,000 HUC 12 sub-watersheds in the continental United States, WBH will require a theoretical running time of 16.8 minutes to compute a HUC 6 watershed on the same personal desktop mentioned above. It is thus obvious that the original implementation of HAND is significantly more efficient than the NWCH in terms of computational cost and resources for HAND layer generation. Once HAND values are computed, NWCH and WBH have comparable computational efficiency in creating the flood maps as the last step, namely a pixel-by-pixel comparison between the water depth and the HAND matrix, is the same in the two approaches. Both approaches will require just a few seconds to create a flood map for a HUC8 watershed.

Last but not least, WBH is just one possible implementation of the original HAND, and more accurate implementations could be designed to avoid the data precision issue (the first factor discussed in this section that caused the underestimation) based on different needs of usage scenarios. For instance, we could build stand-alone applications where data format conversion is not necessary. Or we could preserve the different digits of DEM pixels in separate color channels.

5.2 Evaluation of Multi-Depth Approaches and Drainage Threshold Parameter

Given the performance comparison depicted in Figures 7 to 10, we think that D_A and D_L outperform D_{Local} because these three generate comparable results but D_{Local} necessitates additional computation to determine the areas that each catchment outlet drains so that varied water depths can be applied to the appropriate locations. The D_{Local} and NWCH computations differ in that the FIM 3 versions consider flooding in each selected catchment separately as water is not allowed to spread beyond the catchment boundary. The scope of each catchment is recorded in a GeoPackage file and stored as ‘static’ data that is not updated constantly. By contrast, in D_{Local}, the region each outlet drains is defined by flow directions, which means that the area an outlet controls varies between different thresholds and DEM inputs. Therefore, D_{Local} provides a more accurate representation of possible changes to the topography (such as dredging and land cover changes) and provides stakeholders with greater flexibility.

Previous research has demonstrated that a drainage threshold of 4 km^2 reduces mismatches and improves inundation extent forecasts (Nobre *et al.* 2016). However, the findings of this study reveal that the 4 km^2 does not produce the best results in either study region since it is too small and leads to overestimation.

Conclusion

In this study, we first examined the performance of two implementations of the HAND model, NWCH and WBH, in terms of the extent maps generated from both approaches to investigate the efficacy of the simpler approach of HAND versus a more complex one. The results show that the WBH with the least data input (a single water depth value, an unadjusted drainage threshold, and DEM) can generate comparable inundation extent predictions to the NWCH in areas where the water depths from the synthetic rating curves and those derived from each catchment's separate rating curves are relatively consistent. Otherwise, WBH predictions with the simplest model configuration may be underestimated due to a combination of 1) localized flow direction changes caused by the pit-removing algorithm; 2) inaccuracy of the HAND values at a pixel level due to data format transformation and storage; and 3) differences between the depth estimates for the synthetic rating curves and the rating curves for each catchment. However, this underestimation can be avoided by employing multiple water depths and can also be avoided by applying carefully designed pre-processing steps to keep data precisions.

We also tested the performance of HAND with various model configurations for WBH model. Our results indicate that using 4 km^2 as the drainage threshold value as suggested by early studies results in too many overestimations in both study areas. In our cases, a good threshold falls in the range of 8 to 12 % of the total area. Once exceeding the optimal threshold, the model performance will become stable. We did not see significant performance differences among cases with the three multi-depth approaches (D_A, D_L, and D_{Local}) as no approach can consistently outperform the others considering different flooding scenarios and study areas. We, therefore, believe that D_A and D_L are better than D_{Local} as they require less computation effort. Besides that, all three multi-depth approaches are more robust to factors that may lead to underestimation as compared with D_S.

The results of this study demonstrate the efficacy of the original HAND in flood extent prediction. It requires far fewer computational resources and data dependencies compared to the well-studied NWCH implementation thanks to its simple computation framework. The original HAND is especially suitable for applications where we need acceptable accuracy but fast results, where the inputs, such as DEM, are at constant changes, and where we need to deal with inputs not following pre-defined settings of the NWCH, such as water depth observations collected not at the catchment outlets defined in NHDPlus dataset.

