Effects of High-Quality Elevation Data and Explanatory Variables on the Accuracy of Flood Inundation Mapping via Height Above Nearest Drainage

Fernando Aristizabal[1,2,3], Taher Chegini[4], Gregory Petrochenkov[1,2], Fernando Salas[2], and Jasmeet Judge[3]

[1]Lynker, 338 E Market St, Leesburg, VA, 20176, USA
[2]National Water Center, Office of Water Prediction, National Oceanic and Atmospheric Administration, 205 Hackberry Ln, Tuscaloosa, AL, 35401, USA
[3]Center for Remote Sensing, Agricultural and Biological Engineering, University of Florida, 1741 Museum Rd, Gainesville, FL, 32603, USA
[4]Civil and Environmental Engineering, University of Houston, 4226 Martin Luther King Boulevard, Houston, TX, 77204, USA

Correspondence: Fernando Aristizabal (fernando.aristizabal@noaa.gov)

Abstract. Given the availability of high quality and high spatial resolution digital elevation models (DEMs) from the United States Geological Survey's 3-Dimensional Elevation Program (3DEP) derived from mostly Light Detection and Ranging sensors, we examined the effects of these DEMs at various spatial resolutions on the quality of flood inundation map (FIM) extents derived from a terrain index known as Height Above Nearest Drainage (HAND). We found that using these DEMs improved the quality of resulting FIMs at around 80% of the catchments analyzed when compared to using DEMs from the National Hydrography Dataset Plus High Resolution program. Additionally, we varied the spatial resolution of the 3DEP DEMs from 3, 5, 10, 15, and 20 meters and the results showed no significant overall effect on FIM extent quality across resolutions. However, further analysis at coarser resolutions of 60 and 90 meters revealed a significant degradation in FIM skill, highlighting the limitations of using extremely coarse resolution DEMs. Our experiments demonstrated a significant burden on the computational time to produce HAND and related data at finer resolutions. We fit a multiple linear regression model to help explain catchment scale variation in the four metrics employed and found that the lack of reservoir flooding, or inundation upstream of river retention systems, was a significant factor in our analysis. For validation, we used Interagency Flood Risk Management Base Level Engineering produced FIM extents and streamflows at the 100 and 500 year event magnitudes in a sub-region in Eastern Texas.

1 Introduction

Floods are among the most frequent, damaging, and deadly of natural disasters (Doocy et al., 2013; Strömberg, 2007; Kahn, 2005). The frequency and intensity of flood events as well as the exposure of people and property to them have been increasing in recent times driven by secular changes in climate, infrastructure, and demographics (Berz, 2000; Mallakpour and Villarini, 2015; Downton et al., 2005; Kunkel et al., 1999; Pielke Jr and Downton, 2000; Corringham and Cayan, 2019; Gourevitch

et al., 2023). These upward trends are expected to continue placing additional pressure on hydrological extremes (Kahn, 2005; Tabari, 2020; Milly et al., 2002; Wing et al., 2018; Gourevitch et al., 2023). Floods impact mortality and morbidity through drowning or physical trauma at the individual health scale, while increasing the risk of infectious disease at the public health level (Jonkman, 2005; Beinin, 2012; Alajo et al., 2006; French et al., 1983). Flooding disrupts systems providing human needs such as transportation routes, supply chains, water delivery, waste management, communications, shelter, and energy grids (Wijkman and Timberlake, 2021; Gourevitch et al., 2023). These impacts disproportionately affect certain demographics such as the socioeconomically-disadvantaged, youth, and elderly who are more likely to live in vulnerable areas with less access to educational resources, early warning systems, and the capacity or resources to evacuate impacted areas (Kahn, 2005; Smiley et al., 2022; Strömberg, 2007; Jonkman, 2005; Tellman et al., 2020, 2021). These inequitable impacts further entrench poverty and inequalities (Stallings, 1988; Birkmann et al., 2010). In political terms, severe disasters, including floods, can reduce social order, strain governance systems, collapse social safety nets, and increase the risk of social conflict (Drury and Olson, 1998; Xu et al., 2016; Zahran et al., 2009). These dire consequences motivate adaption and mitigation efforts such as early warning systems, protective infrastructure (e.g. storage, defenses, drainage, infiltration), public awareness and education, and zoning regulations (Tumbare, 2000; Tauhid and Zawani, 2018; Charlesworth and Warwick, 2011).

Due to these growing flood risks, early warning systems, or forecasting systems, can help understand future conditions and provide intelligence to furnish adequate warnings to protect life, prevent damages, and enhance resilience (Strömberg, 2007; Cools et al., 2016; UNISDR, 2015; Baudoin et al., 2014; Golnaraghi, 2012; UNEP, 2012; Liu et al., 2018a; Schumann et al., 2013). The early warning of flood disasters at national scales often requires the use of continental-scale, forecast hydrology models and modeling frameworks that span intranational political boundaries. The applications of these models extend beyond early warning systems to provide historical trends for applications in infrastructure planning, public planning, insurance underwriting, and more. The Office of Water Prediction (OWP), an office of the National Oceanic and Atmospheric Administration (NOAA) along with partners at the National Center for Atmospheric Research (NCAR), developed such a continental-scale model known as the United States (US) National Water Model (NWM) (Salas et al., 2018; Gochis et al., 2021; Cosgrove et al., 2019; Cohen et al., 2018; Oceanic and Administration), 2016; of Water Prediction, 2022). The NWM is based on a configuration of the Weather Research and Forecasting Hydro (WRF-Hydro) model that accounts for land surface processes as well as overland and channel routing (Gochis et al., 2021; Salas et al., 2018; Cosgrove et al., 2019). Operationally, the NWM produces streamflow analysis and forecasts at multiple time horizons depending on location which include the conterminous United States (CONUS), Puerto Rico, Hawaii, and portions of Alaska (Cosgrove et al., 2019; Oceanic and Administration), 2016; of Water Prediction, 2022). The NWM routes streamflow across the NWM Version 2.1 (V2.1) stream network, based on the National Hydrography Dataset Plus Version 2 (NHDPlusV2) network, is comprised of more than 5.5 million kilometers (kms) of lines discretized into more than 2.8 million forecast points (Aristizabal et al., 2023c). The NWM V2.1 stream network belongs to the NWM "hydrofabric" defined as a catalog of geospatial layers relevant to hydrology modeling including stream network flowpaths, catchments, reservoirs, and more (of Water Prediction, 2022; Cosgrove et al., 2019). While streamflow is an important variable for engineering and scientific applications of fluvial flooding, flood inundation stages, extents and depths are much more tangible variables to the stakeholders flood events directly impact.

The shallow water equations, a system of two hyperbolic partial differential equations, formally govern the flow of fluvial surface water by conserving both mass (first equation) and momentum (second equation) and can be expressed in both the 1-Dimensional (1D) (Saint Venant Equations) and 2-Dimensional (2D) forms. Solving this system in full 2D form requires numerical methods that can be very cost prohibitive and numerically unstable in an operational setting across continental-scales at high spatial discretizations (10 meter (m) or higher). This use case motivates the implementation of an inundation proxy, also known as a zero-physics or a simplified conceptual model, that is agnostic to the shallow water equations while still computing accurate fluvial inundation extents and depths (Teng et al., 2015; Bates and De Roo, 2000). Height Above Nearest Drainage (HAND) detrends elevations within digital elevation models (DEMs) to compute drainage potentials by normalizing elevations to the nearest, relevant flowpath instead of datums that represent mean sea level (Rennó et al., 2008; Nobre et al., 2011, 2016). HAND as a terrain index has been used extensively for producing flood inundation maps (FIMs) from both modeled or observed stream flows and stages (Nobre et al., 2016; Afshari et al., 2018; Garousi-Nejad et al., 2019; Johnson et al., 2019; Zheng et al., 2018a, b; Zhang et al., 2018; Teng et al., 2015; Li et al., 2022b), as well as for assisting the remote sensing detection of fluvial inundation (Aristizabal et al., 2020; Shastry et al., 2019; Aristizabal and Judge, 2021; Huang et al., 2017; Twele et al., 2016). HAND operates as an inundation proxy by thresholding the relative elevation (or HAND) values with a singular river stage value for each catchment corresponding to the drainage area of a given river reach (Nobre et al., 2016; Garousi-Nejad et al., 2019; Johnson et al., 2019; Zheng et al., 2018a; Teng et al., 2015; Liu et al., 2016; Maidment, 2017; Liu et al., 2018b, 2020, 2018a). When used to generate inundation extents and depths from streamflow, reach-averaged synthetic rating curve (SRC) sample geometric variables along an entire reach and normalize using the length of the reach to create stage-discharge relationships (Zheng et al., 2018b; Aristizabal et al., 2023c; Godbout et al., 2019). These relationships depend on the friction parameter, Manning's n, and are used to convert streamflows to stages for eventual 2D mapping with HAND. Numerous investigations have validated the use of HAND for flood mapping applications as a suitable alternative to more sophisticated physics-based techniques for large scale and high resolution use cases (Johnson et al., 2019; Li et al., 2022a; Aristizabal et al., 2023c; Nobre et al., 2016; Godbout et al., 2019; Afshari et al., 2018; Zhang et al., 2018; Teng et al., 2015, 2017; Diehl et al., 2021; Hocini et al., 2021; Bates et al., 2003).

Several prior and active large-scale HAND implementations catered to operational early warning systems applications including the National Flood Interoperability Experiment (NFIE) (Maidment, 2017; Liu et al., 2016, 2018b), GeoFlood (Zheng et al., 2018a; Hocini et al., 2021; D'Angelo et al., 2022; Carruthers, 2021; Zheng et al., 2022), and PyGFT (Petrochenkov and Viger, 2020; Verdin et al., 2016). The NFIE was a broad, inter-institutional, and pioneering effort to apply HAND to the initial versions of the NWM which leveraged 1/3 arc-second (10 m) seamless elevation data available at the time (Maidment, 2017; Liu et al., 2016, 2018b) from the United States Geological Survey (US Geological Survey)'s National Elevation Dataset (NED) (Gesch et al., 2002; Gesch and Maune, 2007). Zheng et al. (2018a) applied HAND for operational applications with 1/27 arc-second (1 m) elevation data with a novel least cost, geodesic based stream delineation method (Passalacqua et al., 2010, 2012; Zheng et al., 2018a, 2019; Carruthers, 2021; D'Angelo et al., 2022; Zheng et al., 2022). For applications with the NWM, an advanced version of HAND coupled with the use of SRCs, known as OWP FIM, converts NWM analysis, reanalysis, and forecast streamflows to river stages and fluvial inundation depths and extents on an operational basis to CONUS

while extending the modeling domain to Puerto Rico and Hawaii (Aristizabal et al., 2023c, b). OWP FIM utilizes some of the latest datasets including the National Hydrography Dataset Plus High Resolution (NHDPlusHR) (Moore et al., 2019), National Levee Database (NLD) (of Engineers, 2021), and the NWM V2.1 hydrofabric (of Water Prediction, 2022; Oceanic and Administration), 2016; OWP/ESIP, 2021; Gochis et al., 2021). These datasets enforce hydrologically relevant features such as levees and the general location of flowpaths to facilitate conflation with the forecast stream network (Aristizabal et al., 2023c, b). Additionally, OWP FIM advanced a fundamental limitation of HAND that limits sourcing fluvial inundation only from the nearest, relevant flowpath (McGehee et al., 2016; Aristizabal et al., 2023c; Zhang et al., 2018; Li et al., 2022a; Zheng et al., 2018a, b; Nobre et al., 2016). Flowpaths of higher Horton-Strahler stream order that could contribute inundation to a given area have no way of extending beyond catchment lines which creates artificial bottlenecks in inundation extents, especially along junctions of high order rivers with their lower flow tributaries (Aristizabal et al., 2023c; McGehee et al., 2016). To resolve this limitation, OWP FIM disaggregates the NWM V2.1 stream network into segments of effective unit stream order called level paths in a version of HAND called Generalized Mainstems (GMS) (Aristizabal et al., 2023c). In terms of terrain data, OWP FIM uses the 10 m DEM from the NHDPlusHR elevation dataset which is the elevation basis, derived in batches from 3-Dimensional Elevation Program (3DEP), for additional hydrography products within the NHDPlusHR (Aristizabal et al., 2023c; Moore et al., 2019). The previous advances in OWP FIM stopped short of accounting for Light Detection and Ranging (LiDAR) point elevation observations (Aristizabal et al., 2023c) that are now nearing their first collection cycle to form a novel seamless, continental scale DEM from 3DEP (USGS, 2021, 2022).

Broad scale terrain information in the form of DEMs is fundamental to all FIM models and a significant influence on inundation skill (Bales and Wagner, 2009; Dobbs, 2010; Wang and Zheng, 2005; Merwade et al., 2008; Witt III, 2015; Garousi-Nejad et al., 2019; Li et al., 2022b; Neal et al., 2011). The National Geospatial Program, under the US Geological Survey, is the primary authority on collecting, processing, and maintaining terrestrial elevation data within the US in collaboration with Federal partners within the National Digital Elevation Program (NDEP) (Office of Management and Budget, 2016; Dewberry, 2011; Council et al., 2007, 2009; Sugarbaker et al., 2014). The NED (Gesch et al., 2002; Gesch and Maune, 2007), forms the seamless elevation layers of the The National Map (TNM) (Gesch et al., 2009; Archuleta et al., 2017; Arundel et al., 2015a, 2018; Kelmelis et al., 2003). Prior to the introduction of 3DEP, TNM was originally composed of three seamless DEMs at 1/3 (10 m), 1 (30 m), and 2 (90 m) arc-second resolutions produced from a variety legacy sources including digital photogrammetry, cartographic contours, mapped hydrography, and elevations from Shuttle Radar Topography Mission (SRTM) (Gesch et al., 2002; Gesch and Maune, 2007; Arundel et al., 2015a). High quality elevations derived from LiDAR and Interferometric Synthetic Aperture Radar (InSAR) have been integrated into TNM seamless elevation products as made available prior to and after the introduction of 3DEP (Snyder et al., 2013; Gesch et al., 2002; Arundel et al., 2015a). Work by Gesch et al. (2014),(Gesch and Maune, 2007), and Dobbs (2010) illustrated that the inclusion of higher quality elevation data sources had significant improvement in the accuracy of NED data when compared to the National Geodetic Survey (NGS) (Roman et al., 2010). Gesch et al. (2014) identified that the NED 1/3 arc-second DEM, as of April 2013, had a mean error of -0.29 m with an root mean squared error (RMSE) of 1.55 m when compared to over 25 thousand reference points. At the time of evaluation, the NED was subject to legacy, lower quality data sources dating almost a century in the past (Sugarbaker et al., 2014; Gesch et al., 2014;

Gesch and Maune, 2007). This reduction in error and its impact on people and commerce (Dewberry, 2011) motivated action on collection of elevation data from higher quality data sources (Sugarbaker et al., 2014).

3DEP is a national, multi-organizational effort by the NDEP to survey elevations with high quality sensors in response to growing stakeholder needs on a recurring collection cycle of no greater than 8 yrs (years) (Dewberry, 2011; Snyder et al., 2013; Sugarbaker et al., 2014). 3DEP leverages two main collection technologies including LiDAR for the CONUS, Hawaii, and US territories as well as InSAR for Alaska. LiDAR, the collection source of focus in this study, is a light emitting, reflection, and collection technology that beams concentrated powerful light of wavelengths between 1000 - 1600 nanometer (nm) (Muhadi et al., 2020). The reflection of the light is collected while recording the travel time and intensity of return. LiDAR sensors are mounted on top of a variety of mobile or static platforms whose positions are geo-tracked as they collect LiDAR returns (Passalacqua et al., 2015). The travel time of the returns, along with knowledge of the speed of light, serve as a relative positioning of the target(s) referenced to a common vertical datum while the intensities serve as indicators of what the target(s) represent. Modes within the relationship of return intensities with respect to travel time/distance from the LiDAR wave forms can be indicative of vegetation or other land use/land covers (LULCs) that reflect signals at varying distances and magnitudes and influence elevation errors (Gesch et al., 2014). These modes can be discretized into varying DEM products representing bare earth, structures, or canopy elevations. The horizontal and vertical accuracies and the horizontal resolutions of terrain observations derived from LiDAR, and even the consequential economic benefits (Dewberry, 2011, 2022), are dependent on a variety of sensor, platform, target, and collection specifications and practices such as nominal pulse spacing, nominal pulse density, and LULC of the target (Heidemann, 2018; Passalacqua et al., 2015; Smith et al., 2019; Salach et al., 2018; Gesch et al., 2014). LiDAR produces point cloud datasets which are scattered, geo-referenced points representing full wave forms or discretized return intensities. Various assessments of the vertical accuracies of LiDAR point clouds have yielded satisfactory results in agreement with 3DEP requirements (Stoker and Miller, 2022; Kim et al., 2022; Callahan and Berber, 2022; Kim et al., 2022; Salach et al., 2018; Passalacqua et al., 2015). Point clouds must undergo a series of operations to produce analysis ready, seamless DEMs (Passalacqua et al., 2015).

