Extending Height Above Nearest Drainage to Model Multiple Fluvial Sources in Flood Inundation Mapping Applications for the U.S. National Water Model

Fernando Aristizabal[1,2,3], Fernando Salas[3], Gregory Petrochenkov[4], Trevor Grout[1,3], Brian Avant[1,3], Bradford Bates[1,3], Ryan Spies[1,3], Nick Chadwick[3], Zachary Wills[3,5], Jasmeet Judge[2]

[1]Lynker, Leesburg, VA, USA

[2]Center for Remote Sensing, Agricultural and Biological Engineering Department, University of Florida, Gainesville, FL, USA

[3]National Water Center, Office of Water Prediction, National Weather Service, National Oceanic and Atmospheric Administration, Tuscaloosa, AL, USA

[4]Hydrologic Applied Innovations Lab, New York Water Science Center, United States Geological Survey, Troy, NY, USA

[5]Cooperative Institute for Satellite Earth System Studies, University of Maryland, College Park, MD, USA

Key Points:

- Flood maps derived from Height Above Nearest Drainage (HAND) are subject to nearest drainage line limitations that affect inundation skill.
- A means of resolving this limitation is provided by reducing HAND processing units to level paths with effective unit stream order.
- Discretizing the stream network for HAND computation affects the stage-discharge relationship and leads to higher skill inundation.

Corresponding author: Fernando Aristizabal, `fernando.aristizabal@noaa.gov`

Abstract

Height Above Nearest Drainage (HAND), a drainage normalizing terrain index, is a means able of producing flood inundation maps (FIMs) from the National Water Model (NWM) at large scales and high resolutions using reach-averaged synthetic rating curves. We highlight here that HAND is limited to producing inundation only when sourced from its nearest drainage line, thus lacks the ability to source inundation from multiple fluvial sources. A version of HAND, known as Generalized Mainstems (GMS), is proposed that discretizes a target stream network into segments of unit Horton-Strahler stream order known as level paths (LP). The FIMs associated with each independent LP are then mosaiced together, effectively turning the stream network into discrete groups of homogeneous unit stream order by removing the influence of neighboring tributaries. Improvement in mapping skill is observed by significantly reducing false negatives at river junctions when the inundation extents are compared to FIMs from that of benchmarks. A more marginal reduction in the false alarm rate is also observed due to a shift introduced in the stage-discharge relationship by increasing the size of the catchments. We observe that the improvement of this method applied at 4-5% of the entire stream network to 100% of the network is about the same magnitude improvement as going from no drainage order reduction to 4-5% of the network. This novel contribution is framed in a new open-source implementation that utilizes the latest combination of hydro-conditioning techniques to enforce drainage and counter limitations in the input data.

Plain Language Summary

Flooding is one of the most impactful natural disasters on life and property. The United States National Water Model (NWM) provides flood forecasts for the entire country so that adequate warnings can be raised to the public to enable safe evacuations and protective measures. In order to convert flow rates from the NWM to flood inundation maps (FIM), a model, known as Height Above Nearest Drainage (HAND), is used that converts elevation data from height above mean sea-level to height above the nearest river bottom. This model suffers from issues in mapping performance because inundation sourced from rivers is only considered from the nearest river line. We developed a technique that mitigates these errors by removing consideration for neighboring tributaries in the relative elevation computation process. This is done by splitting the stream network into continuous river segments known as level paths (LPs). These LPs have no tributaries,

thus are known to be stream lines with a unit stream order indicating no branching. HAND

is computed independently for each LP and the resulting FIMs are mosaiced together

to form one seamless map. We compared these HAND derived FIMs to maps from physically-

based models and found improvement in mapping performance.

1 Introduction

Flooding is one of the most significant natural disasters in the United States (US)

affecting both the loss of life and property. In 2017 and 2019, river and flash flooding

combined represented the leading cause of death and the second leading cause in 2018

among all natural disasters in the US (Service, 2020b; National Weather Service, 2019,

2018). More than an average of 104 deaths per year are attributed to flood events from

the 10 year period ending in 2019 (Service, 2020a). With respect to property damages,

river and flash flooding have contributed to 60.7, 1.6, and 3.7 billion non-inflation ad-

justed US dollars in the annual periods of 2017 to 2019, respectively (Service, 2020b; Na-

tional Weather Service, 2019, 2018), with the large spike in 2017 attributed to the Hur-

ricane Harvey event along the Gulf Coast. Trends related to flood damages and fatal-

ities have been steadily increasing over recent decades (Mallakpour & Villarini, 2015; Down-

ton et al., 2005; Kunkel et al., 1999; Pielke Jr & Downton, 2000; Corringham & Cayan,

2019). Some are expecting that the hydrologic cycle will intensify due to climate change

which will lead to more extreme precipitation in some areas along with a greater risk of

flooding (Tabari, 2020; Milly et al., 2002; Wing et al., 2018). Increasing trends in fre-

quency and risk are not uniform across spatial regions with work by Slater and Villar-

ini (2016) indicating that trends are increasing across the US Midwest and Great Lakes

regions while decreasing in the coastal Southeast, Southwest, and California.

1.1 Operational Forecasting

Operational flood forecasting systems are primary tools in developing accurate fore-

casts for public awareness prior to life threatening and property damaging events. One

of these operational systems is the Advanced Hydrologic Prediction System (AHPS) main-

tained by the National Oceanic Atmospheric Administration (NOAA) National Weather

Service (NWS) with thousands of forecasting points across the US at typically short fore-

cast horizons of 24 or 72 hours (McEnery et al., 2005). AHPS provides forecasting ser-

vices in the form of ensemble streamflows at more than 3,600 locations and flood inun-

dation maps (FIM) at more than 150 of those points shown in Figure 1. Additionally,
two forecasting networks, Full Resolution (FR) and Mainstems (MS) stream networks,
relevant to the National Water Model (NWM) (see Section 1.2) are rendered in Figure
1. The FR network refers to the entire NWM forecasting domain while MS refers to the
subset of the FR network that is at or downstream of AHPS forecasting points (see Section 1.2). On an approximate basis, there is only one forecast point every 1,450 km of
river (FR) and one forecast point with FIM every 29,000 km of river (FR). Despite the
AHPS advances in operational flood forecasting, it lacks sufficient domain coverage, spatial resolution, and long-range forecast horizons to address the increasingly complex water challenges facing the US.
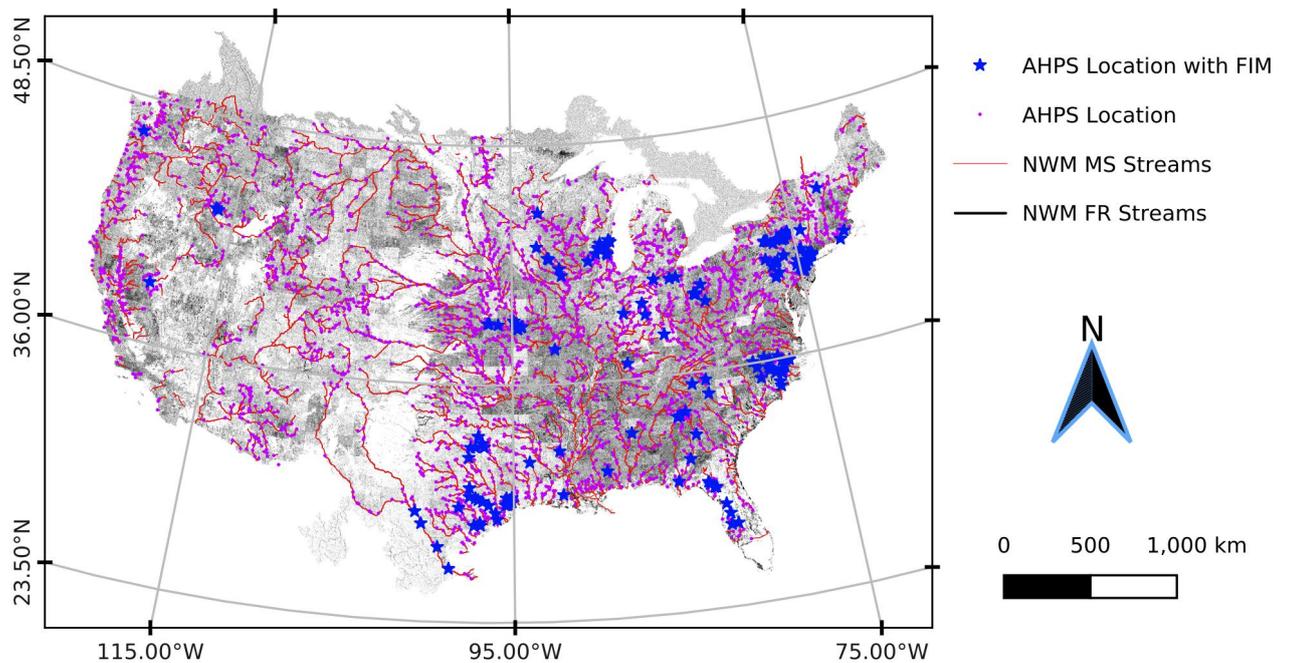

Figure 1. Forecast points with and without Flood Inundation Maps (FIM) in United States'
Advanced Hydrologic Prediction System (AHPS). Note that only a small fraction of the AHPS
forecast points have existing FIM. Also shown are the National Water Model (NWM) stream
networks at the Full Resolution (FR) and Mainstems (MS) resolution. The FR network constitutes the entire NWM stream network while the MS resolution network is the FR network at or
downstream of the AHPS forecast points shown.

1.2 National Water Model

Additional work is required to address the gaps that the AHPS leaves in terms of spatial resolution, spatial coverage, and temporal forecast horizons. In response to growing stakeholder demand for enhanced and integrated water resource forecasts, the Office of Water Prediction (OWP) at the National Water Center (NWC) along with its partners at the National Center for Atmospheric Research (NCAR) have developed and implemented operationally the NWM which is a configuration of the Weather Research and Forecasting Hydrologic Model (WRF-Hydro) (Salas et al., 2018; Gochis et al., 2021; Cosgrove et al., 2019). The NWM forecasts river discharges at more than 2.7 million forecast points at a variety of time horizons including lookback-range (3-28 hrs), short-range (18 hr), medium-range (10 day) and long-range (30 day) forecast horizons. The NWM enhances the spatial and temporal domain of the current AHPS capabilities operated at the 13 River Forecast Centers (RFC) in areas known as 'hydro-blind'. As a complement to the operational NWM, RFC forecasts from AHPS forecast points are assimilated in the NWM and routed downstream to the next downstream AHPS forecast point where the process iterates again. This assimilation into the NWM is used to enhance forecasting skill by leveraging best available regional-scale forecasts. The river network upon which this special assimilation technique operates on is herein referred to as the Mainstem (MS) stream network. Figure 1 shows the NWM V2.1 FR stream network as well as the NWM V2.1 MS network. The MS network contains roughly 120 thousand forecasting points or roughly 4.4% of the reaches of the FR stream network.

The National Hydrography Dataset Plus (NHDPlus) V2.1 is the basis for the "hydrofabric" in the NWM due to its comprehensive use with the hydrologic communities' stakeholders (McKay et al., 2012; *NHDPlusHR GDB*, 2021). The term "hydrofabric" is used within the NWM jargon to describe the subset of hydrography composed of the geospatial datasets required for hydrologic modeling including but not limited to stream networks, catchments, channel properties, and elevation data. The NWM provides stream forecasts at these hydrofabric segments using the Muskingum-Cunge method to reduce computational requirements of a continental scale model but fails to consider backwater dynamics (Bedient et al., 2008; Ponce & Changanti, 1994; Gochis et al., 2021). The need for high resolution FIM at 10 m or better requires additional post-processing from the principal output of the NWM which is forecast river discharges at the reach scale. The use of a 2-dimensional (2D) hydrodynamic model across a continental-scale and high

129 spatial resolutions is very cost prohibitive especially in an operational setting. The Height

130 Above Nearest Drainage (HAND) terrain model is one such technique that can be used,

131 along with synthetic rating curves (SRC), to convert 1-dimensional (1D) riverine discharges

132 to stages, and finally to inundation extents and depths.

1.3 Height Above Nearest Drainage

134 HAND normalizes topography along the nearest drainage path and it has been demon-

135 strated to be a good proxy and indicator of a series of important environmental condi-

136 tions including soil environments, landscape classes, soil gravitational potentials, geo-

137 morphologies, soil moisture, and groundwater dynamics (Rennó et al., 2008; A. Nobre

138 et al., 2011). A. D. Nobre et al. (2016) showed evidence for utilizing the drainage nor-

139 malizing HAND dataset as a proxy for flood potential to make static flood inundation

140 maps from known stages. The terrain index also provides additional utility in the ob-

141 servation of riverine flood inundation mapping from remote sensing especially in areas

142 of high electromagnetic interference such as vegetated and anthropogenic areas (Aristizabal

143 et al., 2020; Shastry et al., 2019; Huang et al., 2017; Twele et al., 2016; Aristizabal &

144 Judge, 2021). Zheng, Tarboton, et al. (2018) developed a methodology for determining

145 stage-discharge relationships known as SRCs by sampling reach-averaged parameters from

146 HAND datasets and inputting into the Manning's equation (Gauckler, 1867; Manning

147 et al., 1890). This collection of methods, coupling HAND with SRCs, have been exper-

148 imented with and compared to other sources of FIM including engineering scale mod-

149 els, in-situ observation, and remote sensing based observation with solid results in large

150 spatial scale applications (Godbout et al., 2019; Johnson et al., 2019; Garousi-Nejad et

151 al., 2019; A. D. Nobre et al., 2016; Afshari et al., 2018; Zheng, Maidment, et al., 2018;

152 Teng et al., 2015, 2017; Zhang et al., 2018).

1.4 HAND's Assumptions and Limitations

154 HAND operates on many underlying assumptions since it can only be used as an

155 inundation proxy or no physics model and thus, not a true hydrodynamic inundation model

156 (A. D. Nobre et al., 2016; Y. Y. Liu et al., 2016; Y. Liu et al., 2020). HAND, to our knowl-

157 edge, has only been applied to natural, inland, and riverine inundation applications thus

158 it is also missing pluvial, coastal, ground water, and dam break components among other

159 possible sources of flooding. Additionally, in order to flood an area, HAND assumes all

160 areas eligible for inundation must drain to some nearest stream line which is used for catch-

161 ment allocation and relative elevation calculation (A. D. Nobre et al., 2016; A. Nobre

162 et al., 2011; Y. Y. Liu et al., 2016; Y. Liu et al., 2020; Maidment, 2017; Garousi-Nejad

163 et al., 2019; Zheng, Tarboton, et al., 2018; Zheng, Maidment, et al., 2018; Johnson et

164 al., 2019; Rennó et al., 2008). Stream thalweg networks must also collectively drain to

165 a singular outlet point for a given processing region (A. D. Nobre et al., 2016; Zheng,

166 Maidment, et al., 2018; Rennó et al., 2008). Since elevations don't naturally do this, they

167 must undergo a long series of hydro-conditioning processes to enforce monotonically de-

168 creasing elevations across an entire processing unit along with hydrologically correct di-

169 rections of flow (A. D. Nobre et al., 2016; A. Nobre et al., 2011; Y. Y. Liu et al., 2016;

170 Y. Liu et al., 2020; Donchyts et al., 2016; Rennó et al., 2008). The level of digital ele-

171 vation map (DEM) manipulation required to enforce this assumption can be substan-

172 tial depending on the region and can be a significant source of error. The drainage en-

173 forcing assumption also interacts with an inability to properly account for fluvial inun-

174 dation in regions of DEM depressions that lack natural drainage to riverine areas (A. D. No-

175 bre et al., 2016; Rennó et al., 2008).

176   When used for FIM applications, HAND assumes only fluvial inundation sourced

177 from its nearest drainage line is accounted for (A. D. Nobre et al., 2016; McGehee et al.,

178 2016). Catchments are independent of one another for FIM purposes meaning a reaches'

179 stage value is only used to threshold the HAND values within its respective catchment

180 (Y. Y. Liu et al., 2016; Zheng, Tarboton, et al., 2018; Zheng, Maidment, et al., 2018).

181 This assumption plays to the "Nearest Drainage" term in HAND and creates a signif-

182 icant limitation within HAND for FIM applications (Zhang et al., 2018; McGehee et al.,

183 2016; Li et al., 2020; A. D. Nobre et al., 2016). At the junction of high stream order and

184 high flow rivers with lower flow tributaries, there can be a lack of inundation extents ex-

185 hibited which is known colloquially in the forecasting community as the "catchment bound-

186 ary problem". The academic community has somewhat referenced this issue before but

187 it has been characterized more as a problem with the stream delineation process that

188 comes from thresholding the drainage accumulation maps (A. D. Nobre et al., 2016; Li

189 et al., 2020). Later in this study, we will re-introduce this problem and demonstrate how

190 we initialize with a stream network (that of the NWM's) and thus avoid having to thresh-

191 old accumulations to some arbitrary value to define stream networks. We illustrate how

192 computing HAND independently for stream lines of unit stream order can significantly

enhance FIM performance by accounting for multiple sources of fluvial inundation that may exist in certain regions and flow scenarios.

1.5 HAND Implementations

Due to significant advances in high performance computing (HPC) and large scale high resolution DEMs such as the 3D Elevation Program (3DEP) seamless at the 1/3 arc-second (approximately 10 m depending on latitude) scale, HAND has been implemented into software for large-scale, continental computation. As part of the OWP's Innovators Program and NWC's Summer Institute, the National Flood Interoperability Experiment (NFIE) generated FIM hydrofabric (will be used interchangeably with the datasets produced by HAND) rapidly on a HPC (Maidment, 2017; Y. Y. Liu et al., 2016). NFIE used open-source dependencies including the Terrain Analysis Using Digital Elevation Models (TauDEM) (Tarboton, 2005) and the Geospatial Data Abstraction Library (GDAL) (Warmerdam, 2008) to compute HAND for the Continental United States (CONUS) at 331 Hydrologic Unit Code (HUC) 6 processing units in 1.34 central processing unit (CPU) years. By allocating 31 nodes at 20 cores per node for a total of 620 available cores to the overall operation, it enabled the production to finish up in 36 hours consuming 3.2 terrabyte (TB) of peak memory and 5 TB of total disk space. Originally, NFIE utilized the NHD Medium Resolution (MR) to etch or burn flowlines prior to further conditioning but more recent work has advanced this to the more current NHDPlus High Resolution (NHDPlusHR) which better agrees with the 10 m DEM from the NHDPlusHR program (Y. Liu et al., 2020). The original NFIE dataset was employed by the NWC as an unofficial demonstration to produce forecast FIM from the NWM for additional guidance in hydro-blind regions. Further work by Djokic (2019), implemented a series of improvements to HAND including equidistant reaches, updates to use with NHDPlusHR hydrography, and AGREE DEM reconditioning (Hellweger & Maidment, 1997) into an ESRI Arc-Hydro workflow with use in ArcGIS. More notably the software added the ability to derive HAND on both the NWM FR and MS stream networks to consider multiple sources of fluvial inundation along high impact rivers of primary forecasting concern.