Software availability

Software: NOAA Flood inundation mapping and evaluation software

Developer: NOAA-OWP, National Water Center

Operating systems: Windows, Linux, and Mac OSX

Dependent software: Docker

Availability: The software is publicly available and can be accessed from the GitHub repository at https://github.com/NOAA-OWP/inundation-mapping

Software: HAND Multi-Depth Module

Developer: Zhouyayan Li, Ibrahim Demir

Availability: The module is publicly available and can be accessed from the GitHub repository at https://github.com/uihilab/HandMultiDepth.

Readers can access https://hydroinformatics.uiowa.edu/lab/handmultidepth for an interactive demonstrating webpage to see how the module works.

Acknowledgement

We would like to acknowledge the assistance and support provided by experts at the Iowa Flood Center for the reference community maps. We also would like to thank NOAA Office of Water Prediction, Lynker, and National Water Center for the NWCH model and data.

Funding

This project is supported by the COMET Outreach Program from the University Corporation for Atmospheric Research (UCAR) which is supported by National Oceanic and Atmospheric Administration (NOAA). Award No: SUBAWD002442.

Appendix

Table A1. List of Abbreviations

Abbreviation	**Definition**
B	Bias, an index to evaluate the model's tendency of making positive (flooded) predictions versus negative (dry) predictions
D_A	A multi-depth approach that takes the average of depth values from rating curves of multiple catchments weighted by catchment areas
D_L	A multi-depth approach that takes the average of depth values from rating curves of multiple catchments weighted by catchment stream length
D_{Local}	A multi-depth approach that applies depth values to the corresponding provider

	catchments without taking the average
D_S	A single-depth approach that uses depth value derived from synthetic rating curve to compare with HAND value for the entire study area
F	Fitness Statistics, an index evaluating model performance that focuses more on flooded pixels
H	Hit Rate, indicating how many flooded pixels are recognized by a model
HAND	Height Above Nearest Drainage
K	Kappa value, an index evaluating model performance that focuses more on dry pixels
NWC	National Water Center
NWM	National Water Model
NWCH	HAND framework implemented at NWC
PC	Proportion Correct, an index indicating how many predictions are correct
WBH	Web-based HAND, a web framework designed to run the original HAND on-the-fly

Pseudocode for Multi-Depth Module

HAND—HAND grid, recording HAND values of each pixel in a study area.

drainageMx—drainage Matrix, recording how many upstream pixels each pixel drains, same dimension with the HAND grid. Generated along with the HAND grid.

flowMx—flow direction Matrix, recording the D8 flow direction of each pixel, same dimension with the HAND grid. Generated along with the HAND grid.

pointList—a list of position or index of all pixels in a study area, same dimension with the HAND grid. Encoded before all computation starts.

locationList—a nested list of locations where the custom depth values to be applied, each element of the list is the X, Y coordinates or row & column index of the locations. This has to be provided.

depthList—a list of water depths assigned to locations in locationList, same length with locationList Has to be provided

areaList—a list of area drained by each pixel in the location List, same length with locationList. Can be obtained from existing datasets, such as NHDPlus, measurements on a real map or a digital map such as DEM, or be derived using flowMx.

streamList— a list of stream length each pixel controls in the location List, same length with locationList. Can be obtained from existing datasets, such as NHDPlus, or measurements on a real map or a digital map such as DEM.