3DEP extends TNM to include a 1/27 arc-second (1 m), LiDAR derived DEM product for CONUS, Hawaii, and US territories as well as a 1/2 arc-second (5 m) DEM for Alaska derived from InSAR (Sugarbaker et al., 2014; Stoker et al., 2015). To create bare earth DEMs, LiDAR observations must undergo a series of processes that filter out returns from vegetation, anthropogenic, and other features then grid the observations with resampling methods (Passalacqua et al., 2015). The 1 m 3DEP product is a hydrologically conditioned (hydro-flattened), topographic, and bare-earth raster DEM gridded to 1 km square shaped tiles with 6 pixels of overlap (Arundel et al., 2015b). Hydro-flattening refers to a process in which hydrologic features such as lakes, reservoirs, streams, rivers, and more are flattened in elevation for bathymetric regions from lower bank to lower bank represented by breaklines (Archuleta et al., 2017; Maune and Nayegandhi, 2018). This flattening excludes along gradient directions, parallel to the direction of the breaklines, for hydrologic features that naturally exhibit water conveyance such as streams, rivers, and long reservoirs (Arundel et al., 2015b). This process includes elevations underneath bridges that are not accurately observed from topographic LiDAR (Bales and Wagner, 2009). According to specifications, the horizontal accuracy of 1 m 3DEP is within 1 m while the vertical accuracies are within 19.6 centimeters (cms) and 30 cm at the 95% confidence

interval for non-vegetative and vegetative regions, respectively (Arundel et al., 2015b; Heidemann, 2018). Non-vegetative vertical accuracies fall within 10 cm RMSE (Arundel et al., 2015b; Heidemann, 2018). Work by Stoker and Miller (2022), Callahan and Berber (2022), and Kim et al. (2022) have verified the vertical accuracies and general quality of the DEMs for 3DEP specifications.

The quality of FIM extents are subject to a wide variety of terrain related factors including collection technology, gridding methods, resampling techniques, hydrological conditioning processes, presence of bathymetry, vertical accuracies and horizontal resolutions (Merwade et al., 2008). The main enhancement of including 3DEP data within HAND based OWP FIM is the broader availability of high quality data sources for elevations such as LiDAR with enhanced vertical accuracies and horizontal resolutions (Arundel et al., 2015b; Stoker and Miller, 2022; Archuleta et al., 2017). Generally speaking, the literature has demonstrated the sensitivity to, and improved effect of, using 3DEP or LiDAR data on the quality of FIM extents, mostly due to the enhanced vertical accuracies that these data sources provide (Podhorányi and Fedorcak, 2015; Bales and Wagner, 2009; Merwade et al., 2008; Witt III, 2015; Mason et al., 2007; Zheng et al., 2018a). Limitations have been noted with respect to vertical accuracies in areas with vegetation, buildings, bridges, or classified as bathymetric (Merwade et al., 2008; Mason et al., 2007; Bales and Wagner, 2009; Podhorányi and Fedorcak, 2015). FIM extents in areas of low topographic relief or areas behind natural or anthropogenic flow divides can be very sensitive to vertical accuracies (Sanyal and Lu, 2004; Garousi-Nejad et al., 2019; Godbout et al., 2019; Jafarzadegan and Merwade, 2017; Papaioannou et al., 2017). Specifically for HAND, research by Zheng et al. (2018a) and Garousi-Nejad et al. (2019) noted improvement when utilizing higher resolution LiDAR derived DEMs for HAND based FIM.

The spatial resolution of topography likely interacts with many other sources of FIM extent uncertainties including but not limited to elevation source quality, LULC, streamflow intensities, physics employed, and model parameterizations (Fewtrell et al., 2008; Savage et al., 2016; Neal et al., 2011; Thomas Steven Savage et al., 2016). Numerous researchers have evaluated resolution more generally across the spectrum of FIM models to focus more on urban areas where resolution could play an integral part into determining extents (Fewtrell et al., 2008; Neal et al., 2011; Ozdemir et al., 2013; Muthusamy et al., 2021; Savage et al., 2016; de Almeida et al., 2018; Dixon and Earls, 2009). While the studies evaluating HAND are extensive (Afshari et al., 2018; Nobre et al., 2011; Garousi-Nejad et al., 2019; Godbout et al., 2019; Speckhann et al., 2018; McGrath et al., 2018; McGehee et al., 2016; Li et al., 2020, 2022b, a; Liu et al., 2016, 2018b, a; Li and Demir, 2022; Liu et al., 2020; Aristizabal et al., 2023c; Maidment, 2017; Zheng et al., 2018a, b, 2019; Diehl et al., 2021; Johnson et al., 2019; Jafarzadegan and Merwade, 2019), only a few studies have investigated the effects of high quality DEMs and their spatial resolutions on FIM extents when derived from HAND. Li et al. (2022b) evaluated HAND based FIM over a small domain and concluded that resampled LiDAR performed best at the 5 m spatial resolution when compared to coarser, resampled grids. Zheng et al. (2018a) incorporated LiDAR derived elevations while also incorporating a novel stream delineation method and concluded that both combined performed better than utilizing legacy NED 10 m datasets with NHDPlusV2 hydrography as the datum for HAND computation. Lastly, working in flat areas with some anthropogenic influence, Garousi-Nejad et al. (2019), used a 3 m DEM and found improvement in FIM quality extents when compared to the use of a 10 m DEM derived from different sources. In contrast to the other studies, Speckhann et al. (2018) evaluated the sensitivity of HAND based FIM extents to DEM

resolution in Brazil using DEMs from SRTM and found little to no effect given this region. Both Garousi-Nejad et al. (2019) and Zheng et al. (2018a) highlighted the importance of high resolution elevations and novel stream delineation tools to avoid the negative effects of little to no bathymetric information. Due to the interacting uncertainties and the dearth of research on this question with respect to HAND, it is difficult to conclude what the effect would be on the quality of HAND based OWP FIM by incorporating the latest 3DEP data at varying spatial resolutions.

As the spatial coverage of the 3DEP 1 m product rapidly approaches CONUS scale in 2023, we investigate the integration of 3DEP data into OWP FIM for model specific evaluation (USGS, 2021, 2022). We use 3DEP data for the HAND computation process to generate the FIM hydrofabric. OWP FIM uses novel combination of input datasets, hydrological conditioning (hydro-conditioning) processes, level path scale processing, and parameterizations to produce HAND and these specific combinations of methods could interact with terrain related variables including source and resolution. Additionally, we investigated the utility of varying spatial resolutions from 3, 5, 10, 15, and 20 m, specifically its effect on FIM extents. HAND depends on the drainage assumptions which requires DEMs to undergo a long series of enforcement processes to ensure monotonically decreasing elevations with hydrologically correct flow directions (Garousi-Nejad et al., 2019; Nobre et al., 2011, 2016; Aristizabal et al., 2023c). The resampling of DEMs into varying spatial resolutions could interact with these hydro-conditioning operations thus influencing the FIM hydrofabric and the resulting quality of the FIMs produced. OWP FIM is scheduled for public release in 2023 for a region covering 10% of the US population. Evaluations are needed specifically for this region and how 3DEP elevations at varying resolutions affect skill. As validation, we used 1D Hydrologic Engineering Center River Analysis Center (HEC-RAS) modeled flood inundation extents from both the Base Level Engineering (BLE) published by the Interagency Flood Risk Management (InFRM) team. By varying the spatial resolution of 3DEP DEMs, we seek to quantify the relationship in an empirical fashion between spatial resolution and FIM skill produced from HAND that requires significant DEM manipulations to satisfy inherent assumptions. For analysis purposes, we consider a series of potential, catchment scale explanatory variables with multi-variable regression analysis to help explain some the of the catchment scale variation in the metrics we employed that describe agreement with the BLE FIMs.

2 Materials and Methods

2.1 Overview

Investigating the effects of LiDAR derived DEMs and their spatial resolutions involved a multi-step process of data curation, production, evaluation, and analysis. Source information was gathered to produce HAND and its associated datasets most specifically the DEMs from multiple sources and spatial resolutions including 3, 5, 10, 15, and 20 m. Later in the analysis, the resolutions of 60 and 90 m were assessed to help identify if and when spatial resolution begin to influence FIM quality. The FIM hydrofabric, or the collection of datasets required to convert streamflows to FIM extents, was produced using these various DEMs (Aristizabal et al., 2023c; Aristizabal and Judge, 2021). FIMs were produced by intersecting the BLE cross-sections, as described by Aristizabal et al. (2023c), furnished by the InFRM team for both the 1% (100 year (yr)) and 0.2% (500 yr) recurrence flows (FEMA, 2016, 2021a, b; Strategic Alliance for Risk Reduction II, 2019a, b, c, d, e, f, g). These

intersected flows were converted to reach-averaged stages using SRCs (Aristizabal et al., 2023c; Liu et al., 2016, 2018b; Zheng et al., 2018b). Reach-averaged stages were used to threshold HAND values on a per catchment basis which translates to a flooded pixel when the stage value exceeds zero (Zheng et al., 2018b; Aristizabal et al., 2023c). Catchments are defined here as the unique surface drainage areas assigned to each river reach. These extents at the 100 and 500 yr flow magnitudes were then compared to the original BLE furnished extents for the corresponding magnitudes. Agreement statistics and maps were computed for binary categorical variables (inundated = positive and not inundated = negative) then resampled to the catchment scale. A number of covariates and factors were selected at the catchment scale for analysis purposes to explain some of the catchment to catchment variance in the selected metrics with the help of regression models. A high level graphical summary of this explanation is furnished in Figure 1.

2.2 Datasets

We used a wide variety of datasets for investigating the effects of DEM source and spatial resolution on FIMs produced from HAND. We compared legacy DEMs, used by Aristizabal et al. (2023c), sourced from the NHDPlusHR program that were available for the entire NWM domain at the Hydrologic Unit Code (HUC)-4 scale. As the 3DEP program rapidly approaches continental scale availability, the inclusion of 3DEP DEMs were considered here at various spatial resolutions (USGS, 2022; Stoker et al., 2015; USGS, 2021; Chegini et al., 2021; Survey, 2022). Aristizabal et al. (2023c) provided more details for the datasets listed in Figure 1 denoted by "Other Datasets". The remaining datasets used in this study for DEM experimentation and for analysis are elaborated on in Table 1. The analysis datasets are those used to help explain some of the spatial variation in the metrics. NWM catchments were used to resample the agreement maps and metrics down to the catchment scale. Analyses of the importance of various attributes within NWM catchments considered flowpath properties such as channel slope, length, presence of reservoirs, and catchment area. Other attributes, aggregated to the catchment scale for consideration, included terrain slope, imperviousness, overland roughness, and LULC from the National Land Cover Database (NLCD) which were used as either covariates or factors in the statistical analysis of this study.

2.3 Data Retrieval

We used Hydroclimate Data Retriever (HyRiver) (Chegini et al., 2021) for retrieving topographic, land use/land cover, imperviousness, and overland roughness data. HyRiver is a suite of nine open-source Python packages that provide access to a wide variety of hydrology and climatology datasets within the US, through web services. In this study, we used two of these packages: Py3DEP and Python Hydrogeological Datasets (PyGeoHydro).

Py3DEP provides access to 3DEP's dynamic and static services. The dynamic service retrieves topographic data at any resolution using the best available raw elevation data for a requested region, whereas the static service only provides DEM data at 10, 30, and 60 m resolutions. It has some other utilities, including querying availability of raw elevation data at different resolutions from various sources. In this study, we used the dynamic service to obtain DEM at different resolutions, and the raw data availability functionality of Py3DEP for determining the highest resolution data available in the region of our study.

Table 1. Dataset sources, names, descriptions, and citations. Datasets used to generate HAND except for the DEMs are not listed in detail here but explained by Aristizabal et al. (2023c).

Source	Dataset	Description	Citations
USGS	3DEP	Continental-scale high resolution (1 m) DEMs from high quality sources.	USGS (2022, 2021); Stoker et al. (2015); Chegini et al. (2021)
USGS	NHDPlusHR DEM	DEMs available from NHDPlusHR program at the HUC-4 level for the entire country.	USGS (2021); Moore et al. (2019)
Various	Other Datasets (See Figure 1.)	Other datasets used for production of HANDs based FIM hydrofabric.	Aristizabal et al. (2023c)
InFRM	Flood Inundation Extents	Inundation depths produced by InFRM BLE HEC-RAS 1D for 1% and 0.2% recurrence interval events.	FEMA (2021b, 2016, 2021a); Strategic Alliance for Risk Reduction II (2019a, b, c, d, e, f, g)
InFRM	Cross-Sections	BLE HEC-RAS 1D cross-sections for 100 yr (1%) and 500 yr (0.2%) streamflow magnitudes used to intersect with NWM reaches.	FEMA (2021b, 2016, 2021a); Strategic Alliance for Risk Reduction II (2019a, b, c, d, e, f, g)
OWP	NWM Catchments	Surface drainage area corresponding to each reach in the NWM adapted from National Hydrography Dataset Plus (NHDPlus) V2 Catchment feature class.	OWP/ESIP (2021)
OWP	NWM Stream Network	Stream network flowpaths used by NWM for routing and forecasting adapted from NHDPlus V2 NHDFlowline Network feature class.	OWP/ESIP (2021)
USGS	Terrain Slope	Terrain slope (vertical/horizontal) computed from 3DEP DEMs.	Dewitz (2021); Yang et al. (2018); Chegini et al. (2021); Survey (2022)
NLCD	LULC	LULC as produced by the NLCD 2019 at the 30 m resolution derived partly from LandSat imagery.	Dewitz (2021); Yang et al. (2018); Chegini et al. (2021); Multi-Resolution Land Characteristics Consortium (2022)
NLCD	Imperviousness	Urban impervious surface as a percentage of developed surface over 30 m pixels.	Dewitz (2021); Yang et al. (2018); Chegini et al. (2021); Multi-Resolution Land Characteristics Consortium (2022)
NLCD	Overland Roughness	Overland Roughness or Manning's n for given pixel within NLCD.	Dewitz (2021); Yang et al. (2018); Chow (1959); Chegini et al. (2021); Multi-Resolution Land Characteristics Consortium (2022); McCuen et al. (2005); Kalyanapu et al. (2009)

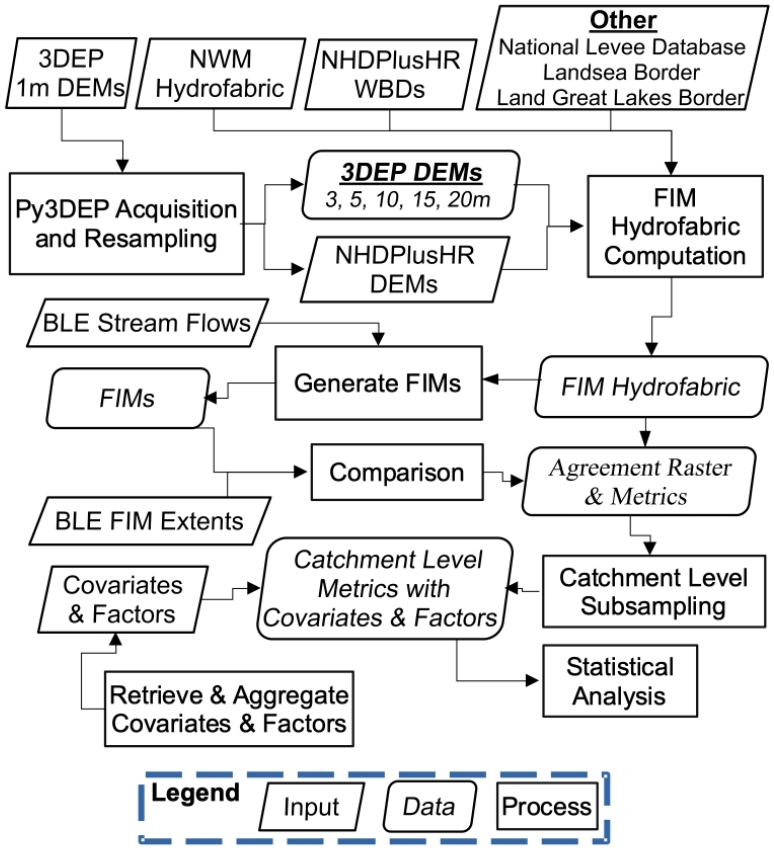

Figure 1. Figure illustrates the overall process for generating HAND and evaluating the use of LiDAR derived DEMs and their resolutions. Input datasets were collected from two different source DEMs, the 3DEP and NHDPlusHR. The 3DEP DEMs were resampled with Python 3-Dimensional Elevation Program (Py3DEP) to 3, 5, 10, 15, and 20 m spatial resolutions. Both source DEMs and their resolutions were used to compute the FIM hydrofabric which is comprised of various datasets used to produce FIM including HAND, catchments, and SRCs. BLE cross-sections were intersected with the NWM stream network to obtain stream flow estimates. These estimates were used to produce FIM using the SRCs coupled with HAND to produce estimates of the 100 and 500 yr extents. These extents were then compared to the extents from the BLE thus removing hydrology related errors that could be introduced if NWM streamflows were used. The agreement statistics were resampled to the NWM catchment level then referenced to a long series of catchment level covariates and factors that were used for statistical analysis and inference.

While Py3DEP is developed only for retrieving topographic data from a single source, PyGeoHydro can query various types of data from different sources, e.g., National Inventory of Dams (United States Army Corps of Engineers, 2023), Watershed Boundary Dataset (United States Geological Survey, 2023), and National Land Cover Database (Multi-Resolution Land

Characteristics Consortium, 2022). It also includes additional functionalities, including a look-up table for associating overland roughness to land cover type based on Liu et al. (2019). In this study, we obtained HUC geometries, reservoirs, LULC, imperviousness, and overland roughness data using PyGeoHydro.