Related to these efforts, the United States Geological Survey (USGS) has invested in relative elevation HAND-like methods via work in the GIS Flood Tool (GFT) that also uses SRCs with cross-sections for stage-discharge relationships (Verdin et al., 2016).

Additional investment by Petrochenkov (2020) was able to successfully scale this approach by transitioning the method to open-source Python source code (PyGFT) and implementing novel interpolation methods to help address some of the catchment boundary discontinuities discussed more in this paper. In addition to the domestic work done in the US, some studies have expanded upon HAND to cover global domains at 30 m resolutions (Yamazaki et al., 2019; Donchyts et al., 2016).

1.6 Office of Water Prediction Flood Inundation Mapping

In order to mitigate the ever increasing threat of flooding to life and property, an operational capability is required to extend NWM streamflow forecasts to river stages, inundation extents, and inundation depths. OWP FIM is introduced here as a continental scale capability that generates these products at high spatial and temporal resolutions. Here we introduce OWP FIM that utilizes a few of the latest techniques in HAND based FIM oriented for use with the NWM in continental scale operational forecasting settings. Within the operational framework of OWP FIM, we introduce research demonstrating how FIM performance skill with HAND can be improved by discretizing stream networks into units of an effective unit Horton-Strahler stream order (Horton, 1945; Strahler, 1952, 1952) for HAND computation contexts. Previous authors dating back to the first HAND for FIM work by A. D. Nobre et al. (2016) have noted a sensitivity of mapping skill to the stream accumulation threshold which is closely related to stream density and the maximum Horton-Strahler stream order (or simply stream order) of the processing unit employed (Zhang et al., 2018; McGehee et al., 2016; Li et al., 2020). Here we demonstrate how reducing a HAND processing unit's stream network into discrete level paths (LPs) of singular, effective stream order, can enhance FIM skill by accounting for multiple possible sources of fluvial inundation. This capability is introduced progressively as MS (whose network represents about 4% of FR network) and to a higher degree Generalized Mainstems (GMS) (covers entire FR network) which will be explained later on. The following methods and results describe the work in more detail and demonstrate its efficacy in producing enhanced FIM for the NWM.

2 Materials and Methods

OWP FIM is a fully operational pipeline of software tools to help acquire datasets, cache hydrofabrics, produce FIMs, and evaluate results. Figure 2 gives a high level overview of the methodology used in OWP FIM and in this study. Input data from multiple sources are preprocessed (not illustrated) and then subset to processing areas based on the model used, FR, MS, or GMS. The standard processing unit of OWP FIM is a HUC8 and the entire NWM FR stream network is used for enforcment. Later, we explain how only the NWM MS stream network is used for the MS version of HAND, while for GMS, the FR stream network is discretized into LPs before computing HAND. A series of hydro-conditioning steps enforces the location of stream lines, monotonically decreasing elevations, excavated bathymetry, stream thalweg breaching, and levee enforcement. After a DEM suitable for HAND's assumptions is conditioned, the FIM hydrofabric is generated including stream network, catchments, HAND, SRCs, and cross-walk table. The FIM hydrofabric is defined as the datasets required to make an inundation map from discharges including the relative elevation model (REM) or HAND grid, the catchments in vector and raster form, and the hydro-table (contains SRC and cross-walk information). In operational circumstances, the NWM streamflows are used in conjunction with the FIM hydrofabric to derive forecast FIMs. However for evaluation purposes, we use the streamflows from the cross-sections of our benchmark model. As later discussed in Section 2.6, independent FIMs from multiple fluvial sources are mosaiced together. For the case of MS, two sources are mosaiced together (FR and MS) while for GMS the inundation from every LP is composited together. The evaluation FIM extents are compared to the extents of the benchmark model and metrics are computed.

NHDPlus HR
Burnline Events
VAA
WBD
DEM's (elev_cm.tif)

NWM Hydrofabric
FR Streams
MS Streams
Catchments
Waterbodies

USACE – NLD
OSM – LandSea Border
GLHD – Land Great Lakes Border

Subset Data
HUC8 – FR, MS
Level paths - GMS

Hydro-Conditioning
1. Stream Network Enforcement (Pruning, Burning, Excavating)
2. Levee Enforcement
3. Depression Filling
4. Stream Thalweg Breaching

Deriving FIM Hydrofabric
1. Flow Directions and Flat Resolution
2. Deriving FIM Stream Network
3. Catchments
4. Height Above Nearest Drainage
5. Synthetic Rating Curves
6. Cross-walking with NWM

FIM
1. FR Inundation Mapping
2. MS FIM = FR + MS
3. GMS FIM By Level Path

NWM Stream Flows

InFRM
HEC-RAS Cross Sections

FIM Hydrofabric
Stream Network
Catchments
HAND
SRCs
Cross-walk Table

Operational FIM

Evaluation FIM

Evaluate
Compute agreement raster
Aggregate primary stats
Compute secondary stats

Metrics

InFRM
HEC-RAS FIM Extents

Legend   Input Data   Data Product   Process

Figure 2. Methodology overview detailing high level steps followed in the study. The flow chart begins with the input data organized by source. Subsetting the data into processing units depends on which model is being considered. FR utilizes the entire NWM stream network processed at HUC8 processing areas. MS only computes HAND using the NWM stream at or downstream of legacy forecasting points. The resulting inundation from the MS HAND is eventually layered with the FIM from FR HAND to account for high levels of inundation contributed by the mainstem. Generalized Mainstem (GMS) discretized NWM streams into level paths (LP) then computes HAND and the FIMs independently only to mosaic them later. This better accounts for multiple possible sources of fluvial inundation. The dotted lines denote the use of NWM streamflow forecasts to produce operational FIM but not used in this study. All acronyms used in the figure are defined in the paper.

2.1 Software Dependencies and Architecture

OWP FIM exclusively utilizes free and open source software dependencies including Python 3, GDAL, TauDEM, Geographic Resource Analysis Support System (GRASS), GNU Parallel, and MPICH (P. C. Team, 2019; contributors, 2020; Tarboton, 2005; G. D. Team, 2020; Tange, 2015; Amer et al., 2021). Within the Python 3 ecosystem, many common packages are employed including but not limited to RichDEM, GeoPandas, Rasterio, Rasterstats, and Numba (Barnes, 2018; Jordahl, 2014; Lam et al., 2015). To simplify setup and enhance portability across host operating systems, OWP FIM packages all dependencies up in a Docker image (Merkel, 2014). A user only needs to install Docker on their host machine and build the image from the provided recipe. Source code is made available for this project on GitHub where a user could consult the Readme.md page for more information on how to acquire the datasets and reproduce the pipeline (Aristizabal et al., 2022b).

2.2 Datasets

Data sources used within OWP FIM are publicly available from a variety of government sources including the USGS, NWC, Federal Emergency Management Agency (FEMA), and US Army Core of Engineers (USACE) to enhance reproducibility and collaboration among government, academia, and industry. Instructions for accessing data processed for OWP FIM are provided on the project's GitHub page via an Amazon Web Services (AWS) S3 bucket furnished by the Earth Science Information Partners (ESIP) (Aristizabal et al., 2022a). The National Hydrography Dataset Plus High Resolution (NHDPlusHR) Beta Version is the latest hydrography dataset used for land surface hydrologic modeling in the US (Moore et al., 2019). We utilized a series of data products from the NHDPlusHR including the BurnLineEvents (*NHDPlusHR GDB*, 2021), Value Added Attributes (VAA) (*NHDPlusHR GDB*, 2021), Water Boundaries (WBD) or HUC Layers (*NHDPlusHR WBD*, 2021), and the DEM elevation rasters (*NHDPlusHR DEM*, 2021). These BurnLines used in conjunction with the hydrofabric of the NWM V2.1 to help define flowlines for OWP FIM while the NWM hydrofabric is also used to define reservoirs for exclusion and catchments to cross-walk against for forecasting purposes (*NWM Hydrofabric V2.1*, 2021). For enforcing levee data, the USACE NLD is used to burn feature elevations into DEMs (ENGINEERS, 2021). Since NHDPlusHR datasets extend beyond land borders into sea and Great Lake regions, we used the land-sea border from

OpenStreetMap (OSM) (*Water polygons*, 2021) and the land-lake border from Great Lakes Hydrography Dataset (GLHD) (*GLHD*, 2020) to exclude those areas from production of FIMs. Additionally, the Base Level Engineering (BLE) datasets within FEMA Region 6 spanning parts of nine states including Colorado, New Mexico, Texas, Oklahoma, Kansas, Arkansas, Louisiana, Missouri and Mississippi at two recurrence intervals, 1% (100 year or yr) and 0.2% (500 year or yr), are used for validation in this study and furnished by the Interagency Flood Risk Management (InFRM) consortium (*Base Level Engineering (BLE) Tools and Resources*, 2021; *estBFE Viewer*, 2021). These BLE datasets are provided at the watershed scale (HUC8) utilizing best available DEMs and simulations from the Hydrologic Engineering Center's River Analysis System (HEC-RAS) model (USACE, 2022). The full input datasets presented by source are listed in Table 1. Areas with all the required data (from the NWM and the USGS) are labeled as the FIM domain which includes 2,188 HUC8s for the FR and GMS networks and 1,604 HUC8s for the MS method. These methods will be explained in more detail later. An enhancement of OWP FIM over previous HAND based FIM versions is the support for Hawaii and Puerto Rico which are expansion domains in the NWM V2.0 and V2.1, respectively.

2.3 Hydro-conditioning

The DEM is subject to a series of hydro-conditioning procedures to enhance its suitability for riverine flood inundation mapping with HAND. These techniques are specific for making OWP FIM and differ from the conditioning methods used by the NHDPlusHR Beta (Moore et al., 2019). HAND inherently requires all areas eligible for inundation to drain to the designated drainage network. So to satisfy this requirment, DEMs must undergo significant manipulation. In other words, all areas within a given processing unit for HAND must have monotonically decreasing elevations to enable eligiblity for flooding. Hydro-conditioning is implemented to obtain many objectives including enforcing the location of hydrologically relevant features such as flowlines, lakes, or drainage divides whether natural or anthropogenic. It can also be used to simulate more accurate bathymetry which is not accounted for in the 10 m DEM (Gesch et al., 2002).

Specifically within the context of OWP FIM, the hydro-conditioning operations that take place in sequential order are presented. Prior to any hydro-conditioning, all input datasets must be subset from their original spatial domain scales into the processing units of size HUC8. The subsetting is done by spatial query for the cases of the levees, DEM,

Table 1. Data sources, names, descriptions, and citations.

Source	Name	Description	Citations
USGS	NHDPlusHR BurnLineEvents	Stream lines used by NHDPlusHR for hydro-enforcement.	(*NHDPlusHR GDB*, 2021)
USGS	NHDPlusHR Value-Added Attributes	Database of additional attributes associated with the BurnLineEvents that enhance navigation, analysis, and display.	(*NHDPlusHR GDB*, 2021)
USGS	NHDPlusHR DEM	DEM used for NHDPlusHR at 1/3 arc-second (10 m) spatial resolution and vertical units in centimeters.	(*NHDPlusHR DEM*, 2021)
USGS	NHDPlusHR WBD	Water Boundaries (WBD) or HUCs used for spatial processing units.	(*NHDPlusHR WBD*, 2021)
NOAA-OWP	NWM Streams	Stream network center lines used by NWM for routing and forecasting adapted from NHDPlus V2 NHDFlowline_Network feature class.	(*NWM Hydrofabric V2.1*, 2021)
NOAA-OWP	NWM Catchments	Surface drainage area corresponding to each reach in the NWM adapted from NHDPlus V2 Catchment feature class.	(*NWM Hydrofabric V2.1*, 2021)
NOAA-OWP	NWM Waterbodies	Waterbodies considered by the NWM as reservoirs or lakes adapted from NHDPlus V2 NHDWaterbody feature class.	(*NWM Hydrofabric V2.1*, 2021)
USACE	NLD	Levee database of locations and elevations.	(ENGINEERS, 2021)
OSM	Land-Sea Border	Border of land and sea.	(*Water polygons*, 2021)
GLHD	Land-Great Lakes Border	Border of land and Great Lakes.	(*GLHD*, 2020)
InFRM	Cross-Sections	HEC-RAS 1D cross-sections used for modeling in BLE datasets. Includes discharges for 1% and 0.2% recurrence interval events.	(*estBFE Viewer*, 2021)
InFRM	Flood Inundation Extents	Inundation depths produced by InFRM BLE HEC-RAS 1D for 1% and 0.2% recurrence interval events.	(*estBFE Viewer*, 2021)

340 and NWM hydrofabric while the NHDPlusHR BurnLineEvents are subset via attribute

341 query for the given reach code's membership in the processing unit. Hydro-conditioning

342 raster operations take place on buffered boundary definitions to avoid edge contamina-

343 tion and effects (Lindsay & Seibert, 2013).

2.3.1 Stream Network Enforcement

345 The location of the stream network is enforced to ensure general agreement with

346 the NWM network which is used for forecasting the streamflow inputs. The NHDPlusHR

347 Beta BurnLineEvent layer is used to enforce stream locations in the NHDPlusHR work-

348 flow and best agrees with thalweg locations in the DEM used so it is also used here for

349 hydro-enforcement (Moore et al., 2019).

350 However, to better match the drainage density of the NWM FR V2.1 stream net-

351 work, which is based on the NHDPlus V2, the BurnlineEvents are pruned utilizing a near-

352 est neighbor search around the NWM flowlines. Headwater points are first derived for

353 the NWM FR V2.1. For every NWM headwater point, the nearest NHDPlusHR point

354 is selected and placed into a set while those excluded are discarded. Only the nearest

355 point on the NHDPlusHR is used so any portion of the NHDPlusHR network upstream

356 of this nearest point is discarded to avoid extending inundation too far above the mod-

357 eling domain. The points in this nearest neighbor set are then traversed downstream.

358 Any headwater portion in the NHDPlusHR or any other stream not traversed are pruned

359 away to better match the resolution and spatial locations of the NWM stream network

360 and its headwater points. The resulting pruned NHDPlusHR stream network gets hydro-

361 enforced in subsequent operations. This procedure is best illustrated in Figure 3 which

362 shows that the pruned NHDPlusHR network corresponds to the NHDPlusHR network

363 at or downstream of NWM V2.1 headwater locations only. Additionally, the NHDPlusHR

364 pruned headwaters are later used for seeding a new FIM drainage network that best agrees

365 with the DEM after all hydro-conditioning takes place. This results in a stream network

366 that has the same density as the NWM V2.1 flowline network but utilizes the locations

367 of the NHDPlusHR Beta BurnLineEvents.
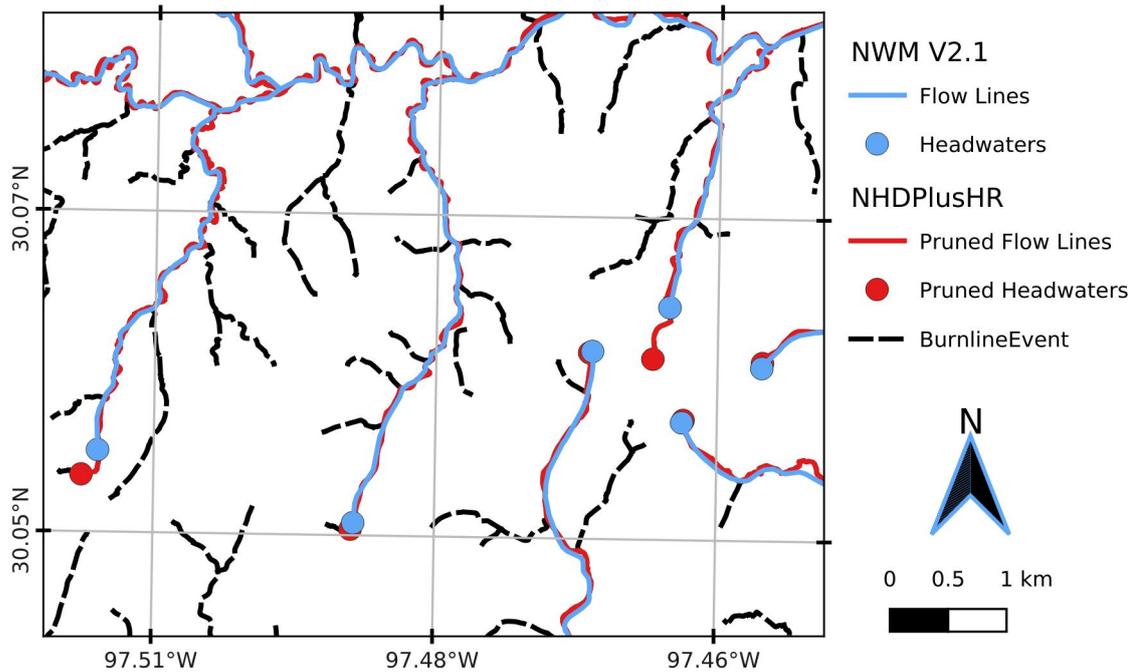
Figure 3. This figure illustrates some of the datasets that result from the pruning of NHD-PlusHR Beta BurnlineEvents (dotted black) to the stream density of the NWM FR V2.1 density (blue). The stream network used for forecasting, NWM FR V2.1, is of lower stream density than that of the NHDPlusHR which has better agreement with the thalweg locations in the DEM used. Thus, we opt to prune the NHDPlusHR network to match the general location and density of the NWM network. The nearest neighbor segment in the NHDPlusHR of each NWM headwater locations and the nearest point on that segment is determined to match the closest point to that of the NWM headwater. These points are then traversed downstream and any segments not traversed are pruned away. The resulting stream network (red) matches the drainage density of NWM V2.1 while corresponding spatially with the NHDPlusHR BurnlineEvents.