depth—a depth value that will control all pixels that are not drained by any custom drainage point in locationList. Has to be provided.

```
Procedure dye(drainageMx, flowMx, locationList, pointList):
    # initial a new matrix filled with zero to record which upstream pixel
    # will be drained by which drainage pixel
    colorMx = zeros(size=HAND.size)

    # get drainage area for all locations in locationList
    drainageList = getDrainArea(drainageMX, locationList)
    for all location in locationList do
        upstreamPoints = getUpstreamPoints(flowMx, location, pointList)
        for all point in upstreamPoints do
            if colorMx[point] == 0 or colorMx[point] > drainageList[location] do
                # Update the color of current point if it has not been changed before
                # or it is controlled by a more downstream point
                colorMx[point] = drainageList[location]
    return colorMx, drainageList

Procedure multiDepth(pointList, HAND, depthList, depth, mode):
    # Initial a new matrix filled with zero to record the final depth of each pixel
    depthMx = zeros(size=HAND.size)
    # Initial a new matrix filled with zero to record the inundation condition of each pixel
```

```
# where 0 is dry and 1 is flooded
floodMx = zeros(size=HAND.size)

colorMx, drainageList = dye(drainageMx, flowMx, locationList, pointList)

# Compute inundation extent using DA
if mode == ‘Da’ do
    depth_DA = weightedAvg(depthList, areaList)
    for all point in pointList do
        if colorMX[point] != 0 do
            if depth_DA > HAND[point] do
                floodMx[point] = 1
        else if depth > HAND[point] do
            floodMx[point] = 1

# Compute inundation extent using DL
elif mode == ‘Dl’ do
    depth_DL = weightedAvg(depthList, streamList)
    for all point in pointList do
        if colorMX[point] != 0 do
            if depth_DL > HAND[point] do
                floodMx[point] = 1
        else if depth > HAND[point] do
            floodMx[point] = 1

# Compute inundation extent using DLocal
else do
    for all point in pointList do
```

```
            # Update the inundation conditions for pixels controlled/not controlled by custom
            # locations in locationList seperately
            if colorMX[point] != 0 do
                depthMx[point] = depthList[index of colorMx[point] in deinageList]
                if depthMx[point] > HAND[point] do
                    floodMx[point] = 1
            else if depth > HAND[point] do
                floodMx[point] = 1
    return floodMx

Procedure getDrainArea(drainageMx, locationList):
    drainList = empty list with the same dimension with locationList
    for all location in locationList do
        append drainageMx[location] to drainList
    return drainList

Procedure getUpstreamPoints(flowMx, location, pointList):
    # Recursively find all upstream points of the given location based on flowMx
    # and put those points in a variable upstreamPoints
    return upstreamPoints

Procedure weightedAvg(depthList, weights):
    avg = 0
    for every item in depthList do
        avg += depthList[item]*weights[item]
    return avg /= sum(weights)
```

Reference

Abdrabo, K.I., Kantosh, S.A., Saber, M., Sumi, T., Elleithy, D., Habiba, O.M., and Alboshy, B., 2022. The Role of Urban Planning and Landscape Tools Concerning Flash Flood Risk Reduction Within Arid and Semiarid Regions. *In*: *Wadi Flash Floods*. 283–316.

Afshari, S., Tavakoly, A.A., Rajib, M.A., Zheng, X., Follum, M.L., Omranian, E., and Fekete, B.M., 2018. Comparison of new generation low-complexity flood inundation mapping tools with a hydrodynamic model. *Journal of Hydrology*, 556, 539–556.

Agliamzanov, R., Sit, M., and Demir, I., 2020. Hydrology@Home: a distributed volunteer computing framework for hydrological research and applications. *Journal of Hydroinformatics*, 22 (2), 235–248.

Alabbad, Y., Mount, J., Campbell, A.M., and Demir, I., 2021. Assessment of transportation system disruption and accessibility to critical amenities during flooding: Iowa case study. *Science of The Total Environment*, 793, 148476.

Alabbad, Y., Yildirim, E., and Demir, I., 2022. Flood mitigation data analytics and decision support framework: Iowa Middle Cedar Watershed case study. *Science of The Total Environment*, 814, 152768.

Di Baldassarre, G., Saccà, S., Aronica, G.T., Grimaldi, S., Ciullo, A., and Crisci, M., 2017. Human-flood interactions in Rome over the past 150 years. *Advances in Geosciences*, 44, 9–13.

Baldassarre, G.D., Nardi, F., Annis, A., Odongo, V., Rusca, M., and Grimaldi, S., 2020. Brief communication: Comparing hydrological and hydrogeomorphic paradigms for global flood hazard mapping. *Natural Hazards and Earth System Sciences*, 20 (5), 1415–1419.

Barnes, R., Lehman, C., and Mulla, D., 2014. An efficient assignment of drainage direction over flat surfaces in raster digital elevation models. *Computers and Geosciences*, 62, 128–135.

Bentivenga, M., Giano, S.I., and Piccarreta, M., 2020. Recent Increase of Flood Frequency in the Ionian Belt of Basilicata Region, Southern Italy: Human or Climatic Changes? *Water 2020, Vol. 12, Page 2062*, 12 (7), 2062.