2.4 DEM Preparation

DEMs underwent a curation procedure prior for use with HAND computation and our experimental design. To match the existing framework with the use of NHDPlusHR DEMs, Py3DEP was used to query the image server to acquire 3DEP elevations at a HUC-4 scale. To counter pixel limitations within the web service, queries were completed using overlapping tiles and mosaiced together using virtual rasters (VRTs) (Survey, 2022). To investigate the effect of varying spatial resolutions on FIM skill and computational performance, queries were elected taken at 3, 5, 10, 15, and 20 m spatial resolutions. As stated previously, the resolutions of 60 and 90 m were also used to help understand if and when spatial resolution begins to affect quality. Utilizing the "check_3dep_availability" tool, we determined that 1 m 3DEP information is available for the entire study region (see Section 2.5). Py3DEP queries a US Geological Survey dynamic web service for the best available DEM when generating its mosaics and resamples them given the user furnished resolution (Survey, 2022). Use of the Py3DEP function "query_3dep_source" confirmed availability of 1 m LiDAR data for the entire study area which means the functionality uses it for resampling purposes (Chegini et al., 2021; Stoker et al., 2015; Survey, 2022; USGS, 2022, 2021). The availability of the 1 m data as well as the other available source DEMs are illustrated in Figure 2. The selected resolutions of 3, 5, 10, 15, and 20 m were bounded by computational demands since further optimizations within the code and external dependencies should be considered prior to transitioning to 1 m elevation information for HAND computation.

2.5 Study Area

The site selection process considered several factors. The location of the site was limited by the availability of validation (discussed in Section 2.6) as well as the availability of 1 m 3DEP information (USGS, 2022, 2021). Additionally, the location of the site was influenced by OWPs plan to release FIM services in stages as a function of the percentage of the population served. The first release will serve 10% of the US population and cover portions of East Texas (TX) as well as the Mid-Atlantic states. On the other hand, the size of the evaluation site was constrained by the computational burdens of producing the FIM hydrofabric at multiple resolutions.

With these criteria in mind, we selected the Neches River sub-region as the study area for this experiment. The HUC-4 (1202) sub-region comprises 7 HUC-8 sub-basins, ranging continuously from 12020001 to 12020007. Located in South-East TX near the Louisiana border, the site stretches from Tyler to Beaumont, and includes the towns of Nacogdoches and Lufkin as depicted in Figure 3. Numerous braided streams and 15 reservoirs, including one of the largest ones, the Sam Rayburn Reservoir, populate the study area. Figure 4 depicts the spatial distribution of LULCs as defined in the 2019 NLCD but grouped to the top tier of categories for visibility and interpretability. The study area features a low slope, with low-lying areas mostly comprising four LULCs: evergreen forests (31.1%), pasture/hay (17.2%), woody wetlands (16.7%), and mixed forests

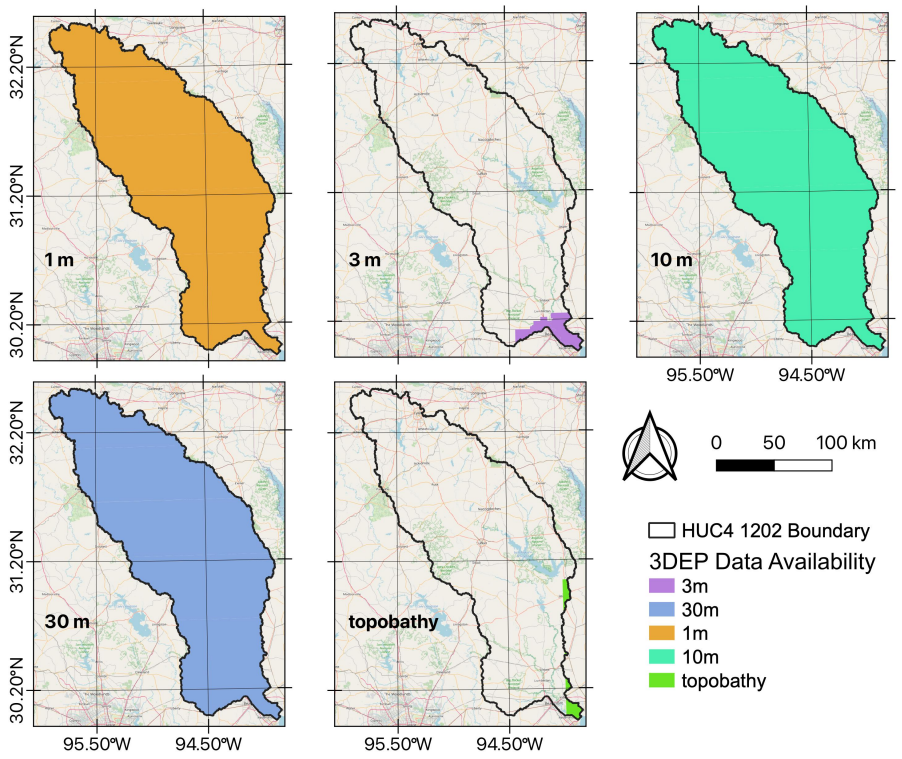

Figure 2. Illustrates the source DEMs available within 3DEP for the study area. High resolution 1 m information is available for the entire study area meaning it was used as the source resolution for resampling to the resolutions used for HAND computation including 3, 5, 10, 15, and 20 m. See Section 2.5 for more information within the study area. ©OpenStreetMap contributors 2023. Distributed under the Open Data Commons Open Database License (ODbL) v1.0

(11.4%). The developed LULCs together account for only 7.3% of the site's area. In summary, the study area has low terrain slope and minimal anthropogenic influence.

2.6 Evaluation

We chose the BLE FIM extents for evaluation, which are HEC-RAS 1D based models provided by InFRM and Federal Emergency Management Agency (FEMA) (FEMA, 2016, 2021a, b; Strategic Alliance for Risk Reduction II, 2019a, b, c, d, e, f, g). FEMA's Region 6 publishes these FIMs, which are available at both 1% (100 yr) and 0.2% (500 yr) flow magnitudes, and they also include cross-sectional information with the associated flows for each level. Despite being a modeled data set, HEC-RAS appears frequently in literature for comparison purposes, as it is an engineering scale model (Cook and Merwade, 2009; Rajib et al., 2016; Zheng et al., 2018a; Afshari et al., 2018; Wing et al., 2017; Criss and Nelson, 2022; Follum et al., 2017). We chose to intersect the cross-sections with NWM flowpaths to remove errors and uncertainties associated with hydrological and meteorological inputs used to produce streamflows within the NWM (Aristizabal et al., 2023c). This process enabled us

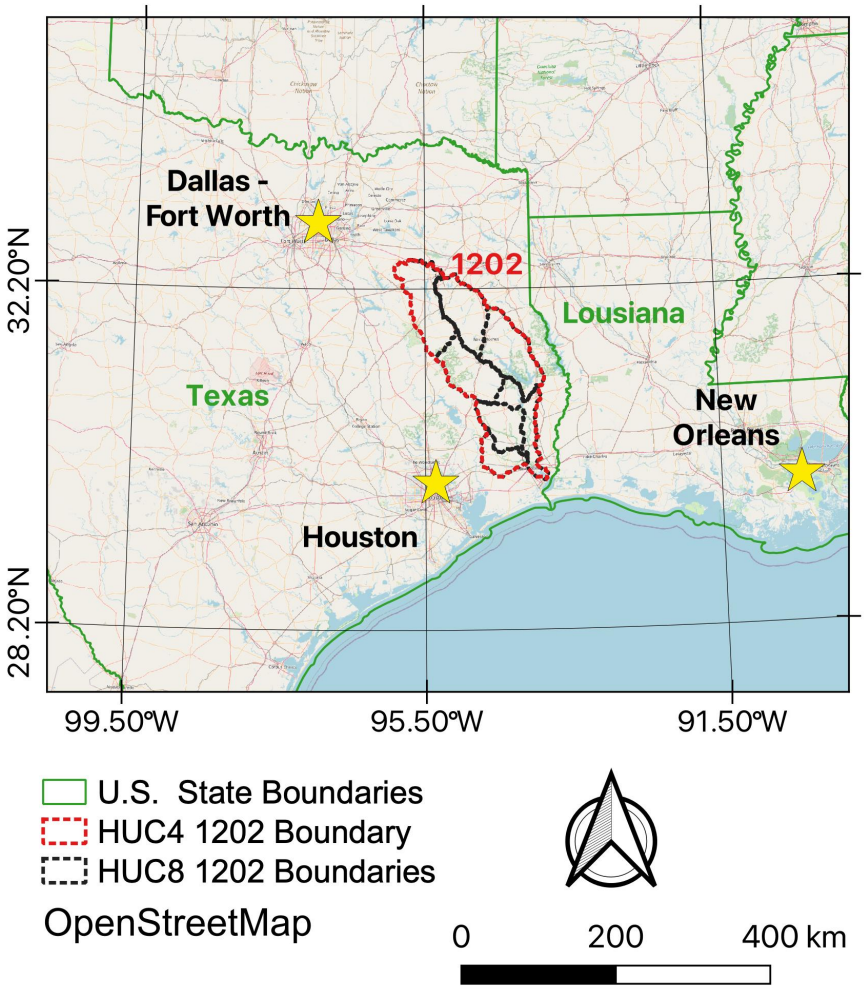

Figure 3. Overview of study area nestled in South East TX near the Louisiana border. Known as the Neches River sub-region or HUC-4 1202, the site is composed of 7 sub-basins or HUC-8s. ©OpenStreetMap contributors 2023. Distributed under the Open Data Commons Open Database License (ODbL) v1.0

to associate BLE derived 100 and 500 yr streamflow magnitudes with NWM forecasting points. If multiple intersections occurred per NWM stream reach, we took the median flow value. Even though this process may lead to conflation errors, we believe it allows for a better comparison with BLE FIM extents by removing any errors introduced from variances in other hydrological processes outside of inundation (Aristizabal et al., 2023c). For more detailed information on this technique and its application, see the evaluation methods in Aristizabal et al. (2023c). It is crucial to emphasize that the BLE benchmark FIMs utilize DEMs derived from high-quality LiDAR data, with a spatial resolution of approximately 1 m, for conducting hydraulic analyses and creating floodplain maps throughout our entire selected study region Strategic Alliance for Risk Reduction II

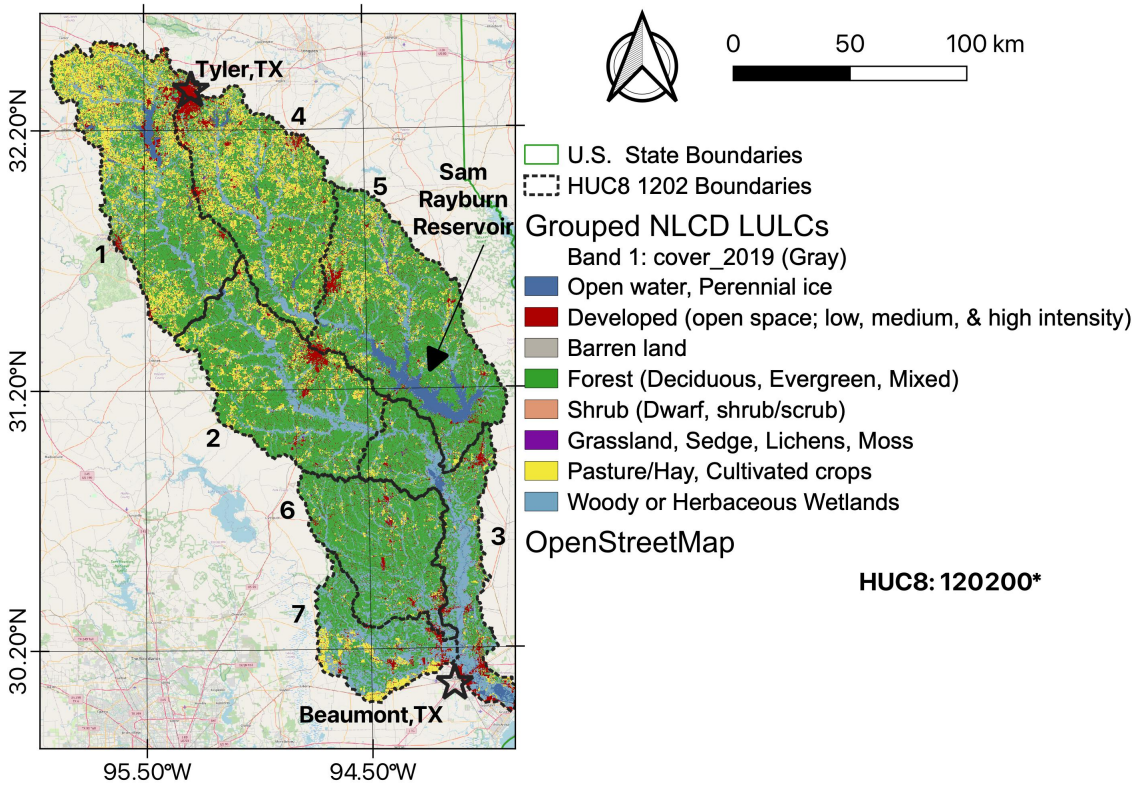

Figure 4. A detailed view showing the spatial distribution of the 2019 NLCD LULCs grouped to the top-tier categories for visibility and interpretability. About three-quarters of the site is made up of just four land covers including evergreen forest (31.1%), pasture/hay (17.2 %), woody wetlands (16.7 %), and mixed forests (11.4 %). Only about 7.2% of the site is considered developed. ©OpenStreetMap contributors 2023. Distributed under the Open Data Commons Open Database License (ODbL) v1.0

(2019a, b, c, d, e, f, g). The benchmark's dependence on LiDAR derived DEMs at 1 m resolutions enables the answering of our central question pertaining to the effect of DEM source and resolution on HAND based FIMs skill. We would like to acknowledge here that producing HAND with 1 m information to match that of the BLE is computationally very expensive, leading to substantial increases in central processing unit (CPU) time and memory usage, which we discuss in Sections 3 and 4.

In order to quantify agreement with the BLE FIM extents, we elected to apply binary contingency statistics. The primary metrics calculated within a contingency table include true positives (TPs), false positives (FPs), false negatives (FNs), and true negatives (TNs). We again note that the positive condition is considered inundated, while the negative condition is considered not inundated. In order to summarize the contingency table into secondary metrics, we employed the commonly used metrics within flood modeling including critical success index (CSI), true positive rate (TPR), and false alarm rate (FAR) shown in Equations 1, 2, and 3, respectively (Gerapetritis and Pelissier, 2004; Schaefer, 1990).

$$CSI = \frac{TP}{TP + FP + TN} \tag{1}$$

$$TPR = \frac{TP}{TP + FN} \tag{2}$$

$$FAR = \frac{FP}{TP + FP} \tag{3}$$

TPR, also known as sensitivity, recall, probability of detection or hit rate, was used to describe a models ability to detect flooding as it represents performance in regions that are considered flooded within the benchmark. It is formally described as the proportion of inundated pixels that are accurately detected as flooded. FAR, also known as false discovery rate, the inverse of precision, or the inverse of positive predictive value, conveys the opposite since it is used to represent over-prediction. This is formally described as the proportion of pixels incorrectly predicted as flooded with respect to the total number of pixels predicted as flooded. Work by Gerapetritis and Pelissier (2004) illustrated how these two metrics, TPR and FAR, are mathematically related to CSI where correctly predicted, non-inundated regions (TNs) are not considered. This leads to CSI being considered as inequitable or exhibiting frequency dependency, which could limit its use in comparing predicted datasets in scenarios with varying frequencies (Gerapetritis and Pelissier, 2004; Schaefer, 1990).

While these 3 widely adopted metrics are considered highly interpretable, we elected to include Matthew's Correlation Coefficient (MCC), shown in Equation 4, which is considered more equitable when dealing with cases of extreme class imbalance (Chicco and Jurman, 2020; Chicco et al., 2021a, b; Boughorbel et al., 2017). However, it does value both conditions (inundated and not inundated) to have equal impact (Chicco and Jurman, 2020; Chicco et al., 2021a, b; Boughorbel et al., 2017).

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{4}$$

2.6.1 Analysis

Evaluations for this HUC-4 study region were conducted at the HUC-8 scale which produces seven HUC-8 metric values across all five spatial resolutions evaluated as well as both flood magnitudes yielding about 70 samples to analyze ($7 \cdot 5 \cdot 2 = 70$). Analysis at this large HUC scale tends to erode away valuable information that could be used if a finer grain unit of measurement were used instead. Under this justification, we opted to sub-sample agreement maps down to the NWM catchment scale and recompute each of the four metrics for each catchment. There are 5,786 NWM catchments available for this study area which generates 405,020 effective samples to analyze ($70 \cdot 5,786 = 405,020$). This yielded a much finer grain spatial distribution of performance but also enabled the introduction of covariates and factors that can help explain some of the catchment to catchment variance in the metrics. Factors are categorical variables in our analysis that have a finite number of

distinct categories. Covariates, on the other hand, are continuous variables that are assumed to have a linear relationship with the dependent variable. The term covariate serves the same function as the factors with the only distinction being that covariates are of continuous data types. We investigated the interaction of explanatory variables by multiplying all possible combinations to capture the variance in the dependent variable more comprehensively. The combination of covariates and factors was carried out by including interaction terms in the regression model. Interaction terms are created by multiplying a covariate and a factor or two factors, which allows us to investigate whether the effect of one variable depends on the level of the other variable. Many of these covariates and factors stemmed from NWM catchments or flowpaths themselves including channel slope, catchment area, stream order, and reservoir.

The term reservoir here is used with respect to catchments that intersect NWM reservoirs. While NWM reservoirs are masked out for evaluation and also not modeled for within OWP FIM, the BLE FIM extents do model reservoir inundation. This creates regions of BLE inundation that extend beyond NWM reservoir definitions thus leading to FNs. NWM catchments that intersect with NWM reservoirs are denoted as reservoir catchments and used as a factor to help account for the performance within these regions. This is better illustrated in Figure 5.