The pruned stream network is then utilized to hydro-enforce the DEM with a methodology developed by Hellweger and Maidment (1997) known as the AGREE DEM Surface Reconditioning System. The AGREE algorithm seeks to burn artificially deep thalweg elevations by a uniform value known as sharp drop. The modification continues by excavating an area of a given buffer distance from the thalweg by a depth proportional to the distance from the channel given by the smooth drop and buffer distance. The resulting enforcement of the thalweg and general bathymetric region results in a cross-section

resembling an inverted triangular notch shape with a significantly lower elevation along the thalweg line only. In total, the AGREE algorithm requires three parameters including the buffer distance, smooth drop, and sharp drop which were set to fixed values of 70 m, 1000 m, and 10 m, respectively, but available to the user via the parameter file. While the values for these parameters are critical to the inundation extents produced, especially for lower flow rates where bathymetric information has more influence, finding their optimal values for OWP FIM was not done since it was out of the main scientific scope of this article. Using the AGREE method as opposed to simple thalweg burning techniques helps prevent distortions in the delineation of streams as well as the catchment boundaries (W. Saunders & Maidment, 1995; W. K. Saunders & Maidment, 1996; Mizgalewicz & Maidment, 1996; Hellweger & Maidment, 1997; Quenzer, 1998; Baker et al., 2006). Baker et al. (2006) noted AGREE produced satisfactory results when compared to other enforcement techniques especially when computational costs are considered. Downsides to the technique include the possibility of exhibiting parallel streams where the burned stream and real stream are both represented (Hellweger & Maidment, 1997; W. Saunders, 1999) and some distortion of the catchment boundaries can also be observed (W. Saunders, 1999; W. K. Saunders & Maidment, 1996). Some of these drawbacks are addressed by additional conditioning techniques applied later on.

2.3.2 Levee Enforcement

Coarse DEM's at 10 m, 30 m, and higher resolutions can lack sufficient representation of fine grain features such as embankments, flood walls, and closure structures (Arundel et al., 2018; Dobbs, 2010; Wang & Zheng, 2005; Sanders, 2007). In order to better represent the influences of these features upon hydraulics and inundation extents, the National Levee Database (NLD) published by USACE was used to enforce elevations within the 1/3 arc-second DEM. The elevations found in the NLD are burned onto the DEM if those elevations were found to exceed those already in place.

2.3.3 Depression Filling

Local depressions are naturally occurring features of a DEM but must be addressed if a connected drainage network with continuous catchments are to be derived for flood modeling purposes with HAND. The partially conditioned DEM was removed of depressions by filling areas with pits while preserving the stream and levee information pre-

viously enforced. Priority-Flood developed by Barnes et al. (2014b) is an algorithm for filling said depressions and shown to have improved performance over early works in the field by Jenson and Domingue (1988) implemented in Tarboton (2005) as well as Planchon and Darboux (2002). The depression filling algorithm used in our pipeline is a Priority-Flood variant developed by (Zhou et al., 2016) with enhanced single-thread performance and a time complexity of O(n log n) for floating point grids. This performance was enabled by limiting the processing queue with a region-growing method to exclude many of the slope cells (Zhou et al., 2016). The depression filling technique employed here does leave the existence of flat regions where pits previously existed thus later requiring the need for resolving these flats. The enhanced variant of Priority-Flood is implemented and made available by Barnes (2018) and Zhou et al. (2015).

2.3.4 Stream Thalweg Elevation Conditioning

Thalweg elevations are critical components of relative elevation based inundation mapping thus much is performed to ensure the best available, monotonically decreasing, elevations are derived prior to the normalizing of elevations. Work on the AGREE DEM method from several authors have illustrated that the AGREE DEM method does not prevent situations where the burned thalweg and the thalweg endemic to the DEM run parallel to one another (Hellweger & Maidment, 1997; Baker et al., 2006; W. Saunders, 1999; W. K. Saunders & Maidment, 1996; Quenzer, 1998; W. Saunders & Maidment, 1995). These works observe that the artificial elevations enforced by the hydrographically based stream network and AGREE DEM disagree with those naturally occurring in the native DEM. In order to mitigate this documented issue, the normalized excavation algorithm (W. Saunders, 1999) is used to seek a zonal (nearest neighbor) elevation minimum on the original, unconditioned DEM for each thalweg pixel. Each zone is defined as the thalweg's pixel nearest neighborhood within a maximum distance of 50 m. The zonal minimum is computed for each thalweg pixel zone and the minimum is used to replace the existing thalweg elevation value. This step essentially enforces an estimate of the native DEM thalweg elevations onto the sharp drop enforced thalweg elevations from the AGREE procedure.

The next step involves conditioning these local minimums along the thalweg to enforce monotonically decreasing thalweg elevations for FIM. Garousi-Nejad et al. (2019) proposed an algorithm that breaches stream thalweg pixel elevations in a depth first man-

438 ner. This procedure was found to increase the Critical Success Index (CSI) of resulting

439 FIMs from HAND and is employed in OWP FIM to enforce monotonically decreasing

440 elevations with thalweg pixel networks.

2.4 Deriving FIM Hydrofabric

442 The FIM Hydrofabric is defined here as the collection of geospatial datasets that

443 are used for converting NWM discharges into inundation extents. These datasets include

444 the HAND or relative elevation model (REM) raster, reach-level catchments raster/polygons,

445 DEM-derived streamlines, SRCs, and cross-walk table. The following sub-sections de-

446 scribe how the subset and hydro-enforced geospatial datasets are converted into the FIM

447 hydrofabric.

2.4.1 Flow Directions and Flats Resolution

449 To facilitate the generation of a connected stream network and its associated catch-

450 ments from the conditioned DEM, the depression-filled DEM is used to derive connec-

451 tivity in the form of D-8 flow directions. D-8 seeks to allocate a drainage direction for

452 every pixel based on the adjacent eight pixel neighborhood with the steepest slope (O'Callaghan

453 & Mark, 1984). The horizontal component of slope is defined as one for the four neigh-

454 boring pixels in the main cardinal directions while the intercardinal pixels are designated

455 a horizontal component of $\sqrt{2}$ by means of the Pythagorean theorem. Flow directions

456 are derived for non-depression filled regions trivially with the above procedure but to de-

457 fine connectivity for every grid cell the remaining flats corresponding to depression-filled

458 cells must be resolved.

459 Flat resolution from flats endemic to the DEM or from depression filled regions is

460 a costly, non-trivial procedure which was originally addressed by Garbrecht and Martz

461 (1997) where flats are resolved by incrementing elevations iteratively. Software imple-

462 mentations have developed means to partition the problem and resolve flats iteratively

463 with communication across processes (Tarboton et al., 2009; Tesfa et al., 2011; Wallis

464 et al., 2009; Tarboton, 2005). The excessive iteration and communication leads to poor

465 computational performance which motivated further work on how to optimize flat res-

466 olution (Survila et al., 2016; Barnes et al., 2014a). The established literature in this niche

467 field of hydrology discusses how prevalent flats can be in given study areas and how dif-

ficult the problem is from both computational and hydrologic stand-points (Garbrecht & Martz, 1997; Tarboton et al., 2009; Tarboton, 2005; Survila et al., 2016; Barnes et al., 2014a; Tesfa et al., 2011; Wallis et al., 2009). OWP FIM utilized a CyberGIS implementation of the D-8 flow direction algorithm with the accelerated resolution of flats which we found to be very efficient and effective (Survila et al., 2016; Y. Liu et al., 2016).

2.4.2 Deriving FIM Stream Network

The derivations of relative elevations and catchments from the newly conditioned DEM involves re-deriving a new, DEM based, FIM stream network. The FIM stream network is of similar drainage density as the NWM V2.1 network and fully converges at all junctions leaving no divergences in the network. This is accomplished by using the seed points generated from the stream network enforcement process (Section 2.3.1). These seeds points are pruned headwater locations of the NHDPlusHR Beta BurnlineEvents layer that spatially correspond to the headwater definitions in the stream network of the NWM V2.1. Feeding the seed points and previously computed flow directions into flow accumulation methods (Wallis et al., 2009; Tarboton, 1997, 2005) yields a stream link accumulation raster that can be converted to a vector file for further processing.

Each stream link in this derived FIM stream network is split into equidistant reaches of 1.5 km in length which is a user exposed parameter. Stream links are defined here as segments of rivers discretized by junctions with other NWM river segments. Stream links are then further segmented at NWM lakes and HUC8 boundaries. Discretizing at NWM lakes isolates reaches and catchments associated with lakes and reservoirs to avoid mapping them using the Manning's equation and could potentially enable volume based mapping in the future as a feature enhancement. Based on previous research, splitting each remaining stream link into equidistant reaches not to exceed a parameterized value of 1.5 km helps improve SRC and mapping skill (Garousi-Nejad et al., 2019; Godbout et al., 2019; Zheng, Maidment, et al., 2018). Small reaches can lead to unrealistic variances in channel geometries while oversized reaches can lead to grouping too much slope variance into one discretization of the stream network. Short stream segments that are introduced as a result of forced network breaks due to reservoir, levee, or HUC boundaries inherit the SRC properties of the upstream or downstream segment, depending on the topology. Section 2.4.5 details the derivation of the SRC and the dependence on channel length. Additionally every reach (and later catchment) is assigned a globally unique

identifier based on the HUC8 membership. This stream network is important since it drives the HAND calculation and derivation of catchments.

2.4.3 Catchments

Catchments were derived using the D8 connectivity established by O'Callaghan and Mark (1984). Outlet points are set at the pixel center points of the delineated stream lines explained in Section 2.4.2. The outlets act as root nodes in a tree structure and the connectivity is traversed to derive the contributing, nearest drainage region for each outlet point. Two sets of catchments are derived, one set of catchments denotes the unique drainage region for each thalweg pixel which is used for relative elevation calculation. The other catchments are derived for the drainage region for each stream reach as defined in Section 2.4.2.

2.4.4 Height Above Nearest Drainage

Once the pixel level catchments are derived, the final relative elevations can be computed. The elevation of every thalweg pixel is subtracted from the elevations of the nonthalweg pixels within the same, corresponding pixel-level catchment described in Section 2.4.3. The DEM used for this operation is the DEM resulting from the thalweg conditioning procedures described in Section 2.3.4. Outside of the excavated channel from the AGREE DEM method, the native non-drainage enforced elevations are used to reduce sources of error in relative elevations due to pit filling (Djokic, 2019). Any negative values resulting from this subtraction with native elevations are replaced by zero. Again, HAND assumes and requires processing areas to drain thus have monotonically decreasing elevations with hydrologically correct flow directions all leading to a singular outlet point. While this is required for the generation of DEM-derived catchments and stream lines, it is not necessarily required for the computation of the relative elevations. Since the use of hydro-conditioning processes to fit the drainage requirement for HAND can be extensive, we found it more fitting to use the native elevations this final HAND computation thus avoid the use of manipulated values that fit modeling assumptions.

2.4.5 Synthetic Rating Curves

A method for converting forecast river discharges from the NWM to stages or river depths at the reach scale is necessary for producing FIMs with HAND. For 1D hydrodynamic models such as the routing methods in the NWM, the typical procedure is to establish the stage-discharge relationship by sampling data from the DEM to derive a SRC at discrete cross-sections (Quintero et al., 2021; Di Baldassarre & Claps, 2011). For this application, we utilized the reach averaged approach for developing SRCs (Zheng, Tarboton, et al., 2018). The reach averaged approach seeks to sample the geometry parameters in the Manning's equation (Gauckler, 1867; Manning et al., 1890) on a reach scale then dividing those by length. Previously not reported in literature to our knowledge in this form, the reach averaged Manning's formula is derived to be

$$Q(y) = \frac{1}{n} \frac{V(y)^{5/3} S^{1/2}}{L B(y)^{2/3}} \tag{1}$$

where Q is discharge at stage y, n is the Manning's n roughness coefficient, V is volume at y, S is channel slope, L is the along-flow reach length, and B is wetted bed area at y. Q, V, and B are taken at specific y values so are more formally written as $Q = Q(y)$, $V = V(y)$, and $B = B(y)$, respectively. All units are international (SI) given the one in the numerator above n. The reach averaged method has been compared to rating curves from HEC-RAS and USGS gages yielding comparable results for estimating the river bottom elevation profile, channel width at given stages, and stage-discharge relationships (Zheng, Tarboton, et al., 2018). The reach averaged geometry parameters including number of wet cells, bed area, and volume are sampled from the thalweg conditioned AGREE DEM using TauDEM's catchhydrogeo utility. Using the split reaches described in Section 2.4.2, the channel slope is sampled from the thalweg conditioned DEM at the end points of the reaches while the same reaches are used to calculate the reach length. While the AGREE DEM is subject to hydro-conditioning processes, it does introduce some notion of bathymetry estimation that the native DEMs lack while being sensitive to additional parameters that could yield further errors in the FIM. We leave this issue open in this study and elaborate on needs with respect to bathymetry and Manning's n values in the Discussion section (Section 4).

Setting of the Manning's n roughness coefficient has precedent in previous continental-scale FIM (CFIM) studies (Maidment, 2017; Y. Y. Liu et al., 2016; Y. Liu et al., 2020; Djokic, 2019; Garousi-Nejad et al., 2019; Zheng, Maidment, et al., 2018) with two noted

values of 0.05 and 0.06 for NFIE and Djokic (2019) respectively. These values are applied universally to the entire forecasting domain across space, time, and discharge profiles. We note significant opportunity to enhance CFIM skill by better localizing Manning's n according to available data including but not limited to land cover, land use, stream order, stream geometry, drainage area, reach length, and discharge percentiles (Garousi-Nejad et al., 2019; Johnson et al., 2019). For now and for the purpose of this study, we examine the SRCs with Manning's n set to both 0.06 and 0.12 which we hope will shed some light on the sensitivity of this parameter to HAND based FIMs. After all the parameters to the Manning's equation have been determined with either hydrofabric sampling or user parameterization, we select 84 stage values (y in Eq. 1) from 0 to 25 meters in depth at a third of a meter increments to calculate the discharge values for each stage value.

2.4.6 Cross-walking with NWM Stream Network

The DEM based stream network derived in Section 2.4.2 must be associated with NWM reach identifiers so that a discharge can be converted to stage and later inundation extents and depths. For the FR version of HAND, we overlap the reach catchments derived in Section 2.4.3 with the NWM catchments matching the ID of the NWM catchment that most overlaps the derived catchment for HAND. For two subsequent HAND methods, MS and GMS, discussed in Sections 2.5.1 and 2.5.2, respectively, we find the mid-point of the derived stream reach line described in Section 2.4.2 and find the NWM catchment that contains the mid-point. Additionally, only relevant catchments from the NWM for the given LP are selected for cross-walking for methods in Sections 2.5.1 and 2.5.2. While these conflation methods are approximate, they can lead to some substantial errors which will be discussed more in Section 4.

2.5 Stream Order Reduction

As previously discussed, HAND based FIMs are subject to many assumptions and limitations in order serve as a suitable inundation proxy for large scale, high resolution domains. HAND produced FIMs are limited in only providing inundation sourced from the nearest drainage line, however, depending on flow conditions and topography, a given area may have multiple contributing fluvial sources of surface water inundation. The forecasting community, in reviewing HAND, have noted significant negative effects at the

confluence of lower flow tributaries with higher flow rivers for which the phrase "catchment boundary issue" has been termed. In previous studies, FIM skill has been shown to be sensitive to the drainage density of the stream network employed as the datum for HAND which is closely related to the maximum Horton-Strahler stream order of the network (Zhang et al., 2018; McGehee et al., 2016; Li et al., 2020; A. D. Nobre et al., 2016). This sensitivity is partly in due to the limitation that catchment boundaries place on inundation extents where only the nearest drainage line can source inundation for any particular area.

Figure 4 illustrates the exact situation our solution proposes to address where two tributaries converge with a higher order stream segment. An actual map with OWP FIM is generated using the NWM full-resolution stream network and compared with a FEMA 100 yr extent (see Section 2.7 for more details) showing significant under-prediction in inundation extent. The higher discharge along the main segment in Figure 4 of 1,900 cubic meters per second (CMS) does not translate to the lower flow rates along the tributaries of 84 and 195 CMS. This is due to a lack of representation of backwater conditions in the hydraulic routing techniques used. As a parallel problem, there is excess water accumulated along the mainstem that cannot extend in either a fluvial or pluvial manner beyond the boundaries of the mainstem catchments.
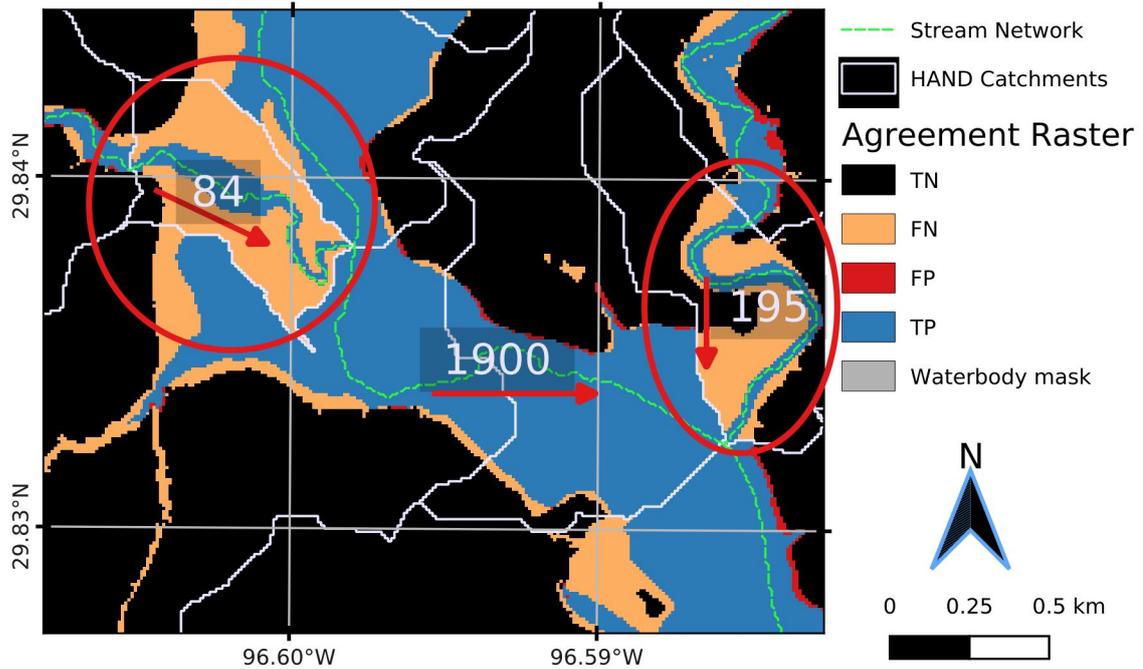
Figure 4. The figure represents an agreement map between a HAND derived FIM and one produced from the Base Level Engineering (BLE) program for a 100 yr magnitude event at HUC8 12090301. Agreement maps are symbolized by false negatives (FN), true negatives (TN), false positives (FP), and true positives (TP) where inundated represents the positive condition (see Section 2.7 for more details). The streamflows associated with each river segment are shown in CMS while the flow directions are symbolized as red arrows. The presence of FNs at the confluence of tributaries (circled in red) with the main segment is associated with lower flow rates in the tributaries that don't account for backwater effects. Additionally, the flow of 1900 CMS from the main segment cannot extend to the neighboring catchments belonging to its tributaries shown here. Water pools up vertically along the catchment boundaries of the higher order segment distorting rating curve behavior (Section 4). Sourcing fluvial inundation from HAND is limited to only its nearest drainage line which is the main issue this study aims to address.