Blöschl, G., Kiss, A., Viglione, A., Barriendos, M., Böhm, O., Brázdil, R., Coeur, D., Demarée, G., Llasat, M.C., Macdonald, N., Retsö, D., Roald, L., Schmocker-Fackel, P., Amorim, I., Bělínová, M., Benito, G., Bertolin, C., Camuffo, D., Cornel, D., Doktor, R., Elleder, L., Enzi, S., Garcia, J.C., Glaser, R., Hall, J., Haslinger, K., Hofstätter, M., Komma, J., Limanówka, D., Lun, D., Panin, A., Parajka, J., Petrić, H., Rodrigo, F.S., Rohr, C., Schönbein, J., Schulte, L., Silva, L.P., Toonen, W.H.J., Valent, P., Waser, J., and Wetter, O., 2020. Current European flood-rich period exceptional compared with past 500 years. *Nature 2020 583:7817*, 583 (7817), 560–566.

Brandt, H.M., Turner-Mcgrievy, G., Friedman, D.B., Gentile, D., Schrock, C., Thomas, T., and West, D., 2019. Examining the Role of Twitter in Response and Recovery during and after Historic Flooding in South Carolina. *Journal of Public Health Management and Practice*, 25 (5), E6–E12.

Carson, A., Windsor, M., Hill, H., Haigh, T., Wall, N., Smith, J., Olsen, R., Bathke, D., Demir, I., and Muste, M., 2018. Serious gaming for participatory planning of multi-hazard mitigation. *International Journal of River Basin Management*, 16 (3), 379–391.

Demiray, B.Z., Sit, M., and Demir, I., 2021. D-SRGAN: DEM Super-Resolution with Generative Adversarial Networks. *SN Computer Science*, 2 (1), 1–11.

Ebert-Uphoff, I., Thompson, D.R., Demir, I., Gel, Y.R., Hill, M.C., Karpatne, A., Guereque, M., Kumar, V., Cabral-Cano, E., and Smyth, P., 2017. A vision for the development of benchmarks to bridge geoscience and data science. *In*: *7th International Workshop on Climate Informatics*.

Ewing, G. and Demir, I., 2021. An ethical decision-making framework with serious gaming: a smart water case study on flooding. *Journal of Hydroinformatics*, 23 (3), 466–482.

Garrote, J., González-Jiménez, M., Guardiola-Albert, C., and Díez-Herrero, A., 2021. The manning's roughness coefficient calibration method to improve flood hazard analysis in the absence of river bathymetric data: Application to the urban historical zamora city centre in spain. *Applied Sciences (Switzerland)*, 11 (19), 9267.

Ghosh, A. and Kar, S.K., 2018. Application of analytical hierarchy process (AHP) for flood risk assessment: a case study in Malda district of West Bengal, India. *Natural Hazards*, 94 (1), 349–368.

Gilles, D., Young, N., Schroeder, H., Piotrowski, J., and Chang, Y.J., 2012. Inundation mapping initiatives of the iowa flood center: Statewide coverage and detailed urban flooding analysis. *Water (Switzerland)*, 4 (1), 85–106.

Godbout, L., Zheng, J.Y., Dey, S., Eyelade, D., Maidment, D., and Passalacqua, P., 2019. Error Assessment for Height Above the Nearest Drainage Inundation Mapping. *Journal of the American Water Resources Association*, 55 (4), 952–963.

Haltas, I., Yildirim, E., Oztas, F., and Demir, I., 2021. A comprehensive flood event specification and inventory: 1930–2020 Turkey case study. *International Journal of Disaster Risk Reduction*, 56.

Hocini, N., Payrastre, O., Bourgin, F., Gaume, E., Davy, P., Lague, D., Poinsignon, L., and Pons, F., 2020. Performance of automated flood inundation mapping methods in a context of flash floods: a comparison of three methods based either on the Height Above Nearest Drainage (HAND) concept, or on 1D/2D shallow water equations. *Hydrology and Earth System Sciences Discussions*, (December), 1–23.

Hu, A. and Demir, I., 2021. Real-time flood mapping on client-side web systems using hand model. *Hydrology*, 8 (2), 65.

Huber, M., Osterkamp, N., Marschalk, U., Tubbesing, R., Wendleder, A., Wessel, B., and Roth, A., 2021. Shaping the Global High-Resolution TanDEM-X Digital Elevation Model. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14, 7198–7212.