In addition to catchment level attributes within the NWM hydrofabric, we collected a variety of datasets associated with hydrological processes including NLCD LULC, imperviousness, overland roughness, and terrain slope. These factors and covariates were obtained utilizing the HyRiver suite of tools described in Section 2.3. Overland roughness was determined by the NLCD LULCs and previous research assigned coefficients for each category (Dewitz, 2021; Yang et al., 2018; Chow, 1959; Chegini et al., 2021; Multi-Resolution Land Characteristics Consortium, 2022; McCuen et al., 2005; Kalyanapu et al., 2009). In order to aggregate to catchment scale, LULC was taken as the dominant category by catchment while the covariates imperviousness, overland roughness, and terrain slope were aggregated by taking the catchment level mean value. This procedure created a total of 10 catchment level covariates and factors summarized as spatial resolution, DEM source, channel slope, catchment area, stream order, reservoir, LULC, imperviousness, overland roughness, and terrain slope. These covariates and factors are collectively known as features, predictors, explanatory variables, or independent variables, and were used to correlate to dependent, response, or outcome variables which were the four metrics of interest in this study. These are described in more detail in Table 2.

Given the fact that we aggregated a variety of catchment scale features for each associated catchment scale metric, we used the regression analysis to help explain the magnitude and significance of the linear relationships between the explanatory variables and the four responses (metrics: MCC, CSI, TPR, and FAR) (Montgomery et al., 2021; Chatterjee and Simonoff, 2013; Merrill et al., 2017). We avoided including the metrics with the NHDPlusHR DEM in the regression analysis since it was already clear that using 3DEP DEMs led to significant skill improvements. To build our regression model, we opted to use forward model selection of all one-way and two-way interactions utilizing the Akaike Information Criterion (AIC) and terminating the model selection after a minimum is reached. Explanatory variables were feature scaled from 0 to 1 prior to fitting to better compare across explanatory variables. Meaning, this procedure added variables to the regression model for each metric first. Finding the explanatory variable that minimized AIC, it left that variable in the model then moved on the

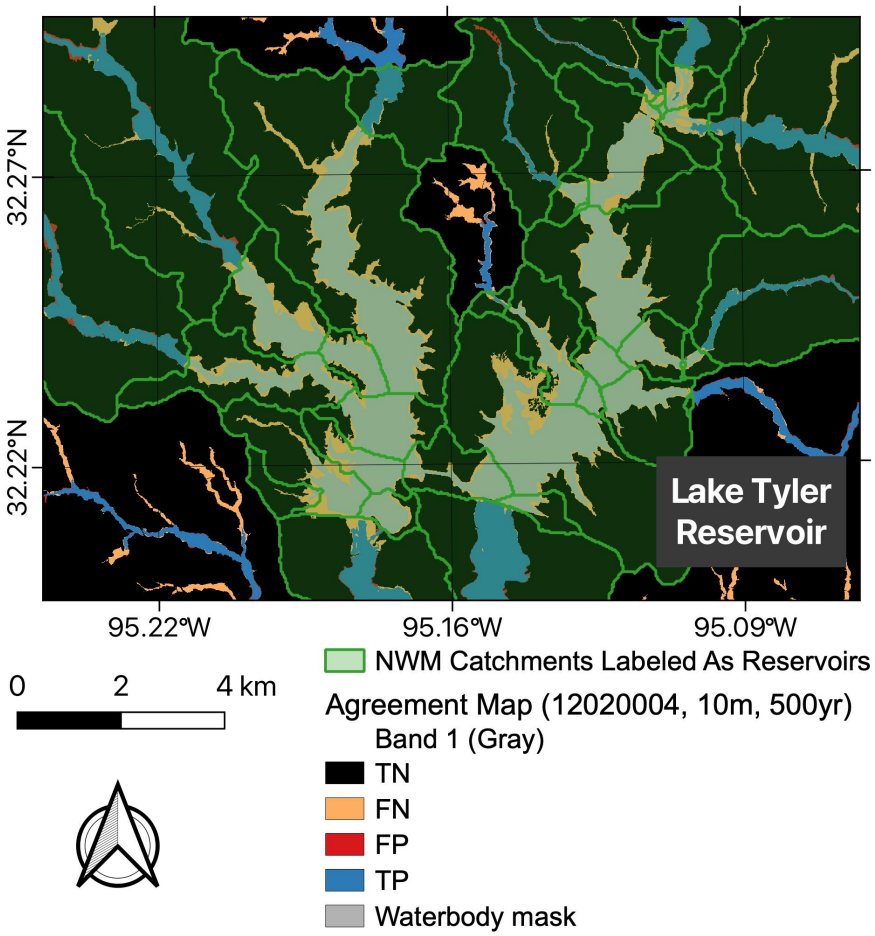

Figure 5. Figure shows Lake Tyler reservoir within HUC-8 12020004 of the study area. Background represents an agreement map between OWP FIM and BLE FIM at a 10 m spatial resolution for the 500 yr magnitude. Gray area shows masked out NWM reservoir since these are not being modeled within OWP FIM. The NWM catchments shaded in green represent catchments associated with this NWM reservoir as they are show to spatially intersect. These reservoir catchments were used in analysis to quantify the catchment variance in performance, partly, due to not accounting for reservoir inundation within BLE FIM.

remaining variables as long as the previous variable explained a minimum of 0.001 more than the previous model. This process helps build models with explanatory power while avoiding unnecessary complexity.

3 Results

Based on the observation of our results, we took an in-depth analysis of the effects of utilizing 3DEP DEMs first when compared to the legacy, NHDPlusHR DEMs. After confirming the positive effect of using 3DEP information, we varied the spatial resolution of these DEMs and observed the impact on performance. To further investigate the effects of additional explanatory

Table 2. Summary of catchment level covariates and factors used for statistical analysis. Includes the dataset, its statistical data type for analysis (factor or covariate), its units if it's a covariate or its levels if it's a factor.

Dataset	Statistical Data Type	Levels for factors or Units for Covariates)
Spatial Resolution	Factor	Five levels in meter units: 3, 5, 10, 15, & 20.
DEM source	Factor	Two levels: NHDPlusHR & 3DEP.
Channel Slope	Covariate	Vertical/horizontal as percentage.
Catchment Area	Covariate	Surface area in km^2.
Horton Strahler Stream Order	Factor	Six levels: 1 - 6
Reservoir	Factor	Two levels: True (1) and False (0).
NLCD 2019 Dominant LULC	Factor	Fifteen levels: Woody Wetlands, Pasture/Hay, Evergreen Forest, Developed, Low Intensity, Shrub-Forest, Mixed Forest, Developed, Open Space, Cultivated Crops, Developed, Medium Intensity, Emergent Herbaceous Wetlands, Water, Herbaceous-Forest, Developed, High Intensity, Deciduous Forest, and Grasslands/Herbaceous.
Imperviousness	Covariate	Percent of pixel area that is impervious surface.
Overland Roughness	Covariate	Unitless: Friction coefficient for overland water flow.
Terrain Slope	Covariate	Vertical/horizontal as percentage.
Stream Order	Covariate	Horton-Strahler stream order as defined in NWM stream network.

variables, we built a multiple linear model with forward model selection to help explain some of the catchment to catchment variance in the four metrics. Lastly, we decided to do an in-depth analysis on a few of these variables that we found of importance.

3.1 3DEP Data

For the given study area, we decided to investigate the effect on HAND based FIMs extents by utilizing the 3DEP data instead of the legacy source DEMs from the NHDPlusHR. We conducted this comparison on a NWM catchment scale in order to have a sense of the distribution of the results across some spatial definition finer than the HUC scale. Additionally, this comparison was conducted by resampling the 3DEP 1 m data to a spatial resolution of 10 m to match that of the legacy DEM. Figure 6 details the results of this comparison in a scatter plot format. Each individual data point represents a sample of the metrics taken at the NWM catchment scale. The points are sampled across two axes representing their performance with NHDPlusHR DEMs on the x-axis and 3DEP DEMs on the y-axis. The 45 degree, diagonal line represents a dividing line where the metric values for both DEMs are the same. Catchment samples symbolized in green represent enhanced FIM extents for that catchment for the given case while sample symbolized in red signify poorer quality extents. We also included descriptive statistics on each sub-figure representing the mean and standard deviation of the metric differences across DEMs (3DEP - NHDPlusHR) as well as the percentage of differences greater than or less than zero.

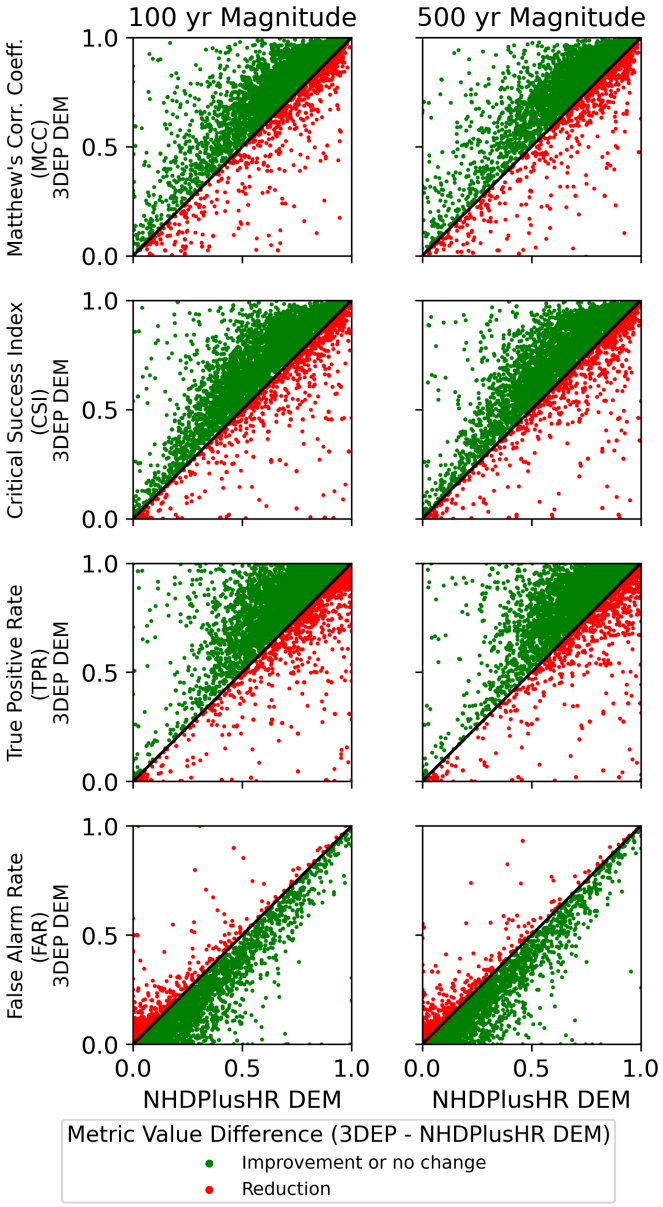

Figure 6. Figure shows catchment scale metric values. The eight sub-figures are organized by magnitude (100 and 500 yr) across the columns and for the four metrics across the rows. These values within each sub-figure are plotted on an axis representing HAND based FIMs generated from the NHDPlusHR DEMs (x-axis) and the same FIMs generated from 3DEP DEMs resampled to the 10 m spatial resolution (y-axis). The diagonal 45 degree line divides catchments that perform better with the legacy DEM (in red) from the catchments that perform better with the 3DEP DEM (in green). The majority of catchments perform better across all four metrics and both magnitudes with the higher quality 3DEP information. Additional descriptive statistics quantifying the distribution of metric differences (3DEP - NHDPlusHR) are also presented including the mean and standard deviation of the differences. We also included the percentage of samples whose difference is greater than or less than zero depending on the metric referenced.

Overall, the use of the higher quality, more recently produced 3DEP DEMs generally enhances FIM extents across all the metrics and magnitudes examined. This is evident by observing the high proportion of catchments represented in green as well as the high percentage of samples greater than zero for the first three metrics. The FAR is minimized so a lower proportion of samples above zero is considered better. Overall, approximately four in every five catchments in considered to benefit from the use of 3DEP when compared to the use of NHDPlusHR. This approximate relationship holds true across metrics and event magnitudes for our given experimental design.

3.2 Regression Analysis

After we established the effect of the new elevation data source on FIM extents, we elected to conduct regression analysis on the remaining explanatory variables of interest. As explained in the methods, we regressed on the four metrics of interest independently and fit the model in a forward selection fashion utilizing AIC as a measure of model fit. Figure 7 represents the resulting models from that forward model selection in graphical form. The four subplots represent the results of the model fit to each metric or response variable. The y-axis labels represent explanatory variables starting with the intercept followed by the remaining variables and their two-way interactions in the order of selection as per the AIC metric. By "two-way interactions," we refer to the statistical interaction between pairs of explanatory variables, implying that the effect of one explanatory variable on the response variable may change depending on the value of another explanatory variable. The points on the graph represent the values of the coefficients while the shape represents the level of significance from $>= 0.05$ (circle), < 0.05 (pentagon), < 0.01 (triangle), and < 0.001 (star). The green and red colors represent the nature of the effects as either positive (direct) or negative (indirect), respectively. Since AIC lacks interpretability, we elected to show the coefficient of determination or R^2 at each step of the forward selection process. Additionally, Table 3 presents the results shown in Figure 7 in a table format Jann (2005).

Further examining Figure 7, we can infer interesting pieces of information as regression analysis is a tool to synthesize data into something interpretable. The coefficient of determination or R^2 values across the metrics vary from about 0.21 to 0.33. Translating this into other terms, one can say that about one fifth to one third of the catchment to catchment variance in the metrics can be explained by the eleven catchment scale, explanatory variables and their two-way interactions selected in this study. Additional, observation from the Figure illustrates the prevalence of certain explanatory variables near the beginning of the selection process that seem to explain a fair amount of the variation as well as exhibiting strong effect sizes. Some of the variables of note include reservoir, stream order, terrain slope, channel slope, and LULC. These explanatory variables and there effect on catchment level performance on FIM will be examined later on in Section 3.4.

3.3 3DEP DEM Spatial Resolution

We investigated the effect of varying the spatial resolution of the 3DEP DEMs on the quality of FIMs produced from HAND. The 3DEP DEMs were varied from 3, 5, 10, 15, and 20 m prior to HAND computation.

Figure 8 examines the relationship of DEM spatial resolution at five levels for each of the four metrics selected. The relationships are illustrated as distributions of catchment scale metric values for both event magnitudes (100 and 500 yr). We

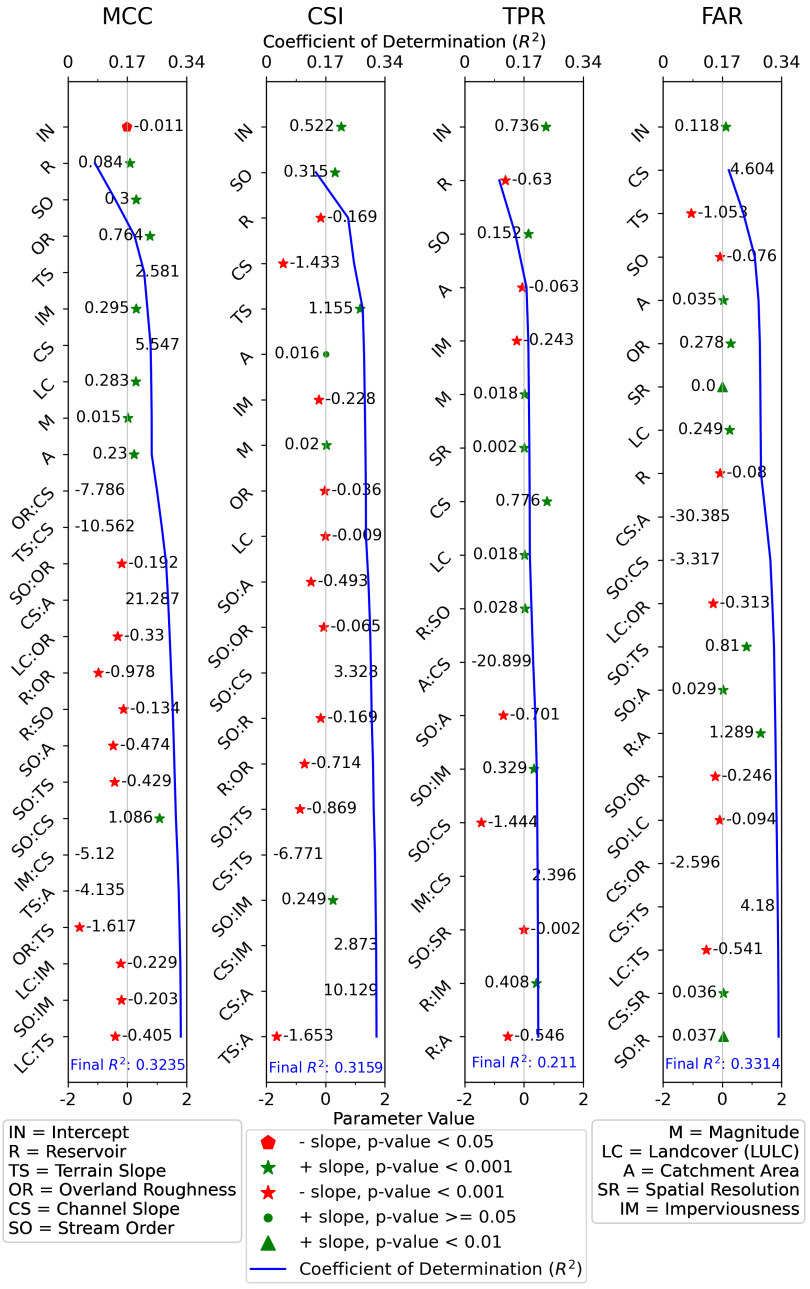

Figure 7. Figure illustrates the coefficients from multiple linear regression models fit on four response variables independently. The points on the graph represent the values of the coefficients while the shape represents the level of significance from $>= 0.05$ (circle), < 0.05 (pentagon), < 0.01 (triangle), and < 0.001 (star). The green and red colors represent the nature of the effects as either positive (direct) or negative (indirect), respectively. The models were built in a step wise fashion using forward model selection and AIC as a criteria for terminating the process. Eleven explanatory variables were considered for these models as well as their two-way interactions. An intercept was also included by default. As the models built, we recorded the R^2 of each successive model and tracked as the complexity of model increased. The final R^2 values for each final model are reported as well for each step of the forward selection.