We seek to resolve this catchment boundary problem or nearest drainage limitation by discretizing the target stream network into stream networks of reduced, unit stream order to avoid the constraining of catchments by those belonging to lower order neighbors. By discretizing the network into stream networks of unit stream order (later defined as LP), we remove the influence of neighboring catchments that constrain the in-

undation extent. This creates much larger and overlapping catchments that can source fluvial inundation from multiple reaches as required by the given river stage at current flow conditions. We present two successive methods, National Weather Service MS (Section 2.5.1) and GMS (Section 2.5.2), implemented that reduce the effective Horton-Strahler stream orders of the networks employed and test our presented hypothesis that unary stream order networks enhance FIM performance skill with HAND by expanding the nearest drainage definition to increase potential inundation areas.

To clarify the phrase "reducing Horton-Strahler stream order" used extensively in this paper, every FIM used in evaluation contains a flood extent sourced from every NWM forecast point in the given evaluation domain. What we do to reduce stream order is discretize the NWM FR network into different units of size, MS network (2.5.1) and GMS LPs (2.5.2), that effectively reduce the HAND computation to independent networks of unit stream order. These independent HAND datasets are later used to produce FIM independently and mosaiced together (see Section 2.6). The inundation from the MS HAND is mosaiced with the inundation from FR HAND, while the inundation of each individual LP from GMS is mosaiced together. The Horton-Strahler stream order is only reduced for HAND computation purposes to reduce the negative effects of the nearest drainage limitation inherent to HAND.

2.5.1 NWS Mainstems

The initial attempt at drainage order reduction to solve the catchment boundary issue was to use a stream network relevant to the NWS forecasting community. The Mainstems (MS) network is a subset of the NWM FR network at and downstream of AHPS forecast points as seen in Figure 1. The MS network comprises about 200 thousand km of stream length which is less than 4% of the FR total stream length of 5.5 million km. It also spans 121,724 reaches across 1,608 HUC8s. In this technique, we derive HAND using the FR stream network as well as the MS network which was originally proposed by Djokic (2019). Inundation is derived independently from the resulting FR and MS HAND hydrofabrics and are mosaiced together using the technique proposed in Section 2.6 to form the MS FIMs. Within each HUC, one might typically only find a MS stream network of uniform stream order but this can vary if more than one AHPS forecasting point is found within or upstream of the HUC in question. So while we may refer to the MS network as that of one with unit stream order, we acknowledge there are many cases

where additional or converging forecast points create multiple branches within a given processing unit.

2.5.2 Generalized Mainstems

Since MS only covers 4% of the entire FR stream network, we sought to expand drainage order reduction techniques to all reaches within the NWM modeling domain. In order to do this, we discretized the NWM network into LPs which when considered individually have unit Horton-Strahler stream orders. LPs group flowlines by maximizing the length of each flow path and minimizing the number of LP identifiers within a given domain (Moore et al., 2019; McKay et al., 2012). In order to derive LPs for the NWM FR stream network at the HUC8 processing area, we first compute arbolate sums which are defined as the cumulative drainage distance of all upstream drainage lines. Arbolate sum is also inclusive of the current drainage reach as well. Arbolate sums are computed by starting at the headwater points and summing up drainage distances as you traverse downstream.

Arbolate sum is critical to discretizing the NWM network into LP identifiers. Starting at a HUC8's outlet, a unique LP is propagated upstream. At every confluence, the direction of maximum arbolate sum is sought to propagate the current LP identifier. For the remaining parent reaches of the given junction, a new LP identifier is assigned and the process recursively continues with them. Figure 5 illustrates how LPs (symbolized by unique colors) are propagated upstream by the value of arbolate sum. The figure shows computed arbolate sums and unique LP identifiers on a HUC12 (120903010404) for clarity but were computed at the corresponding HUC8. The mainstem of the figure runs from the red ellipses to the black one which is the outlet. From the figure, we can see how unique colors are propagated in the direction of the maximum arbolate sum.

Each HUC8 is discretized into LPs independently and the relevant inputs as described in Table 1 are assigned to each LP processing unit given a buffer of seven km. This buffer was selected to avoid edge contamination (Lindsay & Seibert, 2013) and to ensure adequate data availability for wide rivers with large catchments in regions with low slope. Further work could be dedicated to tune this user exposed parameter to better balance its effect on FIM extents and computational expense since larger buffers cre-

676 ate additional floating point calculations and storage requirements. For the time being,

677 we designate this issue to be out of scope.

678   At the LP scale, the methods in Sections 2.3 and 2.4 are executed leaving out any

679 tributaries of the LP in question at the time. The only exception to this is the use of the

680 NWM stream network directly for use with hydro-enforcement by burning these lines

681 and seeding from its headwater points directly instead of going through the NHDPlusHR

682 network as described in Section 2.3.1 . This decision was motivated by the difficulty in

683 deriving LPs in the NWM stream network with high agreement with the LPs derived

684 for the NHDPlusHR stream lines. We found that the same algorithm to compute arbo-

685 late sums and LPs could yield enough disagreements associated with disordered branches

686 or slight differences in arbolate sums that could significantly affect the agreement of the

687 LP identifiers in the NWM and NHDPlusHR networks. This yielded enough error to jus-

688 tify the use of the NWM directly for hydro-enforcement operations.

689   Once the NWM FR stream network is discretized into LPs, we independently com-

690 pute HAND using each LP as the target stream network to be used. To illustrate the

691 GMS procedure, we reference Figure 6 to show how deriving HAND and FIMs from GMS

692 works. In Figure 6a, we uniquely color code the LPs derived for the NWM stream net-

693 work. For each one of these lines, we derive HAND and its associated datasets includ-

694 ing catchments, crosswalks, and rating curves. Each LP is buffered to a polygon with

695 a user-exposed, distance parameter of seven km that is used to subset the original DEM

696 for two selected LPs in Figure 6b. We illustrate two HAND grids for two of the LPs in

697 this HUC8 in Figure 6c. Once the FIM hydrofabrics for each LP are generated, we can

698 inundate them individually also shown in Figure 6d. Lastly, these individual FIMs are

699 mosaiced together as explained in Section 2.6 and shown in Figure 6d.

700   For a more intimate look at the drainage order reduction procedure GMS, and its

701 effects, we allude to Figure 7 which references the same area (in HUC8 12090301) and

702 set of river junctions as in Figure 4. The catchments and stream lines for HAND com-

703 puted at the FR scale are illustrated in Figure 7a where the respective inundation at the

704 100 yr magnitude is heavily constrained by the limited catchment extents especially at

705 junctions. In subsequent sub-figures, we show the same datasets for the HAND compu-

706 tation problem for this region but discretized into independent LPs for the main LP (b),

707 the eastern tributary (c), and the western tributary (d). Notably, inspecting (b), one sees

708   how removing the tributaries creates much larger catchments for the main LP. These catch-

709   ments include drainage areas that would traditionally be considered nearest to the trib-

710   utaries thus ineligible to receive inundation sourced from the main LP. The inundation

711   extents in (b) overlap those of (c) and (d) and are mosaiced together by methods explained

712   in Section 2.6.
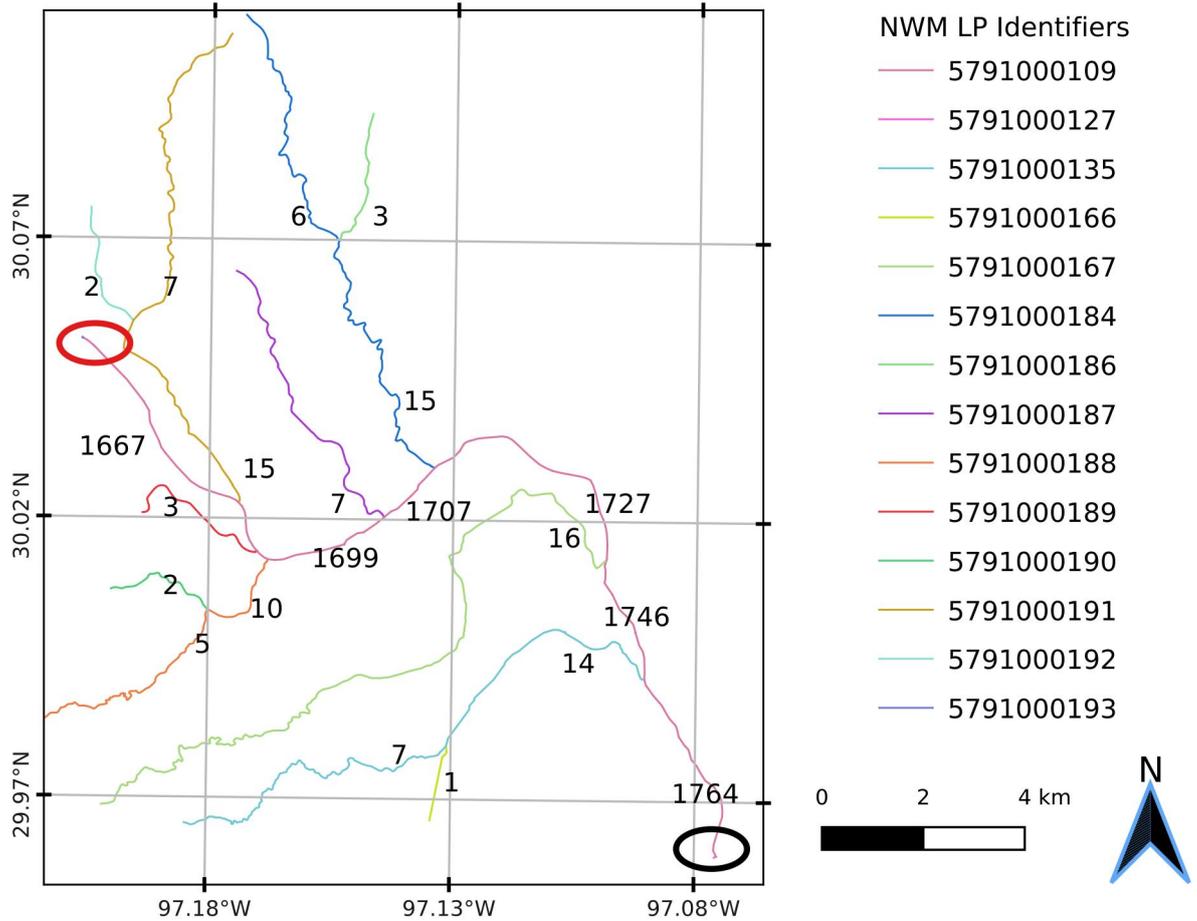
Figure 5. Illustrates the NWM Full Resolution V2.1 stream network discretized into level paths (LP), symbolized by unique colors, as well as the values of the arbolate sums to the nearest whole km distance. The LPs were derived on a HUC8 level (12090301) but only illustrated for a HUC12 (120903010404) for clarity. The mainstem of this HUC12 runs from the red ellipses to the outlet denoted by the black ellipses. Arbolate sums are defined as the cumulative drainage distances of all upstream stream lines. Arbolate sums are computed for the NWM network by starting at the headwater points then traversing downstream and adding the distances cumulatively. LPs are derived by starting at an outlet point with a unique identifier (ID). The unique LP ID is propagated upstream until a junction is reached where the current LP ID is propagated in the direction of maximum arbolate sum. The remaining converging segments at the given junction are each assigned a new unique LP ID and the process is repeated recursively until all reaches have been assigned a LP. Thus, LP serve as a proxy means of assigning membership to a given river when presented with a confluence. Each individual LP has a unit Horton-Strahler stream order thus serves as a great method for our proposed technique.
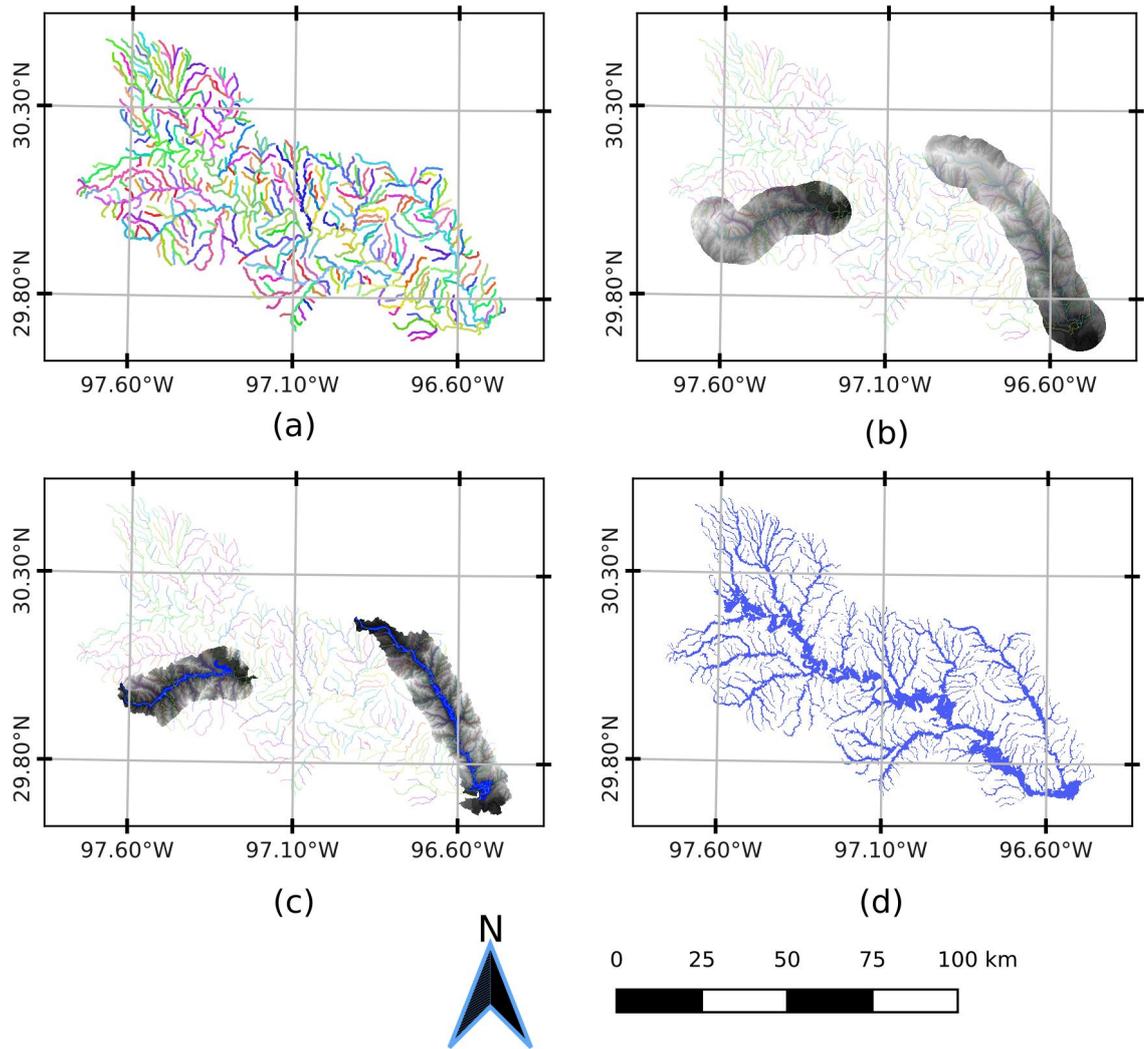
Figure 6. Overall procedure for GMS HAND at HUC8 12090301. In (a), we illustrate all NWM stream lines symbolized by their LP with 372 unique LP IDs in this HUC. Meanwhile (b), demonstrates the DEM clipped to a seven km buffer around two selected LPs. In (c), we show how HAND can be computed just for each one of these two LPs independently. We also show inundation maps created for these two LPs in (c). In (d), we show all the inundation maps for all the LPs mosaiced together.
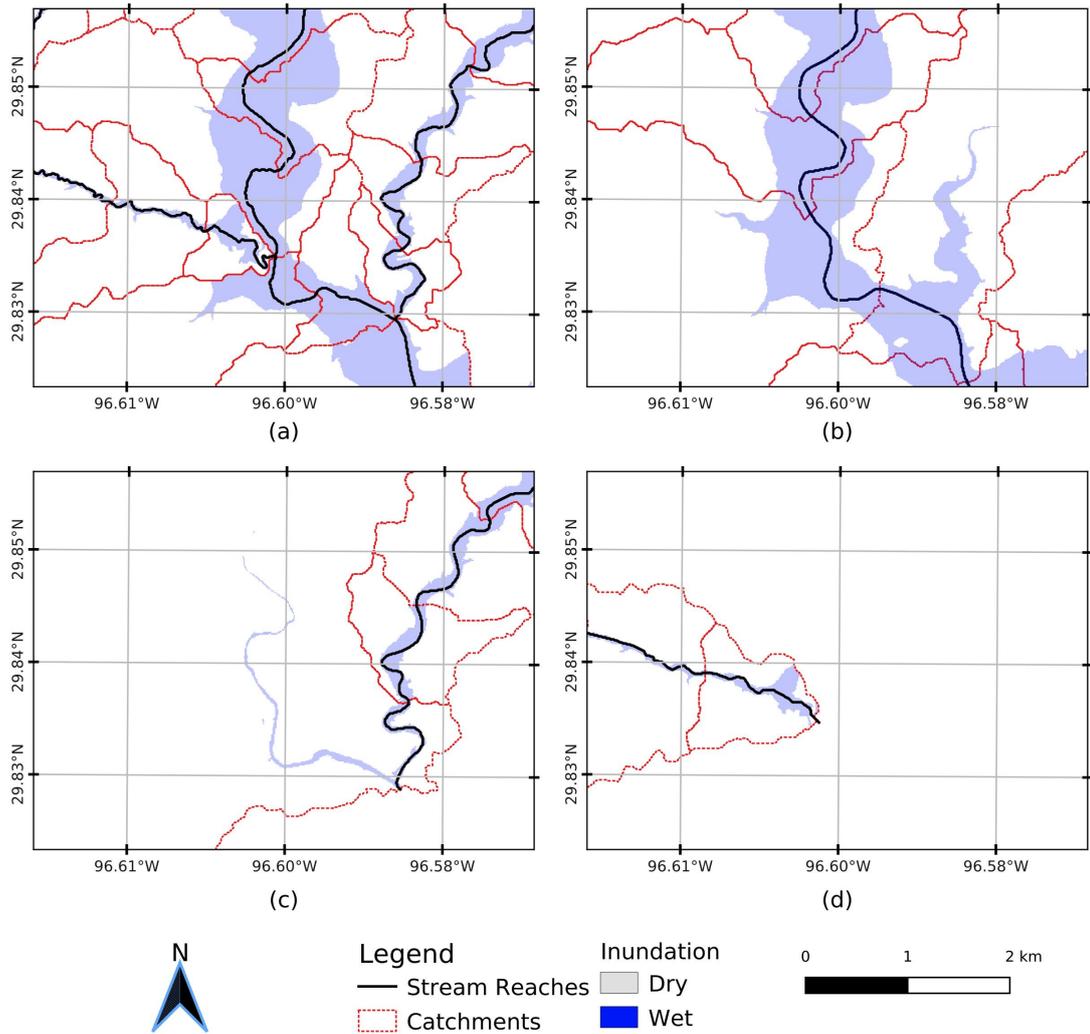
Figure 7. This image, with the same spatial domain as Figure 4 (HUC 12090301), demonstrates how computing HAND on level path (LP) bases leads to larger, independent catchments and more expansive inundation extents (100 yr flows). In (a), the catchments and stream network are shown for HAND computed in Full-Resolution (FR) method which shows constrained inundation extents around the two junctions. (b) demonstrates the LP associated with this region's highest order river. By delineating catchments at this scale independent of the neighboring tributaries, the drainage areas are allowed to expand thus allowing inundation extents to cover previously restricted areas. In (c) and (d), we show the stream lines, catchments, and inundation extents of the two tributaries. Later in Section 2.6, we describe how the inundation in (b), (c), and (d) are mosaiced together to form one seamless inundation map. This process allows for multiple, possible contributing sources of fluvial inundation to be considered thus enhancing FIM skill.