Jafarzadegan, K. and Merwade, V., 2019. Probabilistic floodplain mapping using HAND-based statistical approach. *Geomorphology*, 324, 48–61.

Janizadeh, S., Chandra Pal, S., Saha, A., Chowdhuri, I., Ahmadi, K., Mirzaei, S., Mosavi, A.H., and Tiefenbacher, J.P., 2021. Mapping the spatial and temporal variability of flood hazard affected by climate and land-use changes in the future. *Journal of Environmental Management*, 298, 113551.

Juurlink, D.N. and Detsky, A.S., 2005. Kappa statistic. *CMAJ*, 173 (1), 16–16.

Komolafe, A.A., Olorunfemi, I.E., Akinluyi, F.O., Adeyemi, M.A., and Ajayi, J.A., 2021. Enhanced flood hazard modelling using hydraulic, analytical hierarchical process and height above nearest drainage models in Ogunpa river basin, Ibadan, Southwestern Nigeria. *Modeling Earth Systems and Environment*, 7 (2), 967–981.

Lababidi, S., 2021. AN EVALUATION OF THE HEIGHT ABOVE NEAREST DRAINAGE (HAND) FLOOD MAPPING METHODOLOGY WITH THE IMPLEMENTATION OF MULTIVARIANT LINEAR REGRESSION ALGORITHM. Northeastern University.

Lancia, M., Zheng, C., He, X., Lerner, D.N., Andrews, C., and Tian, Y., 2020. Hydrogeological constraints and opportunities for “Sponge City” development: Shenzhen, southern China. *Journal of Hydrology: Regional Studies*, 28, 100679.

de Lange, J., 2019. 3. History of flood defenses in the Low Countries. *In*: *Security of Flood Defenses*. De Gruyter, 21–33.

Le, T. and Bae, D.H., 2020. Response of global evaporation to major climate modes in historical and future Coupled Model Intercomparison Project Phase 5 simulations. *Hydrology and Earth System Sciences*, 24 (3), 1131–1143.

Leitner, M., Babcicky, P., Schinko, T., and Glas, N., 2020. The status of climate risk management in Austria. Assessing the governance landscape and proposing ways forward for comprehensively managing flood and drought risk. *Climate Risk Management*, 30, 100246.

Li, Z. and Demir, I., 2022. A comprehensive web-based system for flood inundation map generation and comparative analysis based on height above nearest drainage. *Science of The Total Environment*, 828, 154420.

Li, Z., Mount, J., and Demir, I., 2022. Accounting for uncertainty in real-time flood inundation mapping using HAND model: Iowa case study. *Natural Hazards*, 1–28.

Li, Z., Wang, C., Emrich, C.T., and Guo, D., 2018. A novel approach to leveraging social media for rapid flood mapping: a case study of the 2015 South Carolina floods. *Cartography and Geographic Information Science*, 45 (2), 97–110.

Liu, Y.Y., Maidment, D.R., Tarboton, D.G., Zheng, X., and Wang, S., 2018. A CyberGIS Integration and Computation Framework for High-Resolution Continental-Scale Flood Inundation Mapping. *JAWRA Journal of the American Water Resources Association*, 54 (4), 770–784.

Liu, Y.Y., Maidment, D.R., Tarboton, D.G., Zheng, X., Yildirim, A., Sazib, N.S., and Wang, S., 2016. A CyberGIS Approach to Generating High-resolution Height Above Nearest Drainage (HAND) Raster for National Flood Mapping. *The Third International Conference on CyberGIS and Geospatial Data Science*, (August), 1–5.

Maidment, D.R., 2017. Conceptual Framework for the National Flood Interoperability Experiment. *JAWRA Journal of the American Water Resources Association*, 53 (2), 245–257.

Mark, D.M., 1984. Automated detection of drainage networks from digital elevation models. *Cartographica: The International Journal for Geographic Information and Geovisualization,* 21 (2–3), 168–178.

McDougall, K. and Temple-Watts, P., 2012. THE USE of LIDAR and VOLUNTEERED GEOGRAPHIC INFORMATION to MAP FLOOD EXTENTS and INUNDATION. *In*: *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. 251–256.