Table 3. Regression analysis table showing coefficient values and their level of significance for each agreement metric. Coefficient values are based on feature scaled explanatory variables in the range of 0 to 1. Variable definitions: Intercept (IN), Reservoir (R), Terrain Slope (TS), Overland Roughness (OR), Channel Slope (CS), Stream Order (SO), Magnitude (M), Landcover (LULC) (LC), Catchment Area (A), Spatial Resolution (SR), and Imperviousness (IM).

MCC		**CSI**		**TPR**		**FAR**	
Variable	**Coefficient**	**Variable**	**Coefficient**	**Variable**	**Coefficient**	**Variable**	**Coefficient**
IN	-0.011*	IN	0.522***	IN	0.736	IN	0.118***
R	0.084***	SO	0.315***	R	-0.630	CS	4.604***
SO	0.300***	R	-0.169***	SO	0.152	TS	-1.053***
OR	0.764***	CS	-1.433***	A	-0.063	SO	-0.076***
TS	2.581***	TS	1.155***	IM	-0.243	A	0.035***
IM	0.295***	A	0.016	M	0.018	OR	0.278***
CS	5.547***	IM	-0.228***	SR	0.002	SR	0.000**
LC	0.283***	M	0.020***	CS	0.776	LC	0.249***
M	0.015***	OR	-0.036***	LC	0.018	R	-0.08***
A	0.230***	LC	-0.009***	R:SO	0.028	CS:A	-30.385***
OR:CS	-7.786***	SO:A	-0.493***	A:CS	-20.899	SO:CS	-3.317***
TS:CS	-10.562***	SO:OR	-0.065***	SO:A	-0.701	LC:OR	-0.313***
SO:OR	-0.192***	SO:CS	3.328***	SO:IM	0.329	SO:TS	0.81***
CS:A	21.287***	SO:R	-0.169***	SO:CS	-1.444	SO:A	0.029***
LC:OR	0.330***	R:OR	-0.714***	IM:CS	2.396	R:A	1.289***
R:OR	-0.978***	SO:TS	-0.869***	SO:SR	-0.002	SO:OR	-0.246***
R:SO	-0.134***	CS:TS	-6.771***	R:IM	0.408	SO:LC	-0.094***
SO:A	-0.474***	SO:IM	0.249***	R:A	-0.546	CS:OR	-2.596***
SO:TS	-0.429***	CS:IM	2.873***			CS:TS	4.180***
SO:CS	1.086***	CS:A	10.129***			LC:TS	-0.541***
IM:CS	-5.120***	TS:A	-1.653***			CS:SR	0.036***
TS:A	-4.135***					SO:R	0.037**
OR:TS	-1.617***						
LC:IM	-0.229***						
SO:IM	-0.203***						
LC:TS	-0.405***						
R^2	0.3235		0.3159		0.2110		0.3314

$^{***}p <= 0.001$, $^{**}p <= 0.01$, $^{*}p <= 0.05$

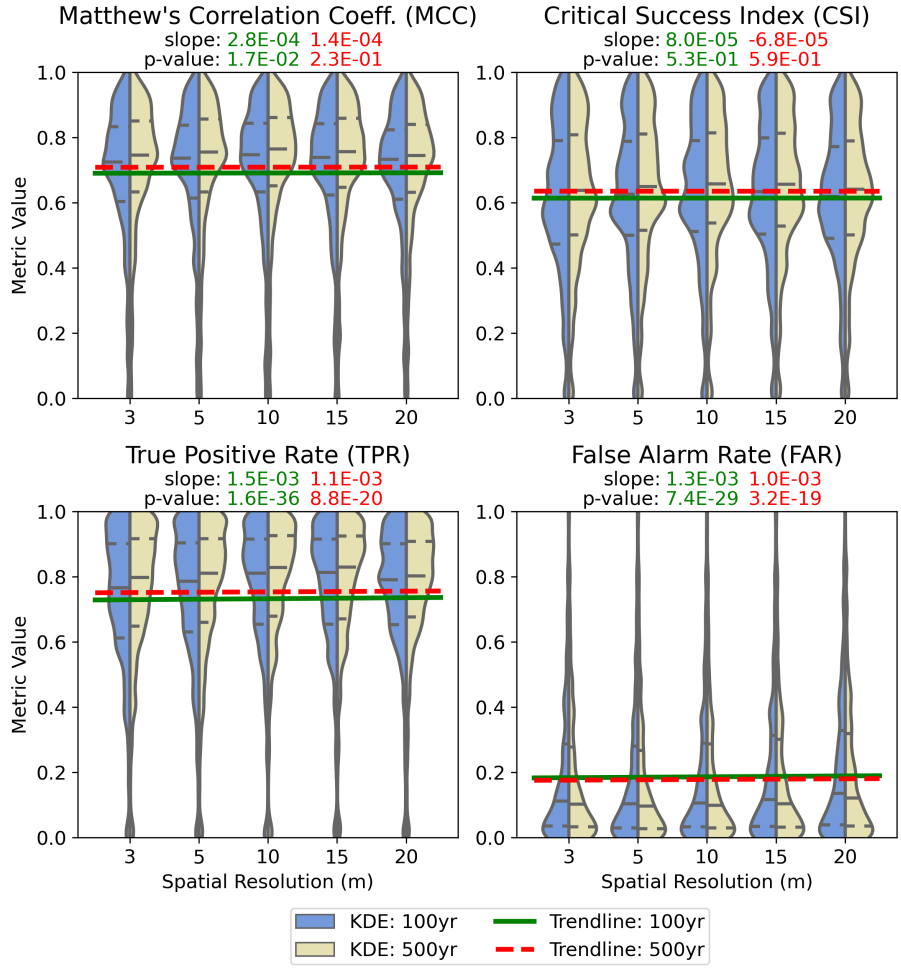

Figure 8. Illustrates the distribution of the four catchment scale metrics as violin plots across every spatial resolution selected including 3, 5, 10, 15, and 20 m. Each half of the violin represents a given magnitude of events (100 and 500 yr). Linear trendlines are fit for each metric-magnitude combination establishing linear relationships between spatial resolutions and metric values at the catchment scale.

computed the distributions as Gaussian kernel density estimations (KDEs) which is a non-parametric statistical technique that determines the probability distribution of a random variable (Virtanen et al., 2020; Scott, 2015; Silverman, 2018; Turlach et al., 1993; Bashtannyk and Hyndman, 2001). For each metric-magnitude distribution of catchment scale metrics, the 75th, 50th, and 25th percentiles are calculated and displayed from top to bottom as dotted, solid, and dotted lines, respectively. Additionally, we fit two linear regression lines, one for each magnitude, for all four metrics relating the linear effects of spatial resolution on metric values. The effect sizes, or the slopes of the regression lines, are displayed as well as their respective p-values. Low p-values denote effect sizes that are unlikely equal to zero.

Table 4. Mean and standard deviations of inundated areas across HUC-8's and magnitudes (100 and 500 yr in km^2 for each spatial resolution in meters.

Spatial Resolution (m)	Mean Inundated Area (km^2)	Standard Deviation of Inundated Area (km^2)
3	653.46	161.58
5	650.58	161.12
10	652.85	153.46
15	654.10	158.22
20	659.15	154.58

Examination of Figure 8 shows statistically significant yet marginal in value effect sizes for the TPR and FAR metrics. For example, the effect size of the TPR and 100 yr case is 0.0015 which represents a 0.0015 increase in the value of TPR for every unit m increase in the magnitude of the resolution. So for approximately 10 m, one would expect TPR to increase by 0.015. While coarser resolution DEMs appear to improve detection of inundation when compared to the BLE FIMs, it also appears have an undesirable effect on FAR as its expected values increase as DEMs are coarsened. These competing effects on TPR and FAR seem to have a canceling effect on the overall performance metrics of MCC and CSI. Both MCC and CSI both have statistically insignificant trendlines which hints little to no overall improvement in catchment scale metrics of HAND based FIM by varying the spatial resolution of the input DEMs used to produce HAND.

Furthermore, we analyzed the mean and standard deviations of the inundated areas for the five spatial resolutions selected. Table 4 shows the HUC-8 level mean and standard deviation of inundated areas in km^2 by spatial resolution and across magnitudes. Very little variation in the inundated areas was seen across the resolutions which hints that while there was an increase in TPR and FAR with coarser DEMs, there is also little change in the inundated areas. This suggests that most of the trade-offs in resolution were related to trading type I errors (FPs) in certain areas with type II errors (FNs) in other areas with little to no overall change in the inundated areas.

A final observation related to the spatial resolution relates to its relative low importance or a lack of interaction variables in the models built for the regression analysis in Section 3.2 and Figure 7. This denotes that spatial resolution provided little to no effect when considering impactful variables such as LULC, imperviousness, stream order, or reservoir.

DEM resolution was found to have a significant effect on the computational demands of producing HAND. We aggregated the times to compute HAND at the seven various DEM spatial resolutions and found a significant effect on CPU time especially at finer resolutions. Figure 9 shows the change in log CPU times in seconds by HUC-8. An almost entire order of magnitude change in CPU time is seconds is observed when using DEMs of 3 versus 20 m resolutions. The number of pixels for a given domain for squared pixels is known to have an inverse relationship with the square of spatial resolution ($numberOfPixels \propto resolution^{-2}$). So reducing spatial resolution from 10 m to 1 m represents a 100x increase in the number of pixels for the fixed domain. It is important to note that all computational benchmarks were computed on an Amazon Web Services' 't3.2xlarge' instance with 8 processing units, 32 gigabytes of memory, and a 2000 gigabyte solid state, elastic block storage unit. The

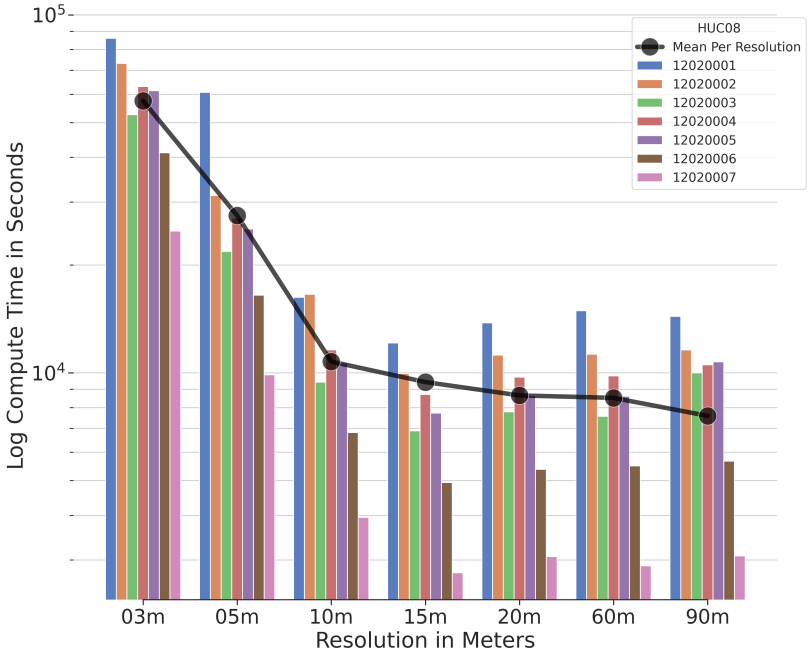

Figure 9. Total log CPU time in seconds across varying DEM spatial resolutions of 3, 5, 10, 15, 20, 60, and 90 m and listed by the seven HUC-8s in the study region. Resolution was found to have a significant effect on total CPU time for computing HAND as nearly an entire order of magnitude reduction in seconds was observed from changing the DEM resolution from 3 to 20 m. All computational benchmarks were computed on an Amazon Web Services' 't3.2xlarge' instance with 8 processing units, 32 gigabytes of memory, and a 2000 gigabyte solid state, elastic block storage unit. The operating system was based on GNU/Linux with an Ubuntu 22.04 distribution on an x86, 64-bit architecture.

operating system was based on GNU/Linux with an Ubuntu 22.04 distribution on an x86, 64-bit architecture. Despite having a minimal observed effect on skill, we found that higher resolutions tended to exhibit an excessive computational cost.

While this initial analysis demonstrates little influence of DEM spatial resolution on FIM extent quality, it also begged the question when might spatial resolution begin to exhibit some significant effect? In order to answer this question, the spatial resolutions of 60 and 90 m were selected and used to produce HAND and subsequent FIMs. These extents were again compared at the catchment scale for the entire study region for both 100 and 500 yr flow magnitudes. Since 10 m is the current standard resolution for elevations within the NHDPlusHR and seamless 3DEP datasets, we illustrate the mean catchment scale metrics across the entire study region for both flow magnitudes in Table 5. These results illustrate how spatial resolution does eventually exhibit a strong negative effect on FIM performance as the resolution is coarsened beyond the standard 10 m which specifically was found to be around the medium resolutions between 20 and 60 m.

Table 5. Mean catchment scale agreement metrics for National Hydrography Dataset (NHD) (10 m resolution) and 3DEP (3, 5, 10, 15, 20, 60, and 90 m resolutions) DEMs across the entire study domain for both 100 and 500 yr flow magnitudes. Values are presented as percentages (%) instead of native decimal values for readability.

	NHD	3DEP					
Agreement Metric	10m	3m	5m	10m	15m	20m	60m	90m
MCC	63.02	70.63	70.69	70.81	70.75	70.67	52.02	61.33
CSI	54.58	62.94	62.99	63.18	63.12	63.01	43.93	52.01
TPR	67.35	76.16	76.18	76.36	76.43	76.38	65.52	72.71
FAR	23.35	17.94	18.23	18.70	18.78	18.98	40.74	32.39

3.4 Explanatory Variable Focus

Since reservoirs and LULCs are valuable for forecasting operations, we elected to focus on those explanatory variables further within this analysis. Other variables, while important, were left out of scope for further analysis for this paper.

3.4.1 Reservoirs

Given the relative importance of reservoirs in explaining catchment to catchment variance in the many of the metrics as shown in Figure 7 and Section 3.2, we isolate this factor out for further analysis here. Figure 10 shows the catchment level distribution of the four metrics across spatial resolutions as violin plots built with KDEs. The halves of the violins are split across catchments that intersect NWM reservoirs and those that don't. The trendlines as well as their displayed slopes and p-values represent catchment scale metric variance as a function of spatial resolution for each reservoir group.

This Figure primarily shows a large statistical difference in catchment scale variation of three metrics, MCC, CSI, and TPR, across the catchments that intersect reservoirs and those that don't. Explaining this variation is simple as OWP FIM does not currently account for reservoir related inundation while the BLE does. While the NWM reservoirs are currently masked out for evaluation purposes, the BLE reservoir inundation extents go beyond these masked regions thus contributing to FNs. Due to this fact, FAR illustrates very little performance difference across reservoir groups as FAR considers FPs and omits FNs (see Equation 3. Another important trend to denote from the previous Figure is the relative lack of interaction between spatial resolution and the reservoir factor shown by the similarity of the slopes of trendlines across reservoir groups. This can be interpreted as spatial resolution having little effect across the reservoir groups which can also be seen in Figure 7 where the selection of a reservoir - spatial resolution predictor was omitted. Until OWP FIM accounts for reservoir flooding or some higher order masking technique is applied, the presences of reservoir related catchments will continue to contribute to a high variance in catchment scale metrics.