₇₁₃   **2.6 Inundation Mapping**

₇₁₄   The FIM hydrofabric consisting of the relative elevations grid, catchments grid, catch-

₇₁₅   ment polygons, rating curve, and cross-walking data are all used to convert forecasts from

₇₁₆   the NWM into forecasts extents. For operational situations, one would cache the FIM

₇₁₇   hydrofabric then either produce libraries of FIM for a sample of discharges or stages or

₇₁₈   also produce the FIM in near real-time (NRT). From the cached FIM hydrofabric and

₇₁₉   design or forecast discharges including those extracted from the NWM, inundation maps

₇₂₀   can be generated at HUC8 spatial processing units in a rapid, parallel operation. The

₇₂₁   discharges are associated with NWM reach identifiers and cross-walked over to reach iden-

₇₂₂   tifiers in the FIM hydrofabric.

₇₂₃   Utilizing the stage-discharge relationships in the SRCs, each forecast for each catch-

₇₂₄   ment identifier is assigned a stage value. The catchments grid encoded with the reach

₇₂₅   identifiers are used to map the stages by thresholding to the forecast stage. We use the

₇₂₆   basic logic already established in previous works to conduct this (A. D. Nobre et al., 2016;

₇₂₇   Y. Y. Liu et al., 2016; Maidment, 2017). Mathematically, the HAND values, H_{ij}, can

₇₂₈   be indexed by the reach identifiers, i, and pixel indices, j. For each forecast stage, S_i, one

₇₂₉   can express the formula for D_{ij}, a continuous variable denoting water depth at a given

₇₃₀   pixel with reach and pixel identifiers i and j respectively in Equation 2. For each fore-

₇₃₁   cast stage, S_i, one can express the formula for F_{ij}, a binary variable denoting inunda-

₇₃₂   tion condition in Equation 3 in terms of D_{ij} by simply thresholding at zero depths.

$$D_{ij} = S_i - H_{ij} \qquad (2)$$

$$F_{ij} = D_{ij} > 0 \qquad (3)$$

₇₃₆   For the cases of MS and GMS, the inundation maps produced for the respective process-

₇₃₇   ing units at lower maximum stream orders must be mosaiced together to form a seam-

₇₃₈   less forecast in the form of a single raster file. For mosaicing the depths, we select the

₇₃₉   maximum inundation depth from the all the contributing areas K index by its lower case

₇₄₀   character, k. Consolidating the depths using a maximum function was decided upon based

₇₄₁   on intuition which we believe to best represent the depth of water in an area with mul-

₇₄₂   tiple contributing fluvial inundation sources. Other aggregation methods could lead to

₇₄₃   different results but were not investigated here. Equation 4 illustrates how the maximum

depth from all the contributing areas, k, to each pixel j in catchment i,

$$D_{ij} = \max_{k=[1,...,K]} D_{ijk} \tag{4}$$

. Equation 5 illustrates the same process but for mosaicing the binary inundation maps,

$$F_{ij} = \max_{k=[1,...,K]} F_{ijk} \tag{5}$$

. For the MS and GMS methods, the contributing areas are defined differently. For MS, the FIM from MS HAND and FR HAND are mosaiced together to form a singular inundation map thus K is set to two for that case. For GMS, all FIMs from all the LPs in a given area are mosaiced together then K is set to this number of LPs. Figures 6a and 6b, illustrate how inundation maps are created for lower stream order processing units then mosaiced together.

2.7 Evaluation

Possible benchmark FIM candidates for evaluation purposes include high water marks, remote sensing observation, crowd-sourced information, and modeled extents. These sources are all subject to limitations for evaluating a continental scale model like OWP FIM such as but not limited to a lack of spatial coverage, signal interference, lack of streamflow data, inaccurate streamflow data, physics-based assumptions, and errors in input data. While in-situ observations such as high water marks offer the highest accuracy, they are often limited in spatial extent and can lack the associated streamflow data necessary to make FIMs to compare to as to isolate out other hydrological factors.

Evaluation of our relative elevation CFIM method is conducted by comparison to the HEC-RAS 1D models produced within FEMA region 6 (*Base Level Engineering (BLE) Tools and Resources*, 2021; *estBFE Viewer*, 2021; USACE, 2022). This dataset was selected due to its large spatial coverage, availability of cross-sections with streamflow information, higher level of sophistication when compared to HAND, engineering scale detail, and a storied use in the literature as an evaluation dataset (Cook & Merwade, 2009; Rajib et al., 2016; Zheng, Maidment, et al., 2018; Afshari et al., 2018; Wing et al., 2017; Criss & Nelson, 2022; Follum et al., 2017). We selected 49 available HUC8s, shown in Figure 8, which span about 185 thousand km^2 across nine states. The maps of the 1% recurrence flow (1 in 100 year) and the 0.2% recurrence flow (1 in 500 year) are furnished by InFRM as well as the corresponding discharges and mapping extents for evaluation.

774   We did exclude NWM V2.1 Reservoirs from evaluation because these are not properly

775   accounted for in the inundation sourced from OWP FIM.

776   By using the same HEC-RAS derived discharges and FIM extents for creating maps

777   with OWP FIM, we are able to separate out errors introduced by NWM inputs and pro-

778   cesses including land surface interactions, groundwater fluxes, atmospheric forcings, hy-

779   draulic routing, and others that would have potentially affected our conclusions if we had

780   used NWM forecasted discharges. Figure 9 illustrates both NWM V2.1 and BLE stream

781   lines as well as the BLE cross-sections that have recurrence discharges associated with

782   them. We elected to spatially intersect the HEC-RAS cross sections with the NWM stream

783   network assigning the 1% and 0.2% flow rates to each NWM reach. To handle multiple

784   intersections, we opted to use a filter to select the median discharge value attributed to

785   each NWM reach. This partially handles the influence of neighboring cross sections that

786   could cause flow discontinuities and mass conservation issues. Additionally, the stream

787   network of the InFRM furnished models are of higher stream densities and bifurcation

788   ratios, as evident in Figure 9, leading to a significant amount of false negatives (FN) (under-

789   prediction) along headwater streams with unit Horton-Strahler order due to the lack of

790   representation of these additional headwater streams in the NWM network. While the

791   limitations are noted, this method does best to detangle the influence of exogenous vari-

792   ables that we do not wish to study in this comparison.
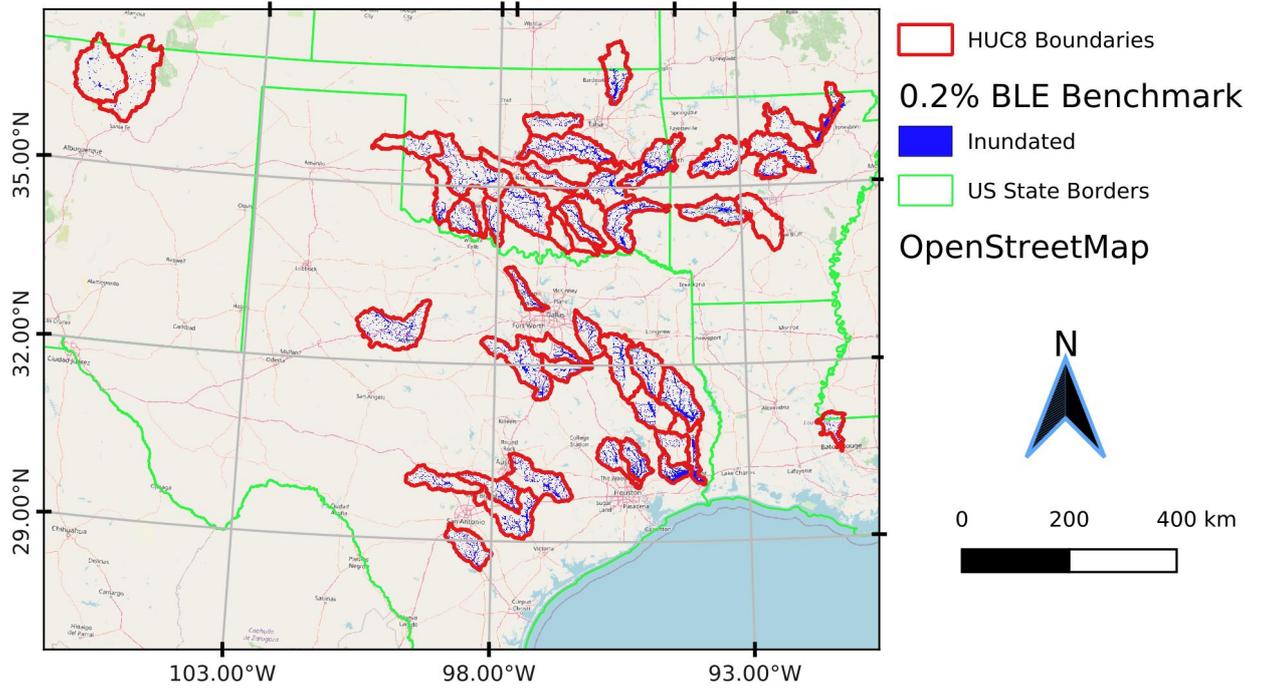
Figure 8. Shows 185 thousand km^2 of modeled areas for the Base Level Engineering (BLE) domain of 49 HUC8s across nine states at 0.2% recurrence magnitude for flow rates. BLE maps are produced for two recurrence flows, 1% (100 yr) and 0.2% (500 yr), using 1D HEC-RAS models. The maps are used as benchmarks for validation purposes of OWP FIM.
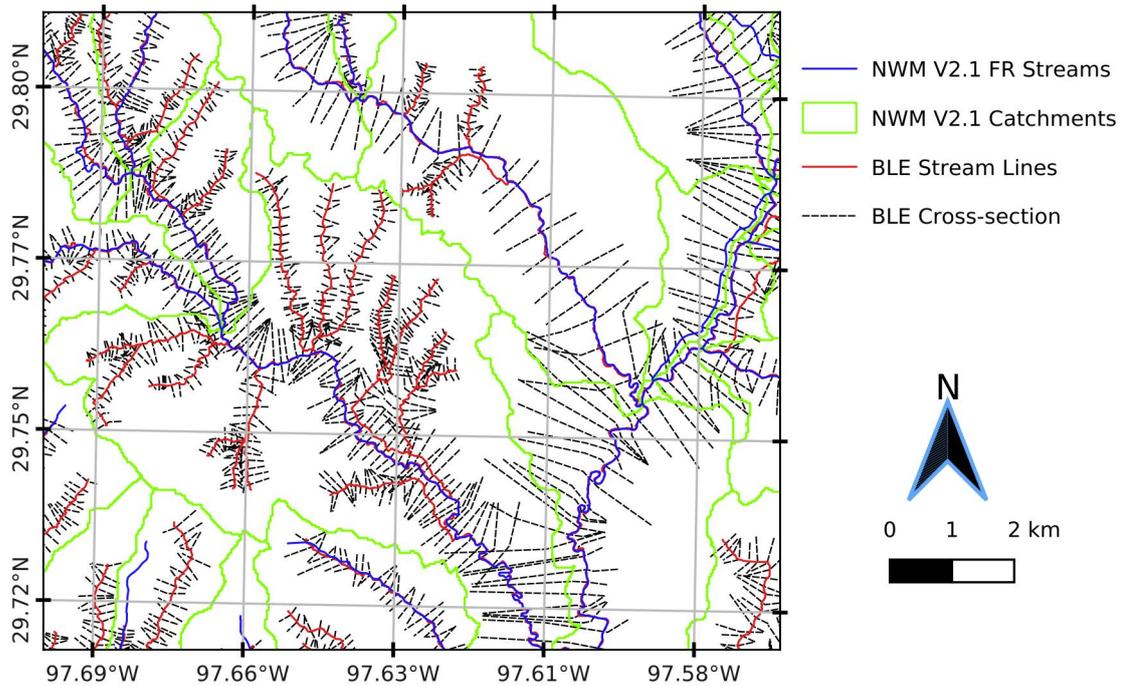
Figure 9. Illustrates Base Level Engineering (BLE) cross sections and stream lines at the HUC8 12100203 near the confluences of West Fork Plum Creek and Clear Fork Plum Creek with Plum Creek. BLE cross sections are intersected with NWM reaches and the median recurrence discharge for 1% and 0.2% levels are selected per NWM V2.1 Full Resolution (FR) stream lines. Additionally, we illustrate the NWM V2.1 catchments to provide a sense of how many cross-sections may intersect a given NWM flowline. The BLE stream network is also shown which is denser than the NWM V2.1 stream lines meaning there are several lower order streams represented in the BLE stream network that are not in the NWM V2.1 stream lines. This creates additional inundation areas in the validation data that are not modeled with our HAND based FIMs.

The metrics employed in this study to evaluate inundation extents include CSI, Probability of Detection (POD), and False Alarm Ratio (FAR) and are presented in Equations 6, 7, 8, respectively. To calculate these secondary metrics, one must define three primary metrics starting with true positives (TP) which is predicted wet and wet in the BLE benchmark dataset. The two types of errors consist of false positives (FP), or type I errors, which is dry in the benchmark but predicted wet and false negatives (FN), or type II errors, which is wet in the benchmark but predicted dry. Lastly, the reader may

come across true negatives (TN) which is defined as dry in both the benchmark and predicted datasets. Maximizing POD indicates a model's ability to detect the given threat of interest, inundation, while minimizing FAR is sought to indicate a models ability in reducing FN errors. In other words, POD is an indicator of model skill in inundated regions while FAR is an indicator of model skill in non-inundated regions. Some work by Gerapetritis and Pelissier (2004) denotes CSI a good proxy for measuring a forecasting system's utility in protecting life and property and has been shown to be optimized mathematically when $POD = 1 - FAR$. We use all three secondary metrics here to add value to the discussion while avoiding aggregating away the meaning of all four primary metrics.

While these metrics are commonly employed in the evaluation of FIM and binary weather prediction communities in general, they do come with some notable limitations including frequency dependence in the case of CSI and FAR (Gerapetritis & Pelissier, 2004; Stephens et al., 2014; Schaefer, 1990; Jolliffe & Stephenson, 2012). Thus, frequency dependent statistics should be used with caution when comparing across sites with varying frequencies. Lastly, approximately six HUC8s do not have NWM MS reaches thus we imputed the metrics for FR for these sites as the best available forecasting capability to compare GMS metrics to.

$$CSI = \frac{TP}{TP + FN + FP} \tag{6}$$

$$POD = \frac{TP}{TP + FN} \tag{7}$$

$$FAR = \frac{FP}{TP + FP} \tag{8}$$

3 Results

3.1 Flood Mapping Performance

We produced FIMs for the entire BLE domain within the 49 HUC8 areas across several states in the south central US. The forecasted FIMs using the discharges for the 1% (100 year) and 0.2% (500 year) recurrence flows directly from HEC-RAS were used to avoid noise and errors from hydrological processes. We computed the statistics, CSI, POD, and FAR, for both 100 and 500 year events for Mannings N set to 0.06 and 0.12. These results are presented in Figure 10 as violin plots and in Table 2 as aggregated metrics with the results for MS and GMS presented as percentage changes from their respective FR values. To be more specific, Table 2 sums the primary metrics, TP, FP, FN, and TN, across all HUC8s then recomputes the secondary metrics which was done to better account for large variances in HUC8 size. The same trends discussed below are consistent across both reporting methods (Figure 10 and Table 2).

The distribution of these flood extent metrics can be examined in Figure 10 as violin plots. Each half of a violin plot represents the kernel density estimation (KDE) for a given model (FR, MS, GMS), Manning's n value (0.06, 0.12), recurrence interval (1%, 0.2%), and performance metric (CSI, POD, FAR). For example, let's examine the violin plot for the row marked CSI and column for Manning's n = 0.06. This sub-figure shows the CSI distributions across all 49 HUC8s when Manning's n is set to 0.06. Each independent, whole violin represents the HUC8 metric value distribution of FR, MS, or GMS while each half of the violin represents the distribution of that data divided up by magnitude (blue for 100 yr and orange for 500 yr). The horizontal dashed and dotted lines represent the 25th, 50th, and 75th percentiles from bottom to top, respectively. Additionally, we show trend lines symbolized in green that for each metric and Manning's n combination denotes the best fit line for the three methods (FR, MS, and GMS). To avoid having two trend lines per sub-figure, we elected to aggregate the two magnitudes together as they tend to observe similar trends across the three models. The slope of each trend line is quantified in the figure by its β_1 value and the p-value for that statistic which tests the significance of that trend (deviation from a zero sloped trend line).

Both Figure 10 and Table 2 contain a fair amount of information that is central to the objectives of this study. As previously stated in Section 2.7, we consider CSI as a general proxy for the skill of the inundation extents with POD denoting skill on in-

undated areas and FAR indicating skill on non-inundated areas. Again, the main objective of the study is to introduce how computing HAND with disaggregated stream networks to those with unit stream order can enhance the fidelity of FIMs by capturing fluvial inundation from multiple sources as opposed to that of just the nearest drainage line. As can be seen in Figure 10 and Table 2, CSI generally increases from FR to MS and MS to GMS for both sets of Manning's n values and flood magnitudes. This increase is primarily driven by an increase in POD thus generally increasing the probability of correctly detecting inundation. Also, we note that FAR is somewhat, albeit marginally, decreased from FR to MS and MS to GMS for both sets of Manning's n values and flood magnitudes. The increases in CSI and POD as well as the decreases in FAR with respect to the methods, FR, MS, and GMS, are not only observed among the trend lines but also in the 25th, 50th, and 75th percentiles (Figure 10). So overall and in other words, the broader distribution of HUC8s improves across the three methods. Due to the means by which FIM is produced utilizing FR, MS, and GMS, we can say that the more we derive HAND on networks of unit stream order and mosaic the resulting FIMs, the better those FIMs perform. We move more details on the relationship between stream order and FIM skill to the Discussion section (Section 4).