McGrath, H., Bourgon, J.F., Proulx-Bourque, J.S., Nastev, M., and Abo El Ezz, A., 2018. A comparison of simplified conceptual models for rapid web-based flood inundation mapping. *Natural Hazards*, 93 (2), 905–920.

Michael Johnson, J., Munasinghe, D., Eyelade, D., and 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.

Mizzell, H., Malsick, M., and Tyler, W., 2017. The Historic South Carolina Rainfall and Major Floods of October 1-5, 2015. *Journal of South Carolina Water Resources*, 3 (1), 2.

Munoz, S.E., Giosan, L., Therrell, M.D., Remo, J.W.F., Shen, Z., Sullivan, R.M., Wiman, C., O'Donnell, M., and Donnelly, J.P., 2018. Climatic control of Mississippi River flood hazard amplified by river engineering. *Nature 2018 556:7699*, 556 (7699), 95–98.

Musser, J.W., Watson, K., Painter, J.A., and Gotvald, A.J., 2016. Flood inundation maps of selected areas affected by the flood of October 2015 in central and coastal South Carolina. *Open-File Report - U. S. Geological Survey*, (October 2015), 81.

Nardi, F., Annis, A., Baldassarre, G. Di, Vivoni, E.R., and Grimaldi, S., 2019. GFPLAIN250m, a global high-resolution dataset of earth's floodplains. *Scientific Data*, 6 (1), 1–6.

Nicholls, R.J., Lincke, D., Hinkel, J., Brown, S., Vafeidis, A.T., Meyssignac, B., Hanson, S.E., Merkens, J.L., and Fang, J., 2021. A global analysis of subsidence, relative sea-level change and coastal flood exposure. *Nature Climate Change 2021 11:4*, 11 (4), 338–342.

NOAA, 2021. NOAA-OWP/cahaba: Flood inundation mapping and evaluation software configured to work with U.S. National Water Model. [online]. Available from: https://github.com/NOAA-OWP/cahaba [Accessed 30 Dec 2021].

Nobre, A.D., Cuartas, L.A., Hodnett, M., Rennó, C.D., Rodrigues, G., Silveira, A., Waterloo, M., and Saleska, S., 2011. Height Above the Nearest Drainage - a hydrologically relevant new terrain model. *Journal of Hydrology*, 404 (1–2), 13–29.

Nobre, A.D., Cuartas, L.A., Momo, M.R., Severo, D.L., Pinheiro, A., and Nobre, C.A., 2016. HAND contour: A new proxy predictor of inundation extent. *Hydrological Processes*, 30 (2), 320–333.

Nohani, E., 2019. Estimating the Manning's Roughness Coefficient in Rivers by Experimental Method. *Foot*, 1 (6), 50.

Papaioannou, G., Vasiliades, L., Loukas, A., and Aronica, G.T., 2017. Probabilistic flood inundation mapping at ungauged streams due to roughness coefficient uncertainty in hydraulic modelling. *Advances in Geosciences*, 44, 23–34.

Provost, F., 1998. Glossary of Terms Special Issue on Applications of Machine Learning and the

Knowledge Discovery Process. *Machine Learning*, 30, 271–274.

Psomiadis, E., Tomanis, L., Kavvadias, A., Soulis, K.X., Charizopoulos, N., and Michas, S., 2021. Potential Dam Breach Analysis and Flood Wave Risk Assessment Using HEC-RAS and Remote Sensing Data: A Multicriteria Approach. *Water 2021, Vol. 13, Page 364*, 13 (3), 364.

Rennó, C.D., Nobre, A.D., Cuartas, L.A., Soares, J.V., Hodnett, M.G., Tomasella, J., and Waterloo, M.J., 2008. HAND, a new terrain descriptor using SRTM-DEM: Mapping terra-firme rainforest environments in Amazonia. *Remote Sensing of Environment*, 112 (9), 3469–3481.

S Chegwidden, O., Rupp, D.E., and Nijssen, B., 2020. Climate change alters flood magnitudes and mechanisms in climatically-diverse headwaters across the northwestern United States. *Environmental Research Letters*, 15 (9), 094048.