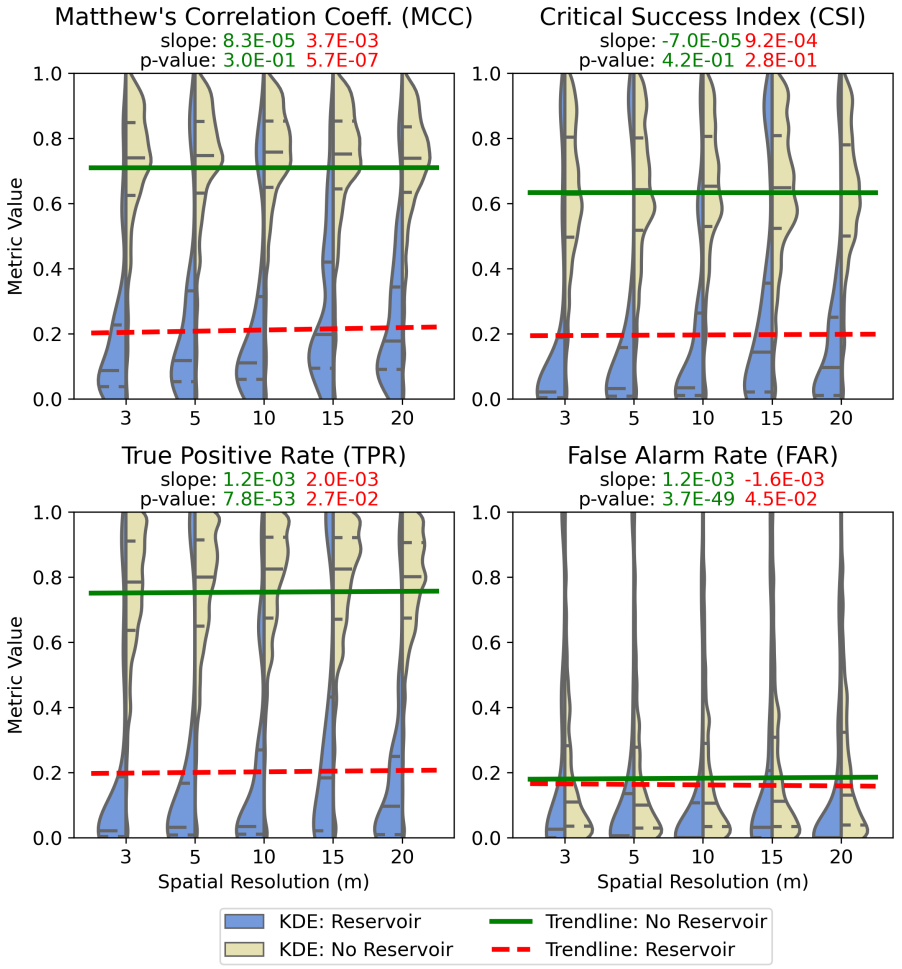

Figure 10. Catchment scale variation illustrated as distributions modeled as KDEs. The distributions are grouped by metric and by spatial resolutions. The halves of the violins are divided by the presesence of a NWM catchment that intersects a NWM reservoir or not. Significant differences are observed between catchments identified with reservoirs and those that are not for all resolutions and metrics employed. Reservoirs are not currently modeled within OWP FIM while the BLE does account for reservoir related inundation extents. While the NWM reservoirs are masked out for evaluation purposes, some of the BLE inundation extents reach beyond these boundaries leading to significant amount of FNs. The trendlines as well as their corresponding slopes and p-values were constructed by regressing the two reservoir groups independently on spatial resolution. Little to now interaction between reservoir groups and spatial resolutions was observed.

3.4.2 Landuse/Landcover

We analyzed catchment scale metrics by taking the dominant landcover per catchment (mode). While the linear analysis in Section 3.2 grouped the NLCD categories into two groups depending on their degree of anthropogenic influence, we decided to ungroup the categories for Figure 11. In this Figure, we illustrate the distribution of the four catchment scale metrics shown as box plots which are grouped both by NLCD LULC and event magnitude. This chart doesn't appear to have a clear trend until further inspection leads one to see a pattern pertaining to the catchment scale agreement and the nature of the LULCs. To reveal this trend, we decided to group LULC categories according to their relative level of anthropogenic influence.

In grouping the LULCs by two categories of "more" and "less" anthropogenic influence we are able to see a clearer trend as to how LULC affects catchment scale agreement. The LULCs grouped into the "more" category include the developed categories (open space, low intensity, medium intensity, and high intensity) and the cultivated crops category which depending on the cropping system can have significant hydrological implications. The remaining LULCs within the study area were placed in the "less" category. Figure 12 shows the distribution of catchment scale metrics sorted by grouped LULC and event magnitude. We fit a multiple linear regression model for each metric using the grouped LULC and magnitude as factors as well as their interaction. The resulting formulas for this linear modeling are shown above each Figure with the parameter values and their relative level of significance. Since only p-values greater than 0.05 and less than 0.001 were encountered, we denoted those with no asterisk and 3 astericks, respectively. Additionally, we plot the trendlines resulting from another regression that associates the metric values to the LULC grouping and does this for each event magnitude independently. Illustrating this regression demonstrates these relationships in a qualitative manner highlighting the lack of interaction of event magnitude and grouped LULC.

Judging from the Figures 11 and 12, there is a clear indication that LULC has a significant influence over catchment scale agreement. Grouped LULCs in Figure 12, show the importance of anthropogenic influence on explaining catchment scale variation in metric values with a negative relationship observed for having "more" relative anthropogenic influence. We found that LULC affected all the metrics except for FAR where over-prediction was found not be as affected by the anthropogenic influence. Under-prediction does appear prevalent in regions of anthropogenic influence which could be explained by a variety of factors including DEM inconsistencies or adverse affects on hydro-conditioning in areas with rapidly varying or uncertain elevations. It does appear that anthropogenic influence also contributes to more variation within the "more" case than the "less" case which could be a result of noise that is inherited from elevation inputs. While the magnitude per se is a significant factor in explaining catchment scale agreement, it does not interact with grouped LULC meaning anthropogenic influence seems to carry a similar effect across event magnitudes. Another interesting observation related to LULC is that the grouped LULCs don't seem to interact with spatial resolution. So for this study area, higher resolutions did not provide an improvement in metrics for regions with more anthropogenic influence. We leave further analysis of the effect of LULC and the anthropogenic influence on catchment scale agreement to future work.

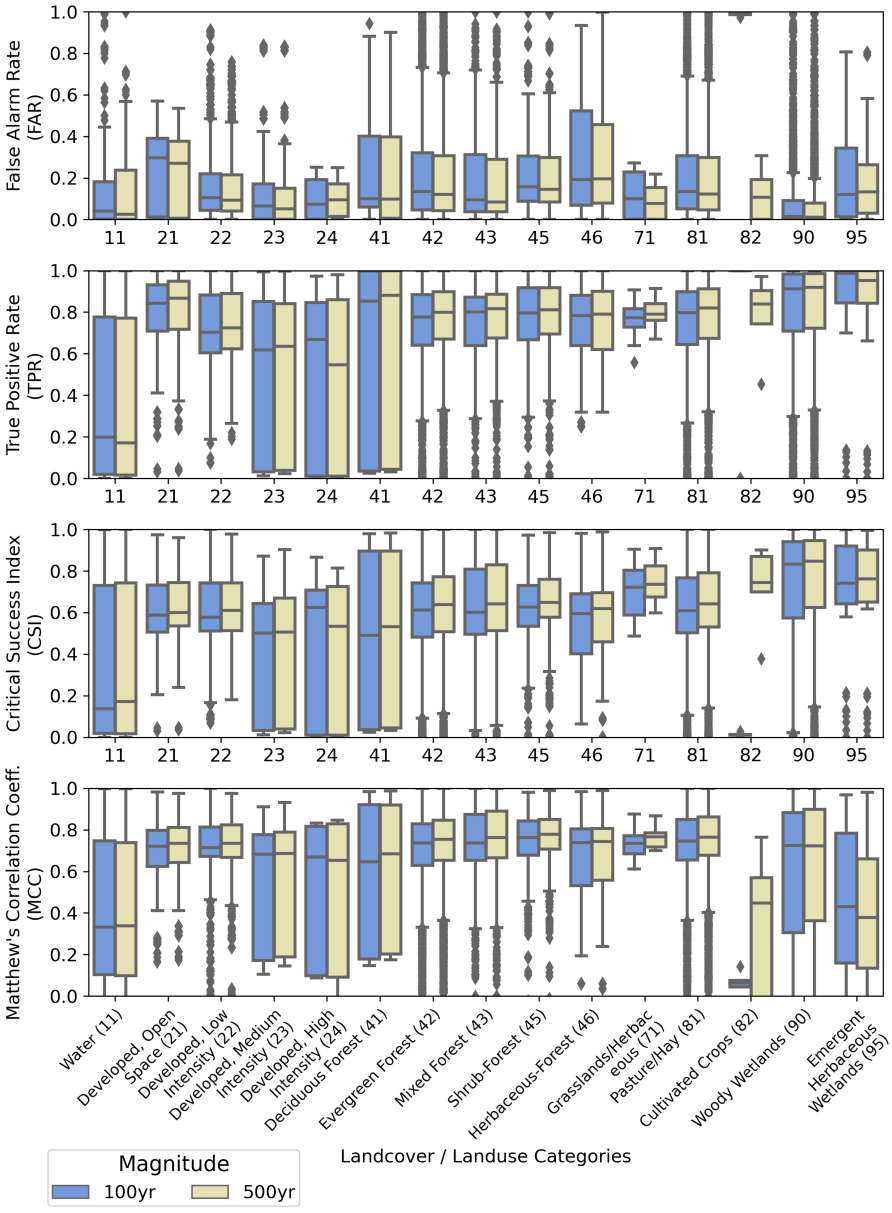

Figure 11. The distribution of four catchment scale agreement metrics are shown as box plots and grouped by the dominant NLCD LULC per catchment as well as the event magnitude.

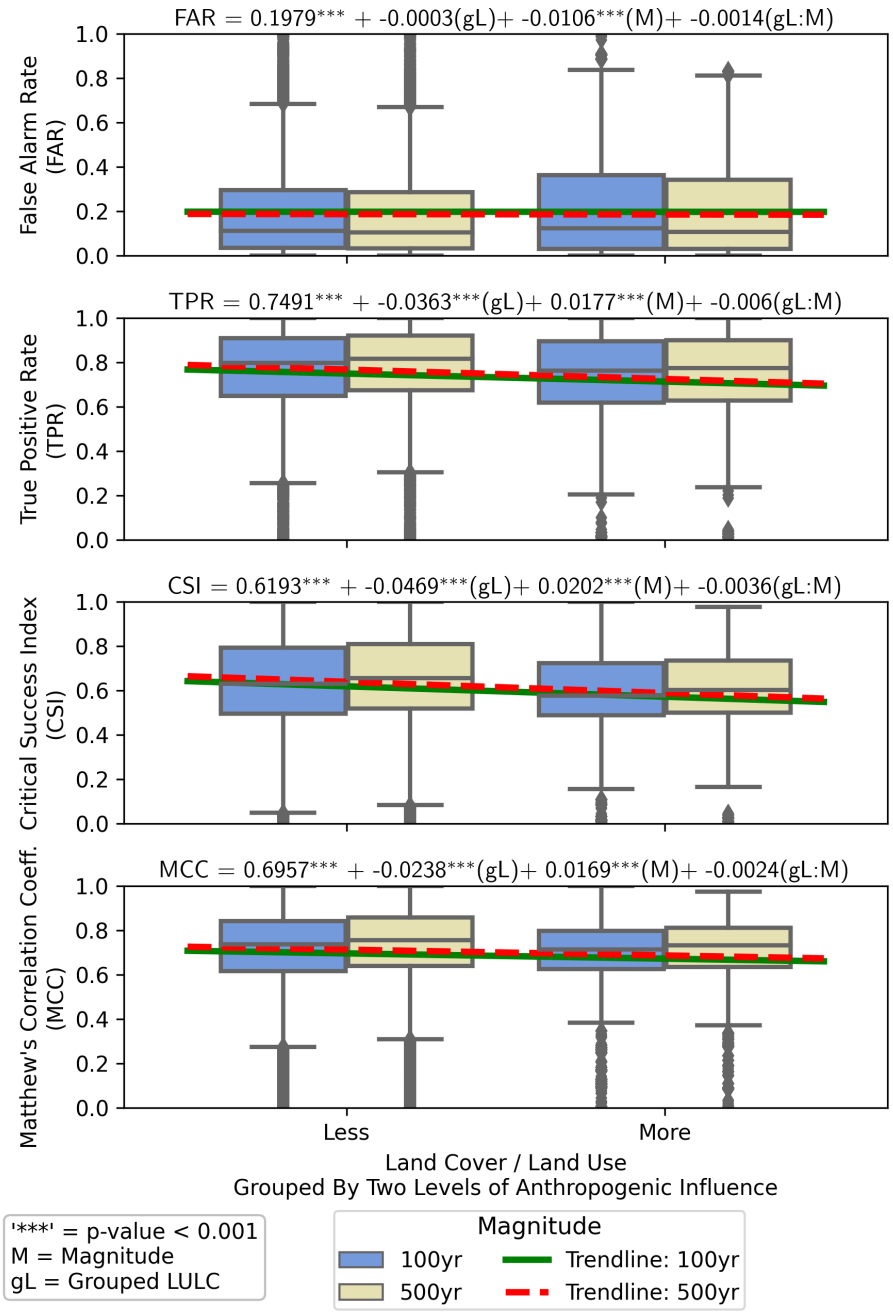

Figure 12. Distribution of catchment scale agreement across four metrics illustrated as box plots and grouped by the level of anthropogenic influence in the NLCD LULCs and the two event magnitudes. The results of a linear model that regress the catchment scale metrics on grouped LULC, magnitude, and their interaction are shown. The coefficients for the model are labeled with their p-values by no asterisk and three astericks for the p-values greater than 0.05 and less than 0.001, respectively. Additionally the trendlines resulting in a regression with catchment scale metrics on grouped LULC are shown per event magnitude. It is important to note that LULC was found not to interact with spatial resolution within our regression analysis, meaning various LULCs perform similarly across varying spatial resolutions.

4 Discussion

Our results and analysis demonstrated several key methods that can improve the agreement of continental scale FIM using HAND when compared to engineering scale FIM models. The inclusion of higher quality terrain information from 3DEP was able to significantly improve the quality of continental scale HAND based FIMs. This finding is consistent with previous studies that have come to similar conclusions (Li et al., 2022b; Zheng et al., 2018a; Garousi-Nejad et al., 2019; Speckhann et al., 2018), while also meeting the goal of the 3DEP objectives set out in justifying the collection effort (Dewberry, 2011; Snyder et al., 2013; Sugarbaker et al., 2014). As the program approaches continental scale availability, 3DEP data can be justified for use directly for HAND computation for the entire US leading to enhanced FIM and forecast quality.

However, varying the spatial resolution of the 3DEP DEMs from their native 1 m was found to have little effect on the quality of HAND based FIMs at least within 20 m resolution. Additional analysis extending comparisons of HAND based FIMs at coarser resolutions, 60 and 90 m, revealed a significant degradation of performance at these scales. Previous studies examining this (Li et al., 2022b; Zheng et al., 2018a; Garousi-Nejad et al., 2019; Speckhann et al., 2018) have varied in their experimental design and modeling assumptions. The modeling assumptions in those studies were different than those used in our HAND methods as we employ different datasets, hydro-conditioning procedures, and compute HAND at finer scales (Aristizabal et al., 2023c). The experimental design of some previous studies (Zheng et al., 2018a; Garousi-Nejad et al., 2019) looked at high resolution terrain data but did not explicitly isolate that factor out for analysis purposes. Additionally, previous studies failed to denote a consistent relationship with spatial resolution and FIM performance (Li et al., 2022b; Speckhann et al., 2018). While the mechanisms of this relationship have not been thoroughly explored with HAND, others have found that spatial resolution may have a spurious relationship with FIM performance due to inherent uncertainties related to this problem (Savage et al., 2016). Future research can expand the analysis of spatial resolution's effect on FIM quality to more study sites across broader domains of interest. Future research can also explore the effect spatial resolution may have on the quality of FIM depths as they likely behave somewhat independently to extents. Additionally, we'd like to motivate alternative benchmark datasets that could further our understanding of how DEM source and resolution affect the quality of HAND based FIMs Afshari et al. (2018). Utilizing 2D hydraulic models along with high-resolution and -quality DEMs could furnish additional insights into this relationship as higher-order physics would produce varying extents compared to that of our 1D HEC-RAS benchmark Afshari et al. (2018).

Further analysis sought to explain some of the catchment level variation in the four agreement metrics with the aim of indicating where future progress can be made in extending FIM quality for continental scale applications. As a result of the regression analysis, a fifth to a third of the catchment scale variation of the agreement metric values was explained by building linear models with eleven explanatory variables and their two-way interactions. These models, while used for analysis purposes in this study, can have predictive performance which could have calibration applications. Previous works have used a variety of methods to help calibrate Manning's n or bathymetry (Zheng et al., 2018b; Johnson et al., 2019; Jian et al., 2017; Neal et al., 2021; Liu et al., 2019). Due to the complex and interconnected nature of the source hydrography, hydro-conditioning

565 operations, reach-averaged channel geometry, and Manning's n, any sort of calibration for SRCs would involve using the same sets of methods and datasets used to produce our version of HAND based FIM.

Some of the explanatory variables were explored in further detail to provide insights on future possible skill improvements for OWP HAND based FIM. Reservoirs were found as one of the leading independent variables in explaining catchment scale variation in three of the four metrics mostly driven by under-prediction or FNs. These errors are caused by not accounting
570 for reservoir inundation within OWP HAND FIM. Several methods exist for accounting for reservoir inundation which could leverage volume computations with the NWM (Gochis et al., 2021; Chen et al., 2018; Shin et al., 2019). It's important to note here that while reservoirs explained a significant amount of variation in the metrics, it does not mean that accounting for reservoir inundation properly would lead to a significant increase in agreement. Agreement will only change in response to the quality of the new method employed as well as the prevalence of reservoir inundation in a given region. Another variable
575 of interest further analyzed included the effect of LULC on agreement metrics. We found that HAND FIMs did not perform as well in catchments that are labeled as developed or cultivated crops. Regions of high anthropogenic influence negatively influence the performance of inundation models by adding extra complexity within the terrain information as well as with the physics employed. Furthering performance in these regions could benefit from use of hyper-resolution models that better account for urban water features (Grimley et al., 2017; Smith et al., 2020; Deo et al., 2018; Gurung et al., 2018; Smith et al.,
580 2021; Leandro et al., 2016; Chegini et al., 2021). Further exploring these and other independent variables could help inform future development directions to help improve the quality of continental scale FIM techniques.