Additional noteworthy trends in Figure 10 center around the all-around better performance of FIMs for those of higher Manning's n values and recurrence flows. The higher Manning's n value enhances performance for both recurrence intervals across all models which seems to better agree with the value of 0.11 used in the BLE model (*Base Level Engineering (BLE) Tools and Resources*, 2021; *estBFE Viewer*, 2021). Most of this improvement is driven by significant increases in POD, but unfortunately, it also leads to a significant amount of over-prediction as observed by the increase in FAR. More work can be invested to better regionalize Manning's n values for FIM purposes with HAND. We also observe additional trends associated with the magnitude or recurrence interval of the flow rates used with the higher flow rates exhibiting better overall CSI, POD, and FAR values than the lower, 100 yr magnitude. We introduce in the Discussion (Section 4) that this skill premium exhibited by higher flow events could be due higher quality elevation data in regions that are not described as bathymetric areas.
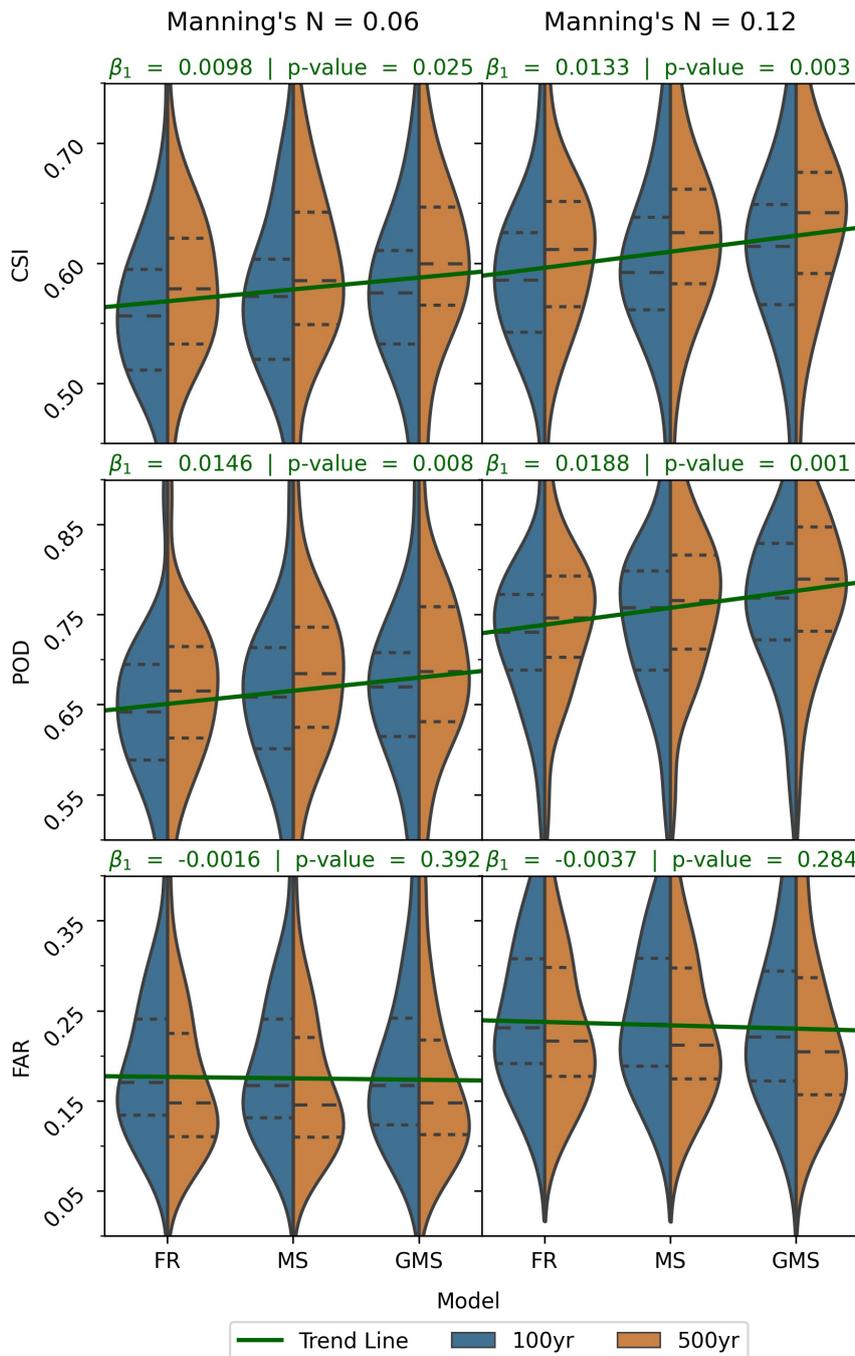
Figure 10. Shows kernel density estimation of the distributions (sample size = 49) for 1% (100 year) and 0.2% (500 year) along with horizontal, dashed lines for the 25th, 50th, and 75th percentiles (in order from bottom to top). The sub-figures separate the combination of three metrics (CSI, POD, and FAR) for two settings of Manning's n (0.06 and 0.12). Trend lines for each combination of metric and Manning's n are shown (sample size = 294) along with associated slope and p-value of slope testing one-tailed significance.

Table 2. Recomputed CSI, POD, and FAR using the primary metrics, TPs, FPs, and FNs, aggregated to the BLE domain. The values for MS and GMS methods are expressed in percentage change (%) from their respective values with the same Manning's n, magnitude, and metric combination in the Full Resolution (FR) method columns. The best value across models is highlighted in bold.

Metric	Manning's n	FR		MS (% Change)		GMS (% Change)	
		100 yr	500 yr	100 yr	500 yr	100 yr	500 yr
CSI	0.06	0.5576	0.5839	2.53	2.59	**3.95**	**4.04**
	0.12	0.5915	0.6149	2.35	2.26	**4.51**	**4.65**
POD	0.06	0.6354	0.6575	2.68	2.74	**4.39**	**4.38**
	0.12	0.7255	0.7446	2.83	2.71	**4.84**	**4.89**
FAR	0.06	0.1800	0.1609	-0.72	**-1.24**	**-1.22**	**-1.24**
	0.12	0.2379	0.2208	-0.21	-0.18	**-2.31**	**-2.72**

3.2 Computational Performance

The NFIE experiments were able to produce HAND for 331 HUC6's in 1.34 CPU years (Y. Y. Liu et al., 2016) and estimates using work from Djokic (2019) put producing HAND at the FR NWM at 0.55 CPU years. For our work, we were able to produce HAND at the full NWM resolution in 0.13 CPU years which represents a substantial speed-up compared to previous works. For the MS resolution, an additional 0.05 CPU years is required on top of this bringing the total to about 0.18 CPU years to produce 2,188 HUC8s that span additional areas not covered in previous HAND versions including Hawaii and Puerto Rico. GMS which generalizes HAND production to the LP scale adds a significant amount of CPU time to the process bringing the estimate total to about 1.17 CPU years.

4 Discussion

Overall, the main observation of this study was how FIM performance can be improved by reducing the Horton-Strahler stream order of the target stream network used for HAND computation. Most of this change is accounted for by substantially increasing POD and inundation extents in some areas thus reducing FNs. We believe, as we later argue, that the increase in POD is primarily driven by an increase in the catchment sizes that is inherently enabled by dividing up the stream network into independent stream networks of unit stream order. Additionally, we note that reducing drainage order also has some minor influence on reducing inundation extents in other areas and the rate of false alarms. We believe that this effect is driven by a change in the stage-discharge relationship where a given streamflow leads to lower river stage values when HAND is computed with target stream networks of unit drainage order. We seek to explain that these two effects, catchment boundary enlargement and stage-discharge curve lowering, are highly interrelated and cannot be easily detangled. Lastly, we discuss the diminishing effects on performance that the MS and GMS techniques may have and also any additional effects including enhanced cross-walking abilities.

As evident in the results of the study in Section 3, a sizable amount of the increase in CSI observed by reducing stream order for HAND computation can be attributed to increases in POD. This can be inferred by close inspection of Figure 10 and Table 2 where changes in POD are significantly higher than that of FAR. Upon investigation of the performance of HAND derived FIM, we note a general increase of catchment sizes for the MS and GMS methods when compared to the FR method as they are now delineated independently of any tributaries that would constrain catchment sizes otherwise. Additionally, we note significantly less water build up along catchment boundaries especially at the junction of lower order tributaries with lower flow rates and higher order rivers with more flow. This allows for inundation extents to expand across regions previously encapsulated by catchments of joining reaches in lower flow tributaries. The water built up along the catchment boundaries can be thought of as a column of water in a cylindrical container (catchment) that has exceeded the elevation of the container's rim which does not represent accurate physics.

Large scale HUC8 level evaluations can fail to demonstrate fine grain enhancements as they aggregate away many changes that are only clear at more local scales. Future

assessments of OWP FIM should consider finer grain evaluation units as well possible impact assessments using asset information such as building footprints to better illustrate fine grain changes in a more relevant manner to stakeholders. For now, we provide Figure 12 which best illustrates the improvement offered by multi-source fluvial flooding capabilities in a more local context. The figure is comprised of two agreement rasters for two different HAND based FIMs compared to the validation dataset for a given region with a high flow mainstem (500 yr recurrence flow) running horizontally along the region. Sub-figure 12a demonstrates the agreement raster for the FR stream network as well as its respective catchment boundary lines symbolized in white and stream network shown in green. Inspection of this sub-figure denotes a clear spatial pattern where TPs or areas correctly inundated are pooled alongside catchment boundary lines. On the other side of the catchment boundary, one can witness large swaths of FNs that should be inundated. The FNs also exhibit a spatial pattern as in they tend to collocate within catchments of the pictured mainstems tributaries. This sort of behavior was introduced early in the paper and shown qualitatively in Figure 4.

As an enhancement, this paper proposes computing HAND for stream networks comprised of level-paths independently of one another. In sub-figure 12b, the agreement raster for the GMS technique is illustrated as well as the stream network lines in green. While the entire mosaiced inundation map from GMS (as described in Section 2.6 and Equation 2) is used to produce this agreement map, we only show the catchments associated with the mainstem of this region that is shown to follow a clear horizontal path. Showing all the catchments for the tributaries that were all derived independently would convolute the image. The main message illustrated here is that the catchments associated with the mainstem of this area significantly increase in size. Since they are not restricted by the catchments of tributaries that lie in the same drainage areas as those of the mainstem, they extend to include those as well. The consequence for inundation extents is a general increase in spatial coverage of the river's water which shows to be in much better agreement with the benchmark map. The TPs are no longer bounded by the catchment lines and allowed to expand to their natural extents.

We note here as a contribution of this study that a major inherent, limitation of HAND is the "nearest drainage" constraint or the idea that a given river reach only drains or, in HAND's case, inundates its immediate, unique drainage area or catchment. In other words, HAND based FIMs are limited to producing fluvial inundation to only their near-

est drainage area or catchment. However, we know that fluvial inundation can be sourced from several streams nearby that also serve as drainage outlets to the area in question. Generally speaking, drainage areas are known to be hierarchical in nature so a given drainage area for a given outlet point can be decomposed into nested drainage areas for outlet points that lie in the original drainage area. A perfect example of this are points that lie on tributary reaches closely neighboring a mainstem. These points lie in the drainage area of reaches in the mainstem but inundation from the mainstem cannot reach these tributary catchments because of the "nearest" assumption in HAND. Hence it's important to state that just like there are different sources of flooding such as fluvial, pluvial, groundwater, reservoir, barrier failure (dam/levee/embankment), and coastal, there can also be multiple sources of a riverine flood. HAND is only equipped to handle riverine flooding from the nearest drainage line. Other relevant drainage lines that produce fluvial flood waters are not considered here especially if the routing model used doesn't consider backwater effects.

The remaining portion of the improvement in CSI was found to come from a marginal yet notable reduction in FAR. Upon investigation of the spatial results in the agreement maps, we found some areas of slight reductions in FPs especially where changes in catchment boundaries may have occurred due to the reduction in effective stream order in computing HAND. These observations pointed to changes in the SRCs introduced by stream order reduction and catchment definition adjustments. Figure 11 illustrates the general effect that stream order reduction has on SRCs. Sub-figure 11a shows how the average SRCs for all reaches for stage values 0 to 25 meters at one-third meter intervals tend to shift the curve down and to the right with ever increasing stream order reduction (FR to MS to GMS). This bias is more pronounced for GMS since that implements stream order reduction down to the unit level for the entire FR network while MS only does so for 4-5% of the network.

Attempting to diagnose this bias in the SRC leads one to Equation 1 which shows the reach averaged SRC relationship between stage and discharge. Across the three methods explored, FR, MS, and GMS, one identifies differences in the inputs and outputs and notes no difference in the stages and Manning's n values. While the channel slope and reach lengths are not exactly the same across methods, their averaged differences are very negligible which only leaves room for deviations in volume and bed area. Again, volume ($V(y)$ or simply V) is synonymous to reach-averaged cross-sectional area and bed area

994 (B(y) or B) is analogous to reach-averaged hydraulic radius but these associations only

995 hold when reach length, L, is considered. Discharge, Q, is directly related to volume and

996 inversely related to bed area and each parameter is weighed according to the magnitude

997 of its exponent which are $\frac{5}{3}$ and $\frac{2}{3}$ respectively (see Equation 1). Figures 11b and 11c

998 show how volume and bed area compare across the three methods with GMS having sig-

999 nificantly greater values than MS which has greater values than FR. Again the relative

1000 discrepancy between FR vs MS and MS vs GMS is explained by the extent of their spa-

1001 tial coverages. Both V and B values increase but are weighed differently by their respec-

1002 tive exponents and pull Q in different directions. We show in Figure 11d the relation-

1003 ship of $\frac{V^{5/3}}{B^{2/3}}$ and plot this ratio against stage, y, to show how these two parameters col-

1004 lectively pull the Q up and changes the SRC accordingly. In other words, the magnitude

1005 and weight of the volume at each stage level exceeds the influence of the magnitude and

1006 weight of the bed area. Both parameters are set to increase mainly due to much larger

1007 catchments leading to more pixels at each stage level as shown in Figure 11e.

1008 Much of the increase in inundated pixels, volume, and bed area can be explained

1009 by much larger catchments that encompass neighboring tributaries. These tributaries

1010 have a significant amount of bathymetry that is low-lying thus easily included in the ge-

1011 ometry for the SRC derivation. They also contribute volume and bed area that is tech-

1012 nically not perpendicular to the flux of streamflow being accounted for in the stream in

1013 question. Careful examination of Figure 12b shows how much larger catchments include

1014 neighboring tributaries and the geometry associated with those tributaries. This geom-

1015 etry is not perpendicular to the flow that is associated with the main reach thus lead-

1016 ing to biases in the SRC. We consider this to have a nuanced effect on skill, while reduc-

1017 ing the rate of FPs it also can lead to FNs due to biases in the SRC.

1018 We note that reducing stream order does suffer from diminishing returns where the

1019 increase in mapping skill for applying stream order reduction to roughly 4-5% of the stream

1020 network is about the same as the increase for applying stream order reduction to the re-

1021 maining 95-96% of the stream network. This motivates further work in identifying what

1022 the optimal coverage of stream order reduction could be and how to parameterize that

1023 coverage. One option could be removing lower stream orders (e.g. 1 and 2) from stream

1024 order reduction and simply using the inundation from FR from these areas.

Additional analysis of Figure 12a, reveals that some catchments don't have inundation or significant inundation. While the cause of these errors can be varied, we assert here that conflating four networks for use in evaluations leads to significant error. Section 2.4.6 details how reach identifiers are conflated for the FIM network back to that of the NWM. One of the issues with the FR version of HAND occurs when a reach of given stream order accidentally conflates to that of a neighboring tributary that is of lower order which leads to areas of FNs. The utilization of MS and GMS only conflates to NWM catchments directly associated with the LP in question which is inherently easy to do with those methods. Thus part of the improvement in MS and GMS methods is due to a slight improvement in cross-walking methodology. The NWM stream network was derived using the NHDPlus V2 dataset which was derived from coarser DEMs than those used here. Additional conflation is identified in cross-walking the stream network used by the BLE maps and those of HAND. Until a singular stream network is used for the NWM, BLE benchmark, and for HAND based FIM, conflation will continue being a source of error.

Our qualitative analysis suggests that the SRCs offer a significant opportunity for improvement in HAND based FIM for future development. The bathymetry of the 10 m DEM from 3DEP is known to be lacking proper representation thus leading to inadequate representation of volume and bed area with all three methods employed. Manning's n which typically accounts for roughness could be tuned to account for these DEM limitations or could be held fixed to some local value associated with a given flood magnitude. Some adjusting parameter must be introduced to enhance the estimation of the bathymetric representation. Lidar DEMs from the USGS at 3 m and 1 m scale could be utilized to derive HAND as well which we conject should show better agreement with higher fidelity FIMs also derived from the same Lidar based DEMs. We suspect that a significant amount of the difference in performance between 100 yr and 500 yr magnitude events can be attributed to poor SRC performance due to poor bathymetric representation. Lower magnitude events are, logically, more susceptible to poor bathymetric data due a greater proportion of the inundation being attributable to areas that are more typically under normal flow conditions. Higher flow events tend to cover regions with more floodplain inundation thus less sensitive to errors from bathymetric data quality. On a related note, the use of the AGREE DEM method discussed in Section 2.3.1 also interacts with the bathymetry issue introducing several artificial geometry param-

1058 eters that affect SRC shape and quality. Due to focus on the nearest drainage problem,

1059 we leave future work related to SRC representation including roughness estimation, bathy-

1060 metric data assimilation, and bathymetry adjustments as opportunities for major enhance-

1061 ments in HAND based FIM.