Savage, J.T.S., Bates, P., Freer, J., Neal, J., and Aronica, G., 2016. When does spatial resolution become spurious in probabilistic flood inundation predictions? *Hydrological Processes*, 30 (13), 2014–2032.

Seo, B.C., Keem, M., Hammond, R., Demir, I., and Krajewski, W.F., 2019. A pilot infrastructure for searching rainfall metadata and generating rainfall product using the big data of NEXRAD. *Environmental Modelling and Software*, 117, 69–75.

Shastry, A., Egbert, R., Aristizabal, F., Luo, C., Yu, C.-W., and Praskievicz, S., 2019. Using Steady-State Backwater Analysis to Predict Inundated Area from National Water Model Streamflow Simulations. *JAWRA Journal of the American Water Resources Association*, 55 (4), 940–951.

Sit, M., Langel, R.J., Thompson, D., Cwiertny, D.M., and Demir, I., 2021. Web-based data analytics framework for well forecasting and groundwater quality. *Science of The Total Environment*, 761, 144121.

Sit, M., Sermet, Y., and Demir, I., 2019. Optimized watershed delineation library for server-side and client-side web applications. *Open Geospatial Data, Software and Standards*, 4 (1), 1–10.

Smith, L., Liang, Q., James, P., and Lin, W., 2017. Assessing the utility of social media as a data source for flood risk management using a real-time modelling framework. *Journal of Flood Risk Management*, 10 (3), 370–380.

Speckhann, G.A., Borges Chaffe, P.L., Fabris Goerl, R., Abreu, J.J. de, and Altamirano Flores, J.A., 2018. Flood hazard mapping in Southern Brazil: a combination of flow frequency analysis and the HAND model. *Hydrological Sciences Journal*, 63 (1), 87–100.

Sun, X., Zhang, H., Hua, D., and Wei, B., 2021. The Influence of Urbanization on Storm Runoff. *In*: *IOP Conference Series: Earth and Environmental Science*. IOP Publishing, 022031.

Tadesse, Y.B. and Fröhle, P., 2020. Modelling of flood inundation due to levee breaches: Sensitivity of flood inundation against breach process parameters. *Water (Switzerland)*, 12 (12), 3566.

Tapete, D., Traviglia, A., Delpozzo, E., and Cigna, F., 2021. Regional-scale systematic mapping of archaeological mounds and detection of looting using cosmo-skymed high resolution dem and satellite imagery. *Remote Sensing*, 13 (16), 3106.

Tarboton, D.G., 1997. A new method for the determination of flow directions and upslope areas in grid digital elevation models. *Water Resources Research*, 33 (2), 309–319.

Teague, A., Sermet, Y., Demir, I., and Muste, M., 2021. A collaborative serious game for water resources planning and hazard mitigation. *International Journal of Disaster Risk Reduction*, 53, 101977.

Teng, J., Jakeman, A.J., Vaze, J., Croke, B.F.W., Dutta, D., and Kim, S., 2017. Flood inundation modelling: A review of methods, recent advances and uncertainty analysis. *Environmental Modelling and Software*.

Terezinha, V., Boulomytis, G., Zuffo, A.C., Gilberto, J., Filho, D., Monzur, &, Imteaz, A., Carlos Zuffo, A., and Alam Imteaz, M., 2017. Estimation and calibration of Manning's roughness coefficients for ungauged watersheds on coastal floodplains. *http://dx.doi.org/10.1080/15715124.2017.1298605*, 15 (2), 199–206.

Wilks, D.S., 2011. *Statistical Methods in the Atmospheric Sciences*. Academic Press.

Xu, H., Demir, I., Koylu, C., and Muste, M., 2019. A web-based geovisual analytics platform for identifying potential contributors to culvert sedimentation. *Science of The Total Environment*, 692, 806–817.

Yildirim, E. and Demir, I., 2021. An Integrated Flood Risk Assessment and Mitigation Framework: A Case Study for Middle Cedar River Basin, Iowa, US. *International Journal of Disaster Risk Reduction*, 56, 102113.

Zheng, X., Tarboton, D.G., Maidment, D.R., Liu, Y.Y., and Passalacqua, P., 2018. River Channel Geometry and Rating Curve Estimation Using Height above the Nearest Drainage. *Journal of the American Water Resources Association*, 54 (4), 785–806.