5 Conclusions

Floods are a significant source of natural disasters in the United States (US) leading to loss of property and lives. The National Oceanic and Atmospheric Administration (NOAA)'s Office of Water Prediction (OWP) has implemented a National Water
585 Model (NWM) to help forecast streamflows at nearly three million locations across the continuous US, Hawaii, Puerto Rico, and portions of Alaska at hourly time steps and multiple forecast horizons. The OWP has developed its own version of Height Above Nearest Drainage (HAND) that accounts for multiple fluvial sources of inundation instead of just that from the local, nearest flowpath. The United States Geological Survey (US Geological Survey)'s 3-Dimensional Elevation Program (3DEP) is rapidly approaching continental scale availability so we evaluated its use at the 10 meter (m) spatial resolution within the
590 derivation of HAND and found significant increases in the quality of flood inundation map (FIM) performance. Additionally, we varied the resolution to include 3, 5, 15, and 20 m but did not find any significant trends on an overall basis leading one to justify its use within this study regions. However, a significant degradation in FIM skill was observed at the coarser resolutions of 60 and 90 m. As one would expect, the computational time increased to compute HAND with the number of digital elevation model (DEM) cells considered, which goes up with the inverse of the DEM resolution squared. We determined
595 that more studies are required in other regions to help explore the potential benefits with the use of higher resolution DEMs along with HAND. A multiple linear regression model fitting eleven factors and covariates to the four agreement metrics all at the catchment scale revealed that about one fifth to one third of the variation can be explained by these explanatory variables.

Code and data availability. HAND data produced for this study can be found on our Earth Science Information Partners (ESIP) backed Amazon Web Services S3 bucket (Aristizabal et al., 2023a). Software used in this study is available on GitHub (Aristizabal et al., 2023b). A permanent version of this code and data is available as well. (Aristizabal, 2023).

Author contributions. Conceptualization, FA, GP, and FS; methodology, FA, FS, GP, and JJ; software, FA, GP, and TC; validation, FA and GP; formal analysis, FA, GP, and FS; investigation, FA, GP, and FS; resources, FS; data curation, FA, GP, and TC; writing—original draft preparation, FA and TC; writing—review and editing, FA, FS, GP, TC, and JJ; visualization, FA; supervision, FS and JJ; project administration, FS and JJ; funding acquisition, FS. All authors have read and agreed to the published version of the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We would like to thank Jason Stoker and his team at the 3DEP for producing analysis friendly services and for answering our questions related to these services. Additionally, we would like to thank Fred Ogden, chief scientist of the Office of Water Prediction (OWP), for his help reviewing and providing scientific direction. We would like to acknowledge the use of large language models (LLMs) in the editing process of this manuscript to enhance the clarity, coherence, and grammatical accuracy of our writing.

References

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

Alajo, S. O., Nakavuma, J., and Erume, J.: Cholera in endemic districts in Uganda during El Niño rains: 2002–2003, African health sciences, 6, 93–97, 2006.

Archuleta, C.-A. M., Constance, E. W., Arundel, S. T., Lowe, A. J., Mantey, K. S., and Phillips, L. A.: The National Map seamless digital elevation model specifications, Tech. rep., US Geological Survey, 2017.

Aristizabal, F.: Data and Software for Effects of High-Quality Elevation Data and Explanatory Variables on the Accuracy of Flood Inundation Mapping via Height Above Nearest Drainage, https://doi.org/https://doi.org/10.4211/hs.3d98a9e5a6d84020b72800fd27c87f9a, 2023.

Aristizabal, F. and Judge, J.: 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, IEEE, 2021.

Aristizabal, F., Judge, J., and Monsivais-Huertero, A.: High-Resolution Inundation Mapping for Heterogeneous Land Covers with Synthetic Aperture Radar and Terrain Data, Remote Sensing, 12, 900, 2020.

Aristizabal, F., Bates, B., Avant, B., Chadwick, N., Grout, T., Spies, R., Luck, M., Salas, F., Pruitt, C., Hanna, R., and Cocks, G.: noaa-nws-owp-fim, s3://noaa-nws-owp-fim/, 2023a.

Aristizabal, F., Bates, B., Avant, B., Chadwick, N., Grout, T., Spies, R., Luck, M., Salas, F., Pruitt, C., Hanna, R., and Cocks, G.: NOAA-OWP/inundation-mapping, https://github.com/NOAA-OWP/inundation-mapping, 2023b.

Aristizabal, F., Salas, F., Petrochenkov, G., Grout, T., Avant, B., Bates, B., Spies, R., Chadwick, N., Wills, Z., and Judge, J.: Extending Height Above Nearest Drainage to Model Multiple Fluvial Sources in Flood Inundation Mapping Applications for the US National Water Model, Water Resources Research, p. e2022WR032039, 2023c.

Arundel, S., Phillips, L., Lowe, A., Bobinmyer, J., Mantey, K., Dunn, C., Constance, E., and Usery, E.: Preparing The National Map for the 3D Elevation Program–products, process and research, Cartography and Geographic Information Science, 42, 40–53, 2015a.

Arundel, S., Bulen, A., Adkins, K., Brown, R., Lowe, A., Mantey, K., and Phillips, L.: Assimilation of the National Elevation Dataset and launch of the 3D elevation program through the USGS spatial data infrastructure, International Journal of Cartography, 4, 129–150, 2018.

Arundel, S. T., Archuleta, C.-A. M., Phillips, L. A., Roche, B. L., and Constance, E. W.: 1-meter digital elevation model specification, chap. 7, https://doi.org/http://dx.doi.org/10.3133/tm11B7, 2015b.

Bales, J. and Wagner, C.: Sources of uncertainty in flood inundation maps, Journal of Flood Risk Management, 2, 139–147, 2009.

Bashtannyk, D. M. and Hyndman, R. J.: Bandwidth selection for kernel conditional density estimation, Computational Statistics & Data Analysis, 36, 279–298, 2001.

Bates, P., Marks, K., and Horritt, M.: Optimal use of high-resolution topographic data in flood inundation models, Hydrological processes, 17, 537–557, 2003.

Bates, P. D. and De Roo, A.: A simple raster-based model for flood inundation simulation, Journal of hydrology, 236, 54–77, 2000.

Baudoin, M.-A., Henly-Shepard, S., Fernando, N., and Sitati, A.: Early warning systems and livelihood resilience: Exploring opportunities for community participation, Tech. rep., 2014.

Beinin, L.: Medical consequences of natural disasters, Springer Science & Business Media, 2012.

Berz, G.: Flood disasters: lessons from the past—worries for the future, in: Proceedings of the institution of civil engineers-water and maritime engineering, vol. 142, pp. 3–8, Thomas Telford Ltd, 2000.

Birkmann, J., Buckle, P., Jaeger, J., Pelling, M., Setiadi, N., Garschagen, M., Fernando, N., and Kropp, J.: Extreme events and disasters: a window of opportunity for change? Analysis of organizational, institutional and political changes, formal and informal responses after mega-disasters, Natural hazards, 55, 637–655, 2010.

Boughorbel, S., Jarray, F., and El-Anbari, M.: Optimal classifier for imbalanced data using Matthews Correlation Coefficient metric, PloS one, 12, e0177 678, 2017.

Callahan, D. and Berber, M. M.: Vertical accuracy of the USGS 3DEP program data: study cases in Fresno County and in Davis, California, Boletim de Ciências Geodésicas, 28, 2022.

Carruthers, A.: The assessment, improvement, and application of the GeoFlood flood inundation mapping framework, Ph.D. thesis, 2021.

Charlesworth, S. M. and Warwick, F.: 15 Adapting to and Mitigating Floods Using Sustainable Urban Drainage Systems, Flood Hazards: Impacts and Responses for the Built Environment, p. 207, 2011.

Chatterjee, S. and Simonoff, J. S.: Handbook of regression analysis, John Wiley & Sons, 2013.

Chegini, T., Li, H.-Y., and Leung, L. R.: HyRiver: Hydroclimate Data Retriever, Journal of Open Source Software, 6, 1–3, https://doi.org/10.21105/joss.03175, 2021.

Chen, W., Nover, D., He, B., Yuan, H., Ding, K., Yang, J., and Chen, S.: Analyzing inundation extent in small reservoirs: A combined use of topography, bathymetry and a 3D dam model, Measurement, 118, 202–213, 2018.

Chicco, D. and Jurman, G.: The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation, BMC genomics, 21, 1–13, 2020.

Chicco, D., Starovoitov, V., and Jurman, G.: The benefits of the Matthews correlation coefficient (MCC) over the diagnostic odds ratio (DOR) in binary classification assessment, Ieee Access, 9, 47 112–47 124, 2021a.

Chicco, D., Warrens, M. J., and Jurman, G.: The Matthews correlation coefficient (MCC) is more informative than Cohen's Kappa and Brier score in binary classification assessment, IEEE Access, 9, 78 368–78 381, 2021b.

Chow, V. T.: Development of uniform flow and its formulas, Open-channel hydraulics, McGraw-Hill Book Company, edited by: Harmer, DE, USA, pp. 89–114, 1959.

Cohen, S., Praskievicz, S., and Maidment, D. R.: Featured collection introduction: National water model, 2018.

Cook, A. and Merwade, V.: Effect of topographic data, geometric configuration and modeling approach on flood inundation mapping, Journal of hydrology, 377, 131–142, 2009.

Cools, J., Innocenti, D., and O'Brien, S.: Lessons from flood early warning systems, Environmental science & policy, 58, 117–122, 2016.

Corringham, T. W. and Cayan, D. R.: The effect of El Niño on flood damages in the western United States, Weather, Climate, and Society, 11, 489–504, 2019.

Cosgrove, B., Gochis, D., Graziano, T. M., Clark, E. P., and Flowers, T.: The Evolution of NOAA's National Water Model: An Overview of Version 2.1 and Future Operational Plans, AGUFM, 2019, H51D–01, 2019.

Council, N. R. et al.: Elevation data for floodplain mapping, National Academies Press, 2007.

Council, N. R. et al.: Mapping the zone: Improving flood map accuracy, National Academies Press, 2009.

Criss, R. E. and Nelson, D. L.: Stage-based flood inundation mapping, Natural Hazards, pp. 1–17, 2022.

D'Angelo, C., Passalacqua, P., Fiori, A., and Volpi, E.: Identification of flood-prone areas with GeoFlood: Lessons learned from the Tiber River case study, Journal of Flood Risk Management, 15, e12 795, 2022.

de Almeida, G. A., Bates, P., and Ozdemir, H.: Modelling urban floods at submetre resolution: challenges or opportunities for flood risk management?, Journal of Flood Risk Management, 11, S855–S865, 2018.

Deo, I., Modi, P., Zarekarizi, M., and Valle, J.: Sensitivity of urban flooding to presence of subsurface storm drainage systems in hydrologic models for low-gradient watersheds, National Water Center Innovators Program Summer Institute Report, 53, 2018.

Dewberry: Final Report of the National Enhanced Elevation Assessment, 2011.

Dewberry: 3D Nation Elevation Requirements and Benefits Study - Final Report, Tech. rep., Dewberry, 8401 Arlington Boulevard, Fairfax, VA, 22031-4666, https://www.dewberry.com/docs/default-source/documents/3d-nation-elevation-requirements-and-benefits-study/3d_nation_study_final_report.pdf, 2022.

Dewitz, J, U. G. S.: National Land Cover Database (NLCD) 2019 Products (ver. 2.0, June 2021): U.S. Geological Survey data release, https://doi.org/https://doi.org/10.5066/P9KZCM54, 2021.

Diehl, R. M., Gourevitch, J. D., Drago, S., and Wemple, B. C.: Improving flood hazard datasets using a low-complexity, probabilistic floodplain mapping approach, PloS one, 16, e0248 683, 2021.

Dixon, B. and Earls, J.: Resample or not?! Effects of resolution of DEMs in watershed modeling, Hydrological Processes: An International Journal, 23, 1714–1724, 2009.

Dobbs, K. E.: Evaluation of the Usgs National Elevation Dataset and the Kansas Biological Survey's FLDPLN (" Floodplain") Model for Inundation Extent Estimation, Ph.D. thesis, University of Kansas, 2010.

Doocy, S., Daniels, A., Murray, S., and Kirsch, T. D.: The human impact of floods: a historical review of events 1980-2009 and systematic literature review, PLoS currents, 5, 2013.

Downton, M. W., Miller, J. Z. B., and Pielke Jr, R. A.: Reanalysis of US National Weather Service flood loss database, Natural Hazards Review, 6, 13–22, 2005.

Drury, A. C. and Olson, R. S.: Disasters and political unrest: An empirical investigation, Journal of Contingencies and Crisis Management, 6, 153–161, 1998.

FEMA: Guidance for Flood Risk Analysis and Mapping: Automated Engineering, Guidance Document 27, Federal Emergency Management Agency (FEMA), 2016.

FEMA: Base Level Engineering (BLE) Tools and Resources, https://www.fema.gov/media-collection/base-level-engineering-ble-tools-and-resources, 2021a.

FEMA: estBFE Viewer, https://webapps.usgs.gov/infrm/estBFE/, 2021b.

Fewtrell, T., Bates, P. D., Horritt, M., and Hunter, N.: Evaluating the effect of scale in flood inundation modelling in urban environments, Hydrological Processes: An International Journal, 22, 5107–5118, 2008.

Follum, M. L., Tavakoly, A. A., Niemann, J. D., and Snow, A. D.: 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, 280–299, 2017.

French, J., Ing, R., Von Allmen, S., and Wood, R.: Mortality from flash floods: a review of national weather service reports, 1969-81., Public Health Reports, 98, 584, 1983.

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

Gerapetritis, H. and Pelissier, J. M.: On the behavior of the critical success index, 2004.

Gesch, D., Oimoen, M., Greenlee, S., Nelson, C., Steuck, M., and Tyler, D.: The National Elevation Dataset: Photogrammetric engineering and remote sensing, Photogrammetric engineering and remote sensing, 68, 5–32, 2002.

Gesch, D., Evans, G., Mauck, J., Hutchinson, J., Carswell Jr, W. J., et al.: The national map—Elevation, US geological survey fact sheet, 3053, 2009.

Gesch, D. B. and Maune, D.: Digital elevation model technologies and applications: the DEM users manual, The national elevation dataset, 2nd edn. American Society for Photogrammetry and Remote Sensing, Bethesda, pp. 99–118, 2007.

Gesch, D. B., Oimoen, M. J., Evans, G. A., et al.: Accuracy assessment of the US Geological Survey National Elevation Dataset, and comparison with other large-area elevation datasets: SRTM and ASTER, vol. 1008, US Department of the Interior, US Geological Survey Sioux Falls, SD, USA, 2014.

Gochis, D., Dugger, A., Barlage, M., Cabell, R., FitzGerald, K., McAllister, M., McCreight, J., RafieeiNasab, A., Read, L., Sampson, K., Yates, D., and Zhang, Y.: The WRF-Hydro modeling system technical description,(Version 5.2), Tech. rep., NCAR, 2021.

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

Golnaraghi, M.: An Overview: Building a global knowledge base of lessons learned from good practices in multi-hazard early warning systems, Institutional partnerships in multi-hazard early warning systems, pp. 1–8, 2012.

Gourevitch, J. D., Kousky, C., Liao, Y. P., Nolte, C., Pollack, A. B., Porter, J. R., and Weill, J. A.: Unpriced climate risk and the potential consequences of overvaluation in US housing markets, Climate Change, https://doi.org/https://doi.org/10.1038/s41558-023-01594-8, 2023.

Grimley, L. E., Tijerina, D., Khanam, M., Tiernan, E. D., Frazier, N., Ogden, F. L., Steinke, R. C., Maxwell, R. M., and Cohen, S.: Grid vs Mesh: The case of Hyper-resolution Modeling in Urban Landscapes, in: AGU Fall Meeting Abstracts, vol. 2017, pp. H53F–1541, 2017.

Gurung, D., Goenner, A., Perez, F., and Rouf, T.: Effects of Spatial Resolution on a Distributed Hydrologic Model through Dynamical Forcings: Flood Extent and Depth in Low Gradient Watersheds, National Water Center Innovators Program Summer Institute Report 2018, p. 40, 2018.

Heidemann, H. K.: Lidar base specification, chap. 4, https://doi.org/https://doi.org/10.3133/tm11b4, 2018.

Hocini, N., Payrastre, O., Bourgin, F., Gaume, E., Davy, P., Lague, D., Poinsignon, L., and Pons, F.: Performance of automated methods for flash flood inundation mapping: a comparison of a digital terrain model (DTM) filling and two hydrodynamic methods, Hydrology and Earth System Sciences, 25, 2979–2995, 2021.

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

Jafarzadegan, K. and Merwade, V.: A DEM-based approach for large-scale floodplain mapping in ungauged watersheds, Journal of Hydrology, 550, 650–662, 2017.

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

Jann, B.: Making regression tables from stored estimates, The Stata Journal, 5, 288–308, 2005.

Jian, J., Ryu, D., Costelloe, J. F., and Su, C.-H.: Towards hydrological model calibration using river level measurements, Journal of Hydrology: Regional Studies, 10, 95–109, 2017.

Johnson, J. M., Munasinghe, D., Eyelade, D., and Cohen, S.: An integrated evaluation of the National Water Model (NWM)–Height Above Nearest Drainage (HAND) flood mapping methodology, Natural Hazards and Earth System Sciences, 19, 2405–2420, 2019.

Jonkman, S. N.: Global perspectives on loss of human life caused by floods, Natural hazards, 34, 151–175, 2005.

Kahn, M. E.: The death toll from natural disasters: the role of income, geography, and institutions, Review of economics and statistics, 87, 271–284, 2005.