1062   Lastly, after errors introduced by conflation, poor roughness estimation, bathymet-

1063 ric/elevation adjustment are accounted for, HAND still has another fundamental lim-

1064 itation that is inherently baked into how it works. For HAND to be derived and thus

1065 create a FIM for a given area, that area must entirely drain to the stream network and

1066 the stream network must also drain itself. In other words, an entire area eligible for flood-

1067 ing must monotonically decrease in elevation. DEM's naturally don't do this and the dy-

1068 namics of true flood events don't follow drainage patterns. Enforcing this assumption

1069 for HAND leads to significant amount of DEM manipulations that introduce basic er-

1070 rors. These errors are deep into the assumptions of HAND and thus more difficult to dis-

1071 entangle. Ultimately, the use of more advanced 2D hydrodynamic models should be con-

1072 sidered for dealing with this limitation of HAND but would come at significant expense

1073 at the given high resolution across very large spatial scales and frequent forecast reso-

1074 lutions.
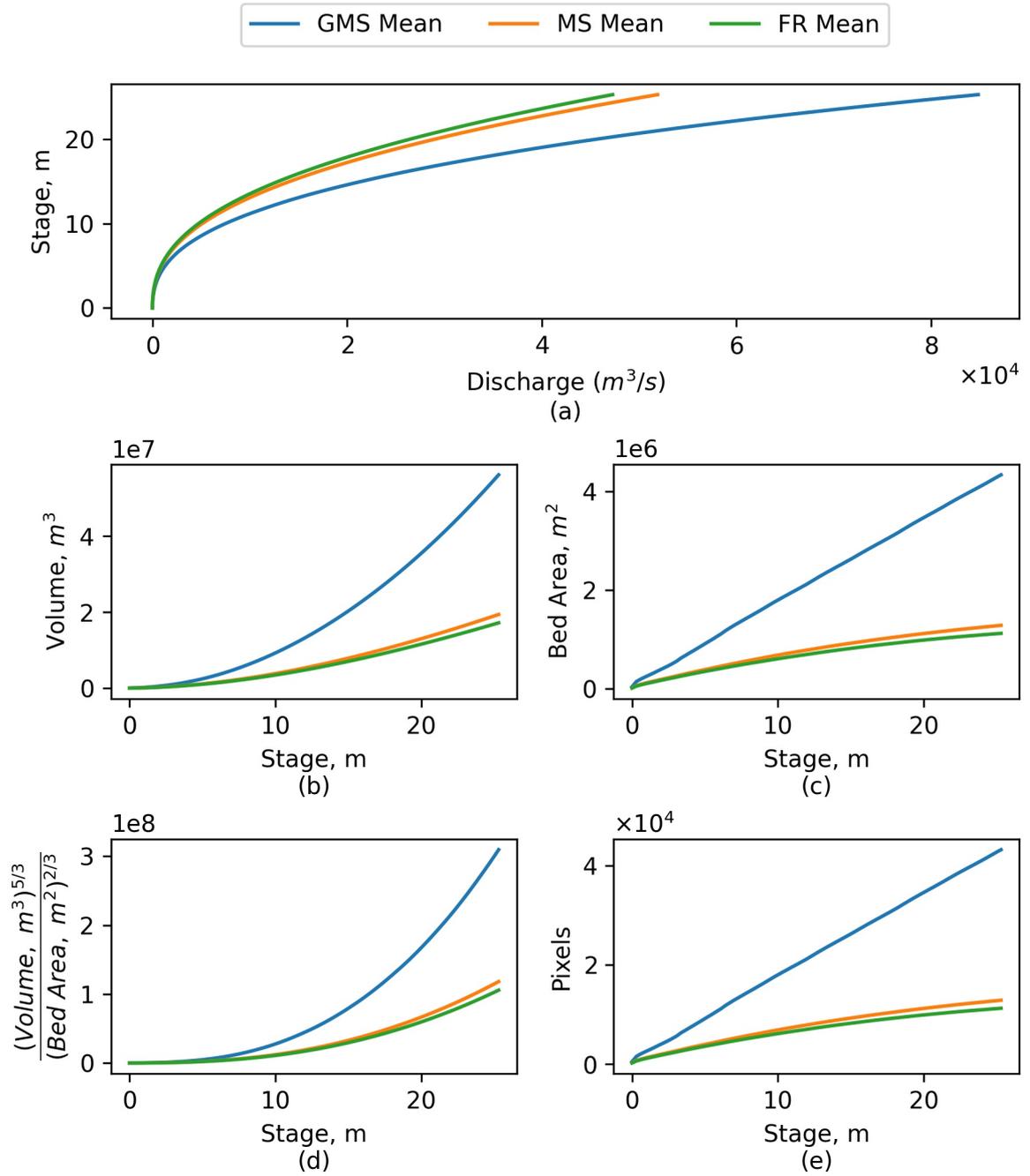
Figure 11. Illustrates average quantities for the three methods, FR, MS, and GMS, for each stage value (m). The values are (a) Discharge $m^3 s^{-1}$, (b) Volume m^3, (c) Bed Area m^2, (d) a function of Volume and Bed Area, and (e) number of pixels.
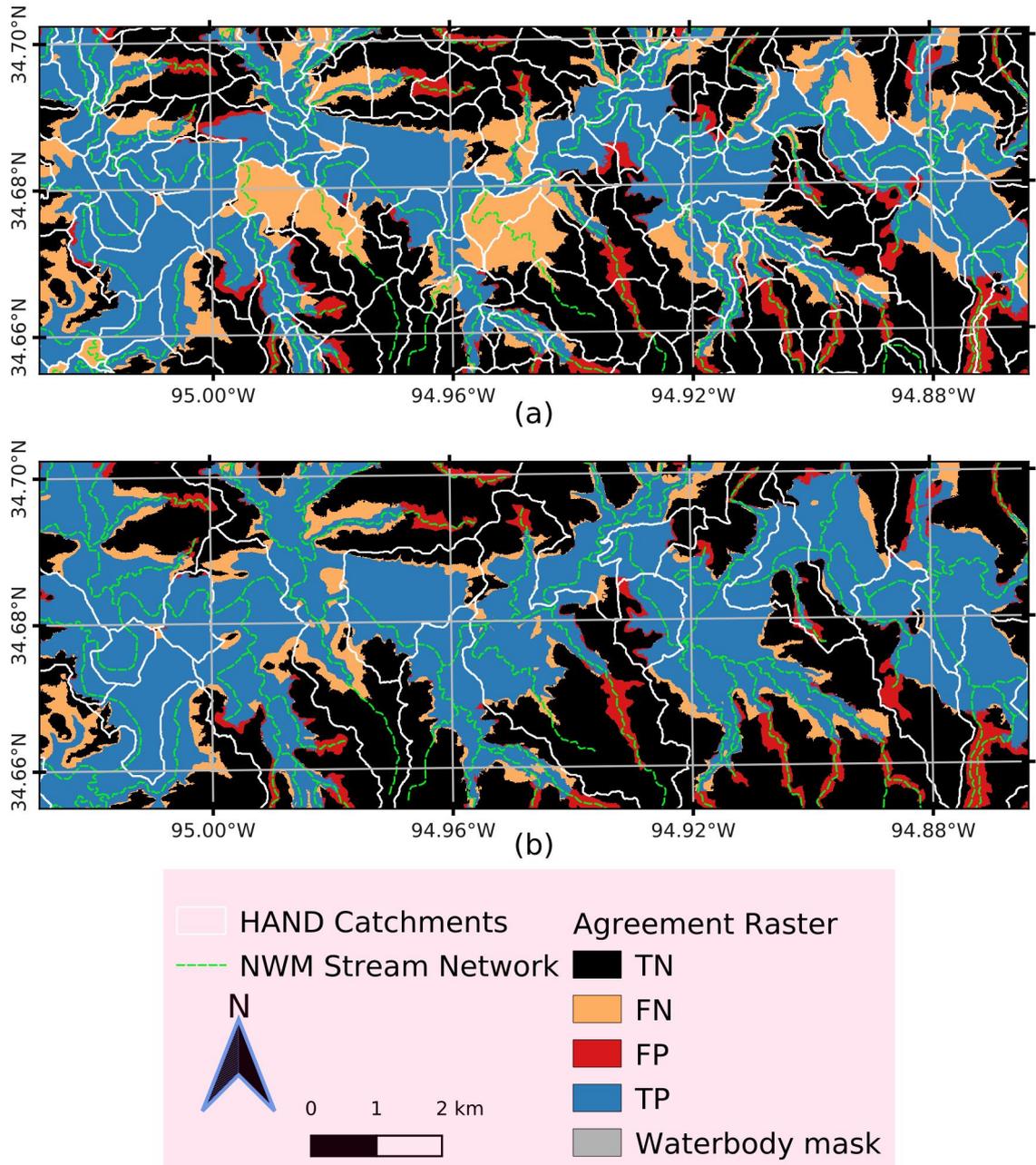
Figure 12. OWP FIM inundation agreement, TP, FP, FN, and TN, with BLE HEC-RAS maps in HUC 11140105 at the 500 yr recurrence magnitude. Catchment boundaries and stream lines are shown in white and dotted green, respectively. Sub-figure (a) shows agreement of FR HAND denoting significant areas of under-prediction due to junctions and catchment boundaries. Meanwhile, (b) shows the agreement for GMS and much larger catchments leading to much better inundation agreement for this given reach. Overall, this illustrates the benefits of stream order reduction for deriving HAND datasets.

5 Conclusions

Floods present a significant, under-served, and expanding risk to life, property, and resources. Previous flood forecasting systems lacked the coverage to adequately inform society of these risks. The National Water Model (NWM), developed by the National Oceanic and Atmospheric Administration's (NOAA) Office of Water Prediction (OWP) and the National Center for Atmospheric Research (NCAR), provides increased spatial coverage and resolution as well as additional forecast time horizons on mostly hourly intervals. Additional modeling is required to convert streamflows from the NWM to river stages and finally to flood inundation maps (FIM). Height Above Nearest Drainage (HAND) is a means of detrending digital elevations maps (DEM) by normalizing elevation to the nearest relevant drainage point. HAND coupled with the use of reach averaged synthetic rating curves (SRC) provide such a means of creating continental scale FIM capabilities at high resolutions (1/3 arc-second, 10 m) and high temporal frequencies (up to 1 hr). Scalable, open-source software, known as OWP FIM, was developed to produce HAND and associated datasets (catchments, SRCs, and cross-walking data) for the NWM forecasting area including Hawaii and Puerto Rico (Aristizabal et al., 2022b). HAND is produced using the latest hydro-conditioning techniques to enforce monotonically decreasing elevations including stream burning, levee enforcement, pit-filling, stream channel excavation, thalweg breaching, headwater seeding, stream reach resampling, and more. Finally, we used this implementation to investigate the skill of the FIMs by varying the scale of the processing units used to derive HAND. FIM skill was evaluated over large spatial scales by comparison to HEC-RAS 1D models.

The main contribution and conclusion of this work centers around a fundamental limitation in HAND based FIM which is a failure to account for multiple possible sources of fluvial inundation since HAND only considers inundation from the nearest drainage line. We illustrate that reducing the Horton-Strahler stream order of a HAND processing unit down to one enhances skill by significantly reducing false negatives at junctions of major streams. In order to reduce stream order of the NWM stream network for HAND computation, we dissected the NWM network into two simpler units of singular Horton-Strahler stream order and mosaiced the resulting FIMs derived from each. The NWM Mainstems (MS) stream network, which covers roughly 4-5% of the NWM Full Resolution (FR) network, spans all established forecasting points in the Advanced Hydrologic Prediction System (AHPS) and downstream reaches. The inundation from MS derived

HAND is mosaiced together with the inundation of FR derived HAND. Extending order reduction to the entire network, the Generalized Mainstem (GMS) technique discretizes the NWM FR network into level paths (LP) of unit stream order for HAND computation. All LP based FIM derived from LP based HAND datasets are mosaiced together to form one seamless FIM. Dissecting the stream network into regions of LPs with unit stream order is necessary because HAND has a "nearest drainage" limitation meaning it only accounts for riverine inundation sourced from the nearest drainage line. In our evaluation of this technique, we observe that HAND based FIM improves in skill as we extend from nearest drainage inundation in FR to multiple drainage support in MS for only 4-5% of the FR network. Extending multiple drainage support to the entire FR network with GMS based HAND improves skill at around the same magnitude that MS improved upon FR. Thus we conclude that deriving HAND with independent stream networks of unit Horton-Strahler stream order enhances the skill of FIM but offers diminishing returns as we extend from 4-5% of the network with MS to 100% of the network with GMS since deriving HAND and FIMs at these localized scale does add computational expense.

This primary contribution also affects the parameters used to compute stage-discharge relationships shifting discharge higher at given stages which reduced the rate of false positives. This shift in SRC behavior is driven by larger catchments that influence reach averaged geometric parameters in the Manning's equation. Related to SRCs, we noted better results and more sensitivity to unit stream order networks with the higher Manning's n value of 0.12 when compared to 0.06 for high magnitude events at 1% (100 year or yr) and 0.2% (500 year or yr) recurrence intervals. Additionally, we noted better performance for more intense 500 yr events which we attribute to a stronger influence of poor quality bathymetric data in 100 yr magnitude inundation extents. While the AGREE DEM procedure is meant to add some bathymetry primarily motivated to enhance catchment and stream line delineation, it does introduce three parameters that have major implications in the quality of SRCs and the resulting FIMs. Utilizing the highest resolution Lidar and bathymetric data should also improve the vertical accuracy of HAND and better account for fine grain features that greatly affect inundation extents. We leave questions related to Manning's n localization as well as bathymetry integration, estimation, and/or calibration open for future research to answer. Two other issues left open for improvement include the integration of higher resolution Lidar-based digital elevation maps

1141   (DEM) as well as the use of physics-based models for continental scale, high resolution

1142   forecasting applications. Due to inherent limitations with HAND, scalable, physics-based

1143   methods are necessary to consider to provide a better representation of flood extent dy-

1144   namics in steady and unsteady conditions.

Open Research

National Water Model (NWM) data used in this study includes the hydrofabric related datasets (*NWM Hydrofabric V2.1*, 2021) including catchments, streamlines, and reservoirs (*NWM Hydrofabric V2.1*, 2021). These are furnished by the National Oceanic and Atmospheric Administration (NOAA) Office of Water Prediction (OWP) via an Earth Science Information Partners (ESIP) Amazon Web Services (AWS) S3 Bucket (Aristizabal et al., 2022a). OWP Flood Inundation Mapping (FIM) capabilities rely extensively on the use of the National Hydrography Plus High Resolution (NHDPlusHR) datasets including BurnLineEvents (*NHDPlusHR GDB*, 2021), value-added attributes (VAA) (*NHDPlusHR GDB*, 2021), water boundaries (WBD) or hydrologic unit code (HUC) geometries (*NHDPlusHR WBD*, 2021), and digital elevation maps (DEM) (*NHDPlusHR DEM*, 2021). Some additional datasets for processing include the National Levee Database (NLD) furnished by the United States Army Core of Engineers (USACE) (ENGINEERS, 2021), Land-sea border from the Great Lakes Hydrography Dataset (GLHD) furnished by the Great Lakes Aquatic Habitat Framework (GLAHF) (*GLHD*, 2020), and a Land-sea border provided by OpenStreetMap (OSM) (*Water polygons*, 2021). Evaluation data was furnished by Interagency Flood Risk Management (InFRM) consortium including cross-sections and flood depths (*Base Level Engineering (BLE) Tools and Resources*, 2021; *estBFE Viewer*, 2021). Additionally, some FIM hydrofabric data including HAND grids, catchments, streamlines, synthetic rating curves, and cross-walk tables are available on the ESIP bucket (Aristizabal et al., 2022a).

Software used in preprocessing data, producing FIM hydrofrabic, generating FIM, computing metrics, and conducting analysis is available from a publicly available Github repository called "inundation-mapping" from the "NOAA-OWP" organization (Aristizabal et al., 2022b).

Acknowledgments

References

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

Amer, A., Balaji, P., Bland, W., Gropp, W., Guo, Y., Latham, R., . . . others (2021). *Mpich user's guide.* Mathematics and Computer Science Division at Argonne National Laboratory. (Version 3.4.1)

Aristizabal, F., Bates, B., Avant, B., Chadwick, N., Grout, T., Spies, R., . . . Cocks, G. (2022a). *noaa-nws-owp-fim.* Earth Science Information Partners (ESIP). Retrieved from `s3://noaa-nws-owp-fim/`

Aristizabal, F., Bates, B., Avant, B., Chadwick, N., Grout, T., Spies, R., . . . Cocks, G. (2022b, Feb). *Noaa-owp/inundation-mapping (v4.0.0.0)* [Software]. GitHub. Retrieved from `https://github.com/NOAA-OWP/inundation-mapping`

Aristizabal, F., & Judge, J. (2021). Mapping fluvial inundation extents with graph signal filtering of river depths determined from unsupervised clustering of synthetic aperture radar imagery. In *2021 ieee international geoscience and remote sensing symposium igarss* (pp. 6124–6127).

Aristizabal, F., Judge, J., & Monsivais-Huertero, A. (2020). High-resolution inundation mapping for heterogeneous land covers with synthetic aperture radar and terrain data. *Remote Sensing*, *12*(6), 900.

Arundel, S., Bulen, A., Adkins, K., Brown, R., Lowe, A., Mantey, K., & Phillips, L. (2018). Assimilation of the national elevation dataset and launch of the 3d elevation program through the usgs spatial data infrastructure. *International Journal of Cartography*, *4*(2), 129–150.

Baker, M. E., Weller, D. E., & Jordan, T. E. (2006). Comparison of automated watershed delineations. *Photogrammetric Engineering & Remote Sensing*, *72*(2), 159–168.

Barnes, R. (2018). *Richdem: High-performance terrain analysis* (Tech. Rep.). PeerJ Preprints.

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

Barnes, R., Lehman, C., & Mulla, D. (2014b). Priority-flood: An optimal depression-filling and watershed-labeling algorithm for digital elevation models. *Computers & Geosciences*, *62*, 117–127.

Base level engineering (ble) tools and resources. (2021, Apr). Interagency Flood Risk Management Consortium, US Federal Emergency Management Agency. Retrieved from `https://www.fema.gov/media-collection/base-level-engineering-ble-tools-and-resources`

Bedient, P. B., Huber, W. C., Vieux, B. E., et al. (2008). *Hydrology and floodplain analysis.* Prentice Hall Upper Saddle River, NJ.

contributors, G. (2020). Gdal/ogr geospatial data abstraction software library [Computer software manual]. Retrieved from `https://gdal.org` (Version 3.1.2)

Cook, A., & Merwade, V. (2009). Effect of topographic data, geometric configuration and modeling approach on flood inundation mapping. *Journal of hydrology*, *377*(1-2), 131–142.

Corringham, T. W., & Cayan, D. R. (2019). The effect of el niño on flood damages in the western united states. *Weather, Climate, and Society*, *11*(3), 489–504.

Cosgrove, B., Gochis, D., Graziano, T. M., Clark, E. P., & Flowers, T. (2019). The evolution of noaa's national water model: An overview of version 2.1 and future operational plans. *AGUFM*, *2019*, H51D–01.

Criss, R. E., & Nelson, D. L. (2022). Stage-based flood inundation mapping. *Natural Hazards*, 1–17.

Di Baldassarre, G., & Claps, P. (2011). A hydraulic study on the applicability of flood rating curves. *Hydrology Research*, *42*(1), 10–19.

Djokic, D. (2019). *Arc hydro: Developing hand from nhdplushr and nwm data – detailed workflow* (1st ed.). Environmental System Research Institute.

Dobbs, K. E. (2010). *Evaluation of the usgs national elevation dataset and the kansas biological survey's fldpln (" floodplain") model for inundation extent estimation* (Unpublished doctoral dissertation). University of Kansas.

Donchyts, G., Winsemius, H., Schellekens, J., Erickson, T., Gao, H., Savenije, H., & van de Giesen, N. (2016). Global 30m height above the nearest drainage. *HAND*, *1000*(0).

Downton, M. W., Miller, J. Z. B., & Pielke Jr, R. A. (2005). Reanalysis of us national weather service flood loss database. *Natural Hazards Review*, *6*(1), 13–

22.

ENGINEERS, U. A. C. O.   (2021, Mar).   *National levee database.*   Retrieved from
https://levees.sec.usace.army.mil/#/

estbfe viewer.   (2021, Apr).   Interagency Flood Risk Management Consortium,
US Federal Emergency Management Agency.   Retrieved from https://
webapps.usgs.gov/infrm/estBFE/

Follum, M. L., Tavakoly, A. A., Niemann, J. D., & Snow, A. D.   (2017).   Autorapid:
a model for prompt streamflow estimation and flood inundation mapping over
regional to continental extents.   *JAWRA Journal of the American Water
Resources Association*, *53*(2), 280–299.

Garbrecht, J., & Martz, L. W. (1997). The assignment of drainage direction over flat
surfaces in raster digital elevation models. *Journal of hydrology*, *193*(1-4), 204–
213.

Garousi-Nejad, I., Tarboton, D. G., Aboutalebi, M., & Torres-Rua, A. F.   (2019).
Terrain analysis enhancements to the height above nearest drainage flood
inundation mapping method. *Water Resources Research*, *55*(10), 7983–8009.

Gauckler, P. (1867). *Etudes théoriques et pratiques sur l'ecoulement et le mouvement
des eaux.* Gauthier-Villars.

Gerapetritis, H., & Pelissier, J. M. (2004). On the behavior of the critical success in-
dex.

Gesch, D., Oimoen, M., Greenlee, S., Nelson, C., Steuck, M., & Tyler, D.   (2002).
The national elevation dataset.   *Photogrammetric engineering and remote
sensing*, *68*(1), 5–32.

Glhd. (2020). Great Lakes Aquatic Habitat Framework contributors. Retrieved from
https://www.glahf.org/download/159/

Gochis, D., Dugger, A., Barlage, M., Cabell, R., FitzGerald, K., McAllister, M.,
. . . Zhang, Y.   (2021).   *The wrf-hydro modeling system technical descrip-
tion,(version 5.2)* (Tech. Rep.).

Godbout, L., Zheng, J. Y., Dey, S., Eyelade, D., Maidment, D., & Passalacqua, P.
(2019).   Error assessment for height above the nearest drainage inundation
mapping.   *JAWRA Journal of the American Water Resources Association*,
55(4), 952–963.

Hellweger, F., & Maidment, D.   (1997).   *Agree-dem surface reconditioning system.* University of Texas, Austin.   Retrieved from `https://www.caee.utexas.edu/prof/maidment/gishydro/ferdi/research/agree/agree.html`

Horton, R. E.   (1945).   Erosional development of streams and their drainage basins; hydrophysical approach to quantitative morphology.   *Geological society of America bulletin*, *56*(3), 275–370.

Huang, C., Nguyen, B. D., Zhang, S., Cao, S., & Wagner, W.   (2017).   A comparison of terrain indices toward their ability in assisting surface water mapping from sentinel-1 data. *ISPRS International Journal of Geo-Information*, *6*(5), 140.

Jenson, S. K., & Domingue, J. O.   (1988).   Extracting topographic structure from digital elevation data for geographic information system analysis.   *Photogrammetric engineering and remote sensing*, *54*(11), 1593–1600.

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

Jolliffe, I. T., & Stephenson, D. B.   (2012).   *Forecast verification: a practitioner's guide in atmospheric science.* John Wiley & Sons.

Jordahl, K.   (2014).   Geopandas: Python tools for geographic data.   *URL: https://github. com/geopandas/geopandas*.

Kunkel, K. E., Pielke Jr, R. A., & Changnon, S. A. (1999). Temporal fluctuations in weather and climate extremes that cause economic and human health impacts: A review. *Bulletin of the American Meteorological Society*, *80*(6), 1077–1098.

Lam, S. K., Pitrou, A., & Seibert, S. (2015). Numba: A llvm-based python jit compiler. In *Proceedings of the second workshop on the llvm compiler infrastructure in hpc* (pp. 1–6).

Li, Z., Mount, J., & Demir, I.   (2020).   Evaluation of model parameters of hand model for real-time flood inundation mapping: Iowa case study.   *EarthArXiv. July*, *1*.

Lindsay, J. B., & Seibert, J.   (2013).   Measuring the significance of a divide to local drainage patterns.   *International Journal of Geographical Information Science*, *27*(7), 1453–1468.

Liu, Y., Tarboton, D. G., & Maidment, D. R. (2020). *Height above nearest drainage (hand) and hydraulic property table for conus* (Tech. Rep.). Oak Ridge National Lab.(ORNL), Oak Ridge, TN (United States). Oak Ridge . . . .

Liu, Y., Yin, D., Gao, Y., & Muite, B. (2016). *Cybergis toolkit.* `https://github.com/cybergis/cybergis-toolkit`. GitHub.

Liu, Y. Y., Maidment, D. R., Tarboton, D. G., Zheng, X., Yildirim, A., Sazib, N. S., & Wang, S. (2016). A cybergis approach to generating high-resolution height above nearest drainage (hand) raster for national flood mapping.

Maidment, D. R. (2017). Conceptual framework for the national flood interoperability experiment. *JAWRA Journal of the American Water Resources Association*, *53*(2), 245-257. Retrieved from `https://onlinelibrary.wiley.com/doi/abs/10.1111/1752-1688.12474` doi: 10.1111/1752-1688.12474

Mallakpour, I., & Villarini, G. (2015). The changing nature of flooding across the central united states. *Nature Climate Change*, *5*(3), 250–254.

Manning, R., Griffith, J. P., Pigot, T., & Vernon-Harcourt, L. F. (1890). *On the flow of water in open channels and pipes.*

McEnery, J., Ingram, J., Duan, Q., Adams, T., & Anderson, L. (2005). Noaa's advanced hydrologic prediction service: building pathways for better science in water forecasting. *Bulletin of the American Meteorological Society*, *86*(3), 375–386.

McGehee, R., Li, L., & Poston, E. (2016). The modified hand method. In D. R. Maidment, A. Rajib, P. Lin, & E. P. Clark (Eds.), *National water center innovators program summer institute report 2016* (Vol. 4).

McKay, L., Bondelid, T., Dewald, T., Johnston, J., Moore, R., & Rea, A. (2012). *Nhdplus version 2: User guide; national operational hydrologic remote sensing center: Washington, dc, 2012.*

Merkel, D. (2014). Docker: lightweight linux containers for consistent development and deployment. *Linux journal*, *2014*(239), 2.

Milly, P. C. D., Wetherald, R. T., Dunne, K., & Delworth, T. L. (2002). Increasing risk of great floods in a changing climate. *Nature*, *415*(6871), 514–517.

Mizgalewicz, P. J., & Maidment, D. R. (1996). *Modeling agrichemical transport in midwest rivers using geographic information systems* (Unpublished doctoral dissertation). Center for Research in Water Resources, University of Texas at

Austin.

Moore, R. B., McKay, L. D., Rea, A. H., Bondelid, T. R., Price, C. V., Dewald, T. G., & Johnston, C. M. (2019). *User's guide for the national hydrography dataset plus (nhdplus) high resolution* (Tech. Rep.). US Geological Survey.

National Weather Service. (2018, Apr). *Summary of natural hazard statistics for 2017 in the united states.* NOAA. Retrieved from `https://www.weather.gov/media/hazstat/sum17.pdf`

National Weather Service. (2019, Apr). *Summary of natural hazard statistics for 2018 in the united states.* NOAA. Retrieved from `https://www.weather.gov/media/hazstat/sum19.pdf`

Nhdplushr dem. (2021, June). United States Geological Survey (USGS). Retrieved from `https://prd-tnm.s3.amazonaws.com/index.html?prefix=StagedProducts/Hydrography/NHDPlusHR/Beta/GDB/`

Nhdplushr gdb. (2021, June). United States Geological Survey (USGS). Retrieved from `https://prd-tnm.s3.amazonaws.com/index.html?prefix=StagedProducts/Hydrography/NHDPlusHR/Beta/GDB/`

Nhdplushr wbd. (2021, June). United States Geological Survey (USGS). Retrieved from `https://prd-tnm.s3.amazonaws.com/StagedProducts/Hydrography/WBD/National/GDB/WBD_National_GDB.zip`

Nobre, A., Cuartas, L., Hodnett, M., Rennó, C., Rodrigues, G., Silveira, A., . . . Saleska, S. (2011). Height above the nearest drainage – a hydrologically relevant new terrain model. *Journal of Hydrology*, *404*(1), 13 - 29. Retrieved from `http://www.sciencedirect.com/science/article/pii/S0022169411002599` doi: https://doi.org/10.1016/j.jhydrol.2011.03.051

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

Nwm hydrofabric v2.1. (2021, Dec). Office of Water Prediction (OWP) and Earth Science Information Partners (ESIP). Retrieved from `s3://noaa-nws-owp-fim/hand_fim/fim_3_0_34_1/inputs/nwm_hydrofabric/`

O'Callaghan, J. F., & Mark, D. M. (1984). The extraction of drainage networks from digital elevation data. *Computer vision, graphics, and image processing*, *28*(3), 323–344.

Petrochenkov, G. (2020). pygft: Rapid flood inundation modeling tool [Computer software manual]. Reston, VA: U.S. Geological Survey. Retrieved from https://code.usgs.gov/gft/python-gis-flood-tool

Pielke Jr, R. A., & Downton, M. W. (2000). Precipitation and damaging floods: Trends in the united states, 1932–97. *Journal of climate*, *13*(20), 3625–3637.

Planchon, O., & Darboux, F. (2002). A fast, simple and versatile algorithm to fill the depressions of digital elevation models. *Catena*, *46*(2-3), 159–176.

Ponce, V. M., & Changanti, P. (1994). Variable-parameter muskingum-cunge method revisited. *Journal of Hydrology*, *162*(3-4), 433–439.

Quenzer, A. M. (1998, May). A gis assessment of the total loads and water quality in the corpus christi bay system.

Quintero, F., Rojas, M., Muste, M., Krajewski, W. F., Perez, G., Johnson, S., . . . Zogg, J. (2021). Development of synthetic rating curves: Case study in iowa. *Journal of Hydrologic Engineering*, *26*(1), 05020046.

Rajib, A., Merwade, V., & Liu, Z. (2016). Large scale high resolution flood inundation mapping in near real-time. In *Proceedings of the 40th anniversary of the association of state flood plain managers national conference, gran rapids, mi, usa* (pp. 19–24).

Rennó, C. D., Nobre, A. D., Cuartas, L. A., Soares, J. V., Hodnett, M. G., & Tomasella, 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.

Salas, F. R., Somos-Valenzuela, M. A., Dugger, A., Maidment, D. R., Gochis, D. J., David, C. H., . . . Noman, N. (2018). Towards real-time continental scale streamflow simulation in continuous and discrete space. *JAWRA Journal of the American Water Resources Association*, *54*(1), 7–27.

Sanders, B. F. (2007). Evaluation of on-line dems for flood inundation modeling. *Advances in water resources*, *30*(8), 1831–1843.

Saunders, W. (1999). Preparation of dems for use in environmental modeling analysis. In *Esri user conference* (p. 24-30).

Saunders, W., & Maidment, D. (1995). Grid-based watershed and stream network delineation for the san antonio-nueces coastal basin. *Proceedings of Texas Water*, *95*, 16–17.

Saunders, W. K., & Maidment, D. R.   (1996). *A gis assessment of non-point source pollution in the san antonio-nueces coastal basin* (Tech. Rep.).   Center for Research in Water Resources, University of Texas at Austin.

Schaefer, J. T.   (1990).   The critical success index as an indicator of warning skill. *Weather and forecasting*, *5*(4), 570–575.

Service, N. W. (2020a, Nov). *Nws preliminary us flood fatality statistics.* NOAA's National Weather Service.   Retrieved from `https://www.weather.gov/arx/usflood`

Service, N. W.   (2020b, Jun). *Summary of natural hazard statistics for 2019 in the united states.*   NOAA.   Retrieved from `https://www.weather.gov/media/hazstat/sum19.pdf`

Shastry, A., Egbert, R., Aristizabal, F., Luo, C., Yu, C.-W., & 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.

Slater, L. J., & Villarini, G. (2016). Recent trends in us flood risk. *Geophysical Research Letters*, *43*(24), 12–428.

Stephens, E., Schumann, G., & Bates, P. (2014). Problems with binary pattern measures for flood model evaluation. *Hydrological Processes*, *28*(18), 4928–4937.

Strahler, A. N. (1952). Hypsometric (area-altitude) analysis of erosional topography. *Geological Society of America Bulletin*, *63*(11), 1117–1142.

Survila, K., Li, T., Liu, Y. Y., Tarboton, D. G., & Wang, S. (2016). A scalable high-performance topographic flow direction algorithm for hydrological information analysis.   In *Proceedings of the xsede16 conference on diversity, big data, and science at scale* (pp. 1–7).

Tabari, H.   (2020).   Climate change impact on flood and extreme precipitation increases with water availability.   *Scientific Reports*, *10*(1), 13768.   Retrieved from `https://doi.org/10.1038/s41598-020-70816-2`   doi: 10.1038/s41598-020-70816-2

Tange, O.   (2015).   Gnu parallel-the command-line power tool.; login: The usenix magazine, 36 (1): 42–47, feb 2011. *URL http://www. gnu. org/s/parallel*, *101*.

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.

Tarboton, D. G. (2005). Terrain analysis using digital elevation models (taudem). *Utah State University, Logan*.

Tarboton, D. G., Schreuders, K., Watson, D., & Baker, M. (2009). Generalized terrain-based flow analysis of digital elevation models. In *Proceedings of the 18th world imacs congress and modsim09 international congress on modelling and simulation, cairns, australia* (Vol. 20002006).

Team, G. D. (2020). Geographic resources analysis support system (grass gis) software [Computer software manual]. USA. Retrieved from `https://grass.osgeo.org`

Team, P. C. (2019). Python: A dynamic, open source programming language [Computer software manual]. Retrieved from `https://www.python.org/` (Python version 3.8.2)

Teng, J., Jakeman, A. J., Vaze, J., Croke, B. F., Dutta, D., & Kim, S. (2017). Flood inundation modelling: A review of methods, recent advances and uncertainty analysis. *Environmental Modelling & Software*, *90*, 201–216.

Teng, J., Vaze, J., Dutta, D., & Marvanek, S. (2015). Rapid inundation modelling in large floodplains using lidar dem. *Water Resources Management*, *29*(8), 2619–2636.

Tesfa, T. K., Tarboton, D. G., Watson, D. W., Schreuders, K. A., Baker, M. E., & Wallace, R. M. (2011). Extraction of hydrological proximity measures from dems using parallel processing. *Environmental Modelling & Software*, *26*(12), 1696–1709.

Twele, A., Cao, W., Plank, S., & Martinis, S. (2016). Sentinel-1-based flood mapping: a fully automated processing chain. *International Journal of Remote Sensing*, *37*(13), 2990–3004.

USACE. (2022). *Hydrologic engineering centers river analysis system (hec-ras).* Retrieved from `https://www.hec.usace.army.mil/software/hec-ras/`

Verdin, J., Verdin, K., Mathis, M. L., Magadzire, T., Kabuchanga, E., Woodbury, M., & Gadain, H. (2016). *A software tool for rapid flood inundation mapping* (Tech. Rep.). US Geological Survey.

Wallis, C., Watson, D., Tarboton, D., & Wallace, R. (2009). Parallel flow-direction and contributing area calculation for hydrology analysis in digital elevation

1478   models. *power*, *11*(8), 7.

1479 Wang, Y., & Zheng, T. (2005). Comparison of light detection and ranging and

1480   national elevation dataset digital elevation model on floodplains of north car-

1481   olina. *Natural Hazards Review*, *6*(1), 34–40.

1482 Warmerdam, F. (2008). The geospatial data abstraction library. In *Open source ap-*

1483   *proaches in spatial data handling* (pp. 87–104). Springer.

1484 *Water polygons.* (2021). Open Street Map (OSM). Retrieved from `https://osmdata`

1485   `.openstreetmap.de/download/water-polygons-split-4326.zip`

1486 Wing, O. E., Bates, P. D., Sampson, C. C., Smith, A. M., Johnson, K. A., & Erick-

1487   son, T. A. (2017). Validation of a 30 m resolution flood hazard model of the

1488   conterminous u nited s tates. *Water Resources Research*, *53*(9), 7968–7986.

1489 Wing, O. E., Bates, P. D., Smith, A. M., Sampson, C. C., Johnson, K. A., Fargione,

1490   J., & Morefield, P. (2018). Estimates of present and future flood risk in the

1491   conterminous united states. *Environmental Research Letters*, *13*(3), 034023.

1492 Yamazaki, D., Ikeshima, D., Sosa, J., Bates, P. D., Allen, G. H., & Pavelsky, T. M.

1493   (2019). Merit hydro: a high-resolution global hydrography map based on latest

1494   topography dataset. *Water Resources Research*, *55*(6), 5053–5073.

1495 Zhang, J., Huang, Y.-F., Munasinghe, D., Fang, Z., Tsang, Y.-P., & Cohen, S.

1496   (2018). Comparative analysis of inundation mapping approaches for the

1497   2016 flood in the brazos river, texas. *JAWRA Journal of the American Water*

1498   *Resources Association*, *54*(4), 820–833.

1499 Zheng, X., Maidment, D. R., Tarboton, D. G., Liu, Y. Y., & Passalacqua, P. (2018).

1500   Geoflood: Large-scale flood inundation mapping based on high-resolution

1501   terrain analysis. *Water Resources Research*, *54*(12), 10–013.

1502 Zheng, X., Tarboton, D. G., Maidment, D. R., Liu, Y. Y., & Passalacqua, P. (2018).

1503   River channel geometry and rating curve estimation using height above the

1504   nearest drainage. *JAWRA Journal of the American Water Resources Associa-*

1505   *tion*, *54*(4), 785–806.

1506 Zhou, G., Sun, Z., & Fu, S. (2015). *Filldem.* `https://github.com/zhouguiyun`

1507   `-uestc/FillDEM`. GitHub.

1508 Zhou, G., Sun, Z., & Fu, S. (2016). An efficient variant of the priority-flood al-

1509   gorithm for filling depressions in raster digital elevation models. *Computers &*

1510   *Geosciences*, *90*, 87–96.