Kalyanapu, A. J., Burian, S. J., and McPherson, T. N.: Effect of land use-based surface roughness on hydrologic model output., Journal of Spatial Hydrology, 9, 2009.

Kelmelis, J. A., DeMulder, M. L., Ogrosky, C. E., Van Driel, N. J., and Ryan, B. J.: The National Map from geography to mapping and back again, Photogrammetric Engineering & Remote Sensing, 69, 1109–1118, 2003.

Kim, M., Stoker, J., Irwin, J., Danielson, J., and Park, S.: Absolute Accuracy Assessment of Lidar Point Cloud Using Amorphous Objects, Remote Sensing, 14, https://doi.org/10.3390/rs14194767, 2022.

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

Leandro, J., Schumann, A., and Pfister, A.: A step towards considering the spatial heterogeneity of urban key features in urban hydrology flood modelling, Journal of Hydrology, 535, 356–365, 2016.

Li, Z. and Demir, I.: U-Net-based Semantic Classification for Flood Extent Extraction using SAR Imagery and GEE Platform: A Case Study for 2019 Central US Flooding, 2022.

Li, Z., Mount, J., and Demir, I.: Evaluation of Model Parameters of HAND Model for Real-Time Flood Inundation Mapping: Iowa Case Study, 2020.

Li, Z., Duque, F. Q., Grout, T., Bates, B., and Demir, I.: Comparative Analysis of Performance and Mechanisms of Flood Inundation Map Generation using Height Above Nearest Drainage, 2022a.

Li, Z., Mount, J., and Demir, I.: Accounting for uncertainty in real-time flood inundation mapping using HAND model: Iowa case study, Natural Hazards, 112, 977–1004, 2022b.

Liu, C., Guo, L., Ye, L., Zhang, S., Zhao, Y., and Song, T.: A review of advances in China's flash flood early-warning system, Natural hazards, 92, 619–634, 2018a.

Liu, Y., Tarboton, D. G., and Maidment, D. R.: 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 . . . , 2020.

Liu, Y. Y., Maidment, D. R., Tarboton, D. G., Zheng, X., Yildirim, A., Sazib, N. S., and Wang, S.: A CyberGIS approach to generating high-resolution height above nearest drainage (HAND) raster for national flood mapping, The Third International Conference on CyberGIS and Geospatial Data Science, Urbana, IL, https://doi.org/https://doi.org/10.13140/RG.2.2.24234.41925/1., 2016.

Liu, Y. Y., Maidment, D. R., Tarboton, D. G., Zheng, X., and Wang, S.: A CyberGIS integration and computation framework for high-resolution continental-scale flood inundation mapping, JAWRA Journal of the American Water Resources Association, 54, 770–784, 2018b.

Liu, Z., Merwade, V., and Jafarzadegan, K.: Investigating the role of model structure and surface roughness in generating flood inundation extents using one-and two-dimensional hydraulic models, Journal of Flood Risk Management, 12, e12 347, 2019.

Maidment, D. R.: Conceptual Framework for the National Flood Interoperability Experiment, JAWRA Journal of the American Water Resources Association, 53, 245–257, https://doi.org/10.1111/1752-1688.12474, 2017.

Mallakpour, I. and Villarini, G.: The changing nature of flooding across the central United States, Nature Climate Change, 5, 250–254, 2015.

Mason, D. C., Horritt, M. S., Hunter, N. M., and Bates, P. D.: Use of fused airborne scanning laser altimetry and digital map data for urban flood modelling, Hydrological Processes: An International Journal, 21, 1436–1447, 2007.

Maune, D. F. and Nayegandhi, A.: Digital elevation model technologies and applications: The DEM users manual, American Society for Photogrammetry and Remote Sensing, 2018.

McCuen, R. H. et al.: Hydrologic analysis and design, vol. 3, Pearson prentice hall Upper Saddle River, NJ, 2005.

McGehee, R., Li, L., and Poston, E.: The Modified HAND Method, in: National Water Center Innovators Program Summer Institute Report 2016, edited by Maidment, D. R., Rajib, A., Lin, P., and Clark, E. P., vol. 4, 2016.

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

Merrill, H. R., Grunwald, S., and Bliznyuk, N.: Semiparametric regression models for spatial prediction and uncertainty quantification of soil attributes, Stochastic Environmental Research and Risk Assessment, 31, 2691–2703, 2017.

Merwade, V., Olivera, F., Arabi, M., and Edleman, S.: Uncertainty in flood inundation mapping: current issues and future directions, Journal of Hydrologic Engineering, 13, 608–620, 2008.

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

Montgomery, D. C., Peck, E. A., and Vining, G. G.: Introduction to linear regression analysis, John Wiley & Sons, 2021.

Moore, R. B., McKay, L. D., Rea, A. H., Bondelid, T. R., Price, C. V., Dewald, T. G., Johnston, C. M., et al.: User's guide for the National Hydrography Dataset plus (NHDPlus) High Resolution., Open-File Report-US Geological Survey, 2019.

Muhadi, N. A., Abdullah, A. F., Bejo, S. K., Mahadi, M. R., and Mijic, A.: The use of LiDAR-derived DEM in flood applications: A review, Remote Sensing, 12, 2308, 2020.

Multi-Resolution Land Characteristics Consortium: https://www.mrlc.gov/data-services-page, 2022.

Muthusamy, M., Casado, M. R., Butler, D., and Leinster, P.: Understanding the effects of Digital Elevation Model resolution in urban fluvial flood modelling, Journal of hydrology, 596, 126 088, 2021.

Neal, J., Schumann, G., Fewtrell, T., Budimir, M., Bates, P., and Mason, D.: Evaluating a new LISFLOOD-FP formulation with data from the summer 2007 floods in Tewkesbury, UK, Journal of Flood Risk Management, 4, 88–95, 2011.

Neal, J., Hawker, L., Savage, J., Durand, M., Bates, P., and Sampson, C.: Estimating river channel bathymetry in large scale flood inundation models, Water Resources Research, 57, e2020WR028 301, 2021.

Nobre, A., Cuartas, L., Hodnett, M., Rennó, C., Rodrigues, G., Silveira, A., Waterloo, M., and Saleska, S.: Height Above the Nearest Drainage – a hydrologically relevant new terrain model, Journal of Hydrology, 404, 13 – 29, https://doi.org/https://doi.org/10.1016/j.jhydrol.2011.03.051, 2011.

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

Oceanic, N. N. and Administration), A.: National Water Model: Improving NOAA's Water Prediction Services, 2016.

of Engineers, U. A. C.: National Levee Database, https://levees.sec.usace.army.mil/, 2021.

of Water Prediction, O.: The National Water Model, https://water.noaa.gov/about/nwm, 2022.

Office of Management and Budget: Coordination of Geographic Information and Related Spatial Data Activities, Tech. Rep. Circular A-16 revised, Washington, DC, https://www.fgdc.gov/policyandplanning/a-16, 2016.

OWP/ESIP: NWM Hydrofabric V2.1, s3://noaa-nws-owp-fim/hand_fim/fim_3_0_34_1/inputs/nwm_hydrofabric/, 2021.

Ozdemir, H., Sampson, C., de Almeida, G. A., and Bates, P.: Evaluating scale and roughness effects in urban flood modelling using terrestrial LIDAR data, Hydrology and Earth System Sciences, 17, 4015–4030, 2013.

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

Passalacqua, P., Do Trung, T., Foufoula-Georgiou, E., Sapiro, G., and Dietrich, W. E.: A geometric framework for channel network extraction from lidar: Nonlinear diffusion and geodesic paths, Journal of Geophysical Research: Earth Surface, 115, 2010.

Passalacqua, P., Belmont, P., and Foufoula-Georgiou, E.: Automatic geomorphic feature extraction from lidar in flat and engineered landscapes, Water Resources Research, 48, 2012.

Passalacqua, P., Belmont, P., Staley, D. M., Simley, J. D., Arrowsmith, J. R., Bode, C. A., Crosby, C., DeLong, S. B., Glenn, N. F., Kelly, S. A., et al.: Analyzing high resolution topography for advancing the understanding of mass and energy transfer through landscapes: A review, Earth-Science Reviews, 148, 174–193, 2015.

Petrochenkov, G. and Viger, R.: pygft: Rapid flood inundation modeling tool, Reston, VA, https://code.usgs.gov/gft/python-gis-flood-tool, 2020.

Pielke Jr, R. A. and Downton, M. W.: Precipitation and damaging floods: Trends in the United States, 1932–97, Journal of climate, 13, 3625–3637, 2000.

Podhorányi, M. and Fedorcak, D.: Inaccuracy introduced by LiDAR-generated cross sections and its impact on 1D hydrodynamic simulations, Environmental Earth Sciences, 73, 1–11, 2015.

Rajib, A., Merwade, V., and Liu, Z.: 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, 2016.

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

Roman, D. R., Wang, Y. M., Saleh, J., Li, X., et al.: Geodesy, geoids, and vertical datums: A perspective from the US National Geodetic Survey, in: FIG Congress, vol. 11, 2010.

Salach, A., Bakuła, K., Pilarska, M., Ostrowski, W., Górski, K., and Kurczyński, Z.: Accuracy assessment of point clouds from LiDAR and dense image matching acquired using the UAV platform for DTM creation, ISPRS International Journal of Geo-Information, 7, 342, 2018.

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

Sanyal, J. and Lu, X. X.: Application of remote sensing in flood management with special reference to monsoon Asia: a review, Natural Hazards, 33, 283–301, 2004.

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

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

Schumann, G.-P., Neal, J. C., Voisin, N., Andreadis, K. M., Pappenberger, F., Phanthuwongpakdee, N., Hall, A. C., and Bates, P. D.: A first large-scale flood inundation forecasting model, Water Resources Research, 49, 6248–6257, 2013.

Scott, D. W.: Multivariate density estimation: theory, practice, and visualization, John Wiley & Sons, 2015.

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

Shin, S., Pokhrel, Y., and Miguez-Macho, G.: High-resolution modeling of reservoir release and storage dynamics at the continental scale, Water Resources Research, 55, 787–810, 2019.

Silverman, B. W.: Density estimation for statistics and data analysis, Routledge, 2018.

Smiley, K. T., Noy, I., Wehner, M. F., Frame, D., Sampson, C. C., and Wing, O. E.: Social inequalities in climate change-attributed impacts of Hurricane Harvey, Nature communications, 13, 1–10, 2022.

Smith, M., Patrick, N., Frazier, N., Kim, J., Flowers, T., and Ogden, F.: Hyper Resolution Modeling of Urban Flood Inundation, 2020.

Smith, M., Patrick, N., Frazier, N., and Kim, J.: Validation of Urban Flood Inundation Models Applied Using Nationally Available Data Sets: Novel Analyses of Observed High Water Information, Journal of Hydrologic Engineering, 26, 04021 039, 2021.

Smith, T., Rheinwalt, A., and Bookhagen, B.: Determining the optimal grid resolution for topographic analysis on an airborne lidar dataset, Earth Surface Dynamics, 7, 475–489, 2019.

Snyder, G. I., Sugarbaker, L., Jason, A. L., and Maune, D. F.: National requirements for enhanced elevation data, US Department of the Interior, US Geological Survey, https://doi.org/http://dx.doi.org/10.3133/ofr20131237, open-File Report 2013–1237, 2013.

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

Stallings, R. A.: Conflict in natural disasters: A codification of consensus and conflict theories, Social science quarterly, 69, 569, 1988.

Stoker, J. and Miller, B.: The Accuracy and Consistency of 3D Elevation Program Data: A Systematic Analysis, Remote Sensing, 14, 940, 2022.

Stoker, J. M., Brock, J. C., Soulard, C. E., Ries, K. G., Sugarbaker, L., Newton, W. E., Haggerty, P. K., Lee, K. E., and Young, J. A.: USGS lidar science strategy: mapping the technology to the science, vol. 10, US Department of the Interior, US Geological Survey, 2015.

Strategic Alliance for Risk Reduction II: Base Level Engineering Analysis: Region 6 Neches River Watershed – Lower Angelina (HUC8 - 12020005), MIP Deliverable 16-09-0654S, Strategic Alliance for Risk Reduction II (STARRII), fEMA IDIQ Contract: HSFE60-15-D-0005, 2019a.

Strategic Alliance for Risk Reduction II: Base Level Engineering Analysis: Region 6 Neches River Watershed – Lower Neches (HUC8 - 12020003), MIP Deliverable 16-09-0654S, Strategic Alliance for Risk Reduction II (STARRII), fEMA IDIQ Contract: HSFE60-15-D-0005, 2019b.

Strategic Alliance for Risk Reduction II: Base Level Engineering Analysis: Region 6 Neches River Watershed – Middle Neches (HUC8 - 12020002), MIP Deliverable 16-09-0654S, Strategic Alliance for Risk Reduction II (STARRII), fEMA IDIQ Contract: HSFE60-15-D-0005, 2019c.

Strategic Alliance for Risk Reduction II: Base Level Engineering Analysis: Region 6 Neches River Watershed – Pine Island Bayou (HUC8 - 12020007), MIP Deliverable 16-09-0654S, Strategic Alliance for Risk Reduction II (STARRII), fEMA IDIQ Contract: HSFE60-15-D-0005, 2019d.

Strategic Alliance for Risk Reduction II: Base Level Engineering Analysis: Region 6 Neches River Watershed – Upper Angelina (HUC8 - 12020004), MIP Deliverable 16-09-0654S, Strategic Alliance for Risk Reduction II (STARRII), fEMA IDIQ Contract: HSFE60-15-D-0005, 2019e.

Strategic Alliance for Risk Reduction II: Base Level Engineering Analysis: Region 6 Neches River Watershed – Upper Neches (HUC8 - 12020001), MIP Deliverable 16-09-0654S, Strategic Alliance for Risk Reduction II (STARRII), fEMA IDIQ Contract: HSFE60-15-D-0005, 2019f.

Strategic Alliance for Risk Reduction II: Base Level Engineering Analysis: Region 6 Neches River Watershed – Village (HUC8 - 12020006), MIP Deliverable 16-09-0654S, Strategic Alliance for Risk Reduction II (STARRII), fEMA IDIQ Contract: HSFE60-15-D-0005, 2019g.

Strömberg, D.: Natural disasters, economic development, and humanitarian aid, Journal of Economic perspectives, 21, 199–222, 2007.

Sugarbaker, L., Constance, E. W., Heidemann, H. K., Jason, A. L., Lucas, V., Saghy, D., and Stoker, J. M.: The 3D Elevation Program initiative: a call for action, US Geological Survey Reston, VA, USA, 2014.

Survey, U. S. G.: 3DEPElevation (ImageServer), https://elevation.nationalmap.gov/arcgis/rest/services/3DEPElevation/ImageServer, 2022.

Tabari, H.: Climate change impact on flood and extreme precipitation increases with water availability, Scientific Reports, 10, 13 768, https://doi.org/10.1038/s41598-020-70816-2, 2020.

Tauhid, F. A. and Zawani, H.: Mitigating climate change related floods in urban poor areas: green infrastructure approach, Journal of Regional and City Planning, 29, 98–112, 2018.

Tellman, B., Schank, C., Schwarz, B., Howe, P. D., and de Sherbinin, A.: Using disaster outcomes to validate components of social vulnerability to floods: Flood deaths and property damage across the USA, Sustainability, 12, 6006, 2020.

Tellman, B., Sullivan, J., Kuhn, C., Kettner, A., Doyle, C., Brakenridge, G., Erickson, T., and Slayback, D.: Satellite imaging reveals increased proportion of population exposed to floods, Nature, 596, 80–86, 2021.

Teng, J., Vaze, J., Dutta, D., and Marvanek, S.: Rapid inundation modelling in large floodplains using LiDAR DEM, Water Resources Management, 29, 2619–2636, 2015.

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

Thomas Steven Savage, J., Pianosi, F., Bates, P., Freer, J., and Wagener, T.: Quantifying the importance of spatial resolution and other factors through global sensitivity analysis of a flood inundation model, Water Resources Research, 52, 9146–9163, 2016.

Tumbare, M. J.: Mitigating Floods in Southern Africa, in: First WAFSA/WATERNET Symposium, Maputo, 2000.

Turlach, B. A. et al.: Bandwidth selection in kernel density estimation: a rewiew, Tech. rep., Humboldt Universitaet Berlin, 1993.

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

UNEP: Early warning systems: a state of the art analysis and future directions, 2012.

UNISDR: Making development sustainable: the future of disaster risk management, 2015.

United States Army Corps of Engineers: National Inventory of Dams, https://nid.sec.usace.army.mil/, 2023.

United States Geological Survey: Watershed Boundary Dataset, https://www.usgs.gov/national-hydrography/watershed-boundary-dataset, 2023.

USGS: NHDPlusHR DEM, https://prd-tnm.s3.amazonaws.com/index.html?prefix=StagedProducts/Hydrography/NHDPlusHR/Beta/GDB/, 2021.

USGS: 3D Elevation Program: FY21 Status of 3DEP Quality Data, https://www.usgs.gov/3d-elevation-program, accessed 2022-09-28, 2021.

USGS: 3D Elevation Program: FY22 Partnerships, https://www.usgs.gov/3d-elevation-program, accessed 2022-09-28, 2022.

Verdin, J., Verdin, K., Mathis, M. L., Magadzire, T., Kabuchanga, E., Woodbury, M., and Gadain, H.: A software tool for rapid flood inundation mapping, Tech. rep., US Geological Survey, 2016.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald,

A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and SciPy 1.0 Contributors: SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python, Nature Methods, 17, 261–272, https://doi.org/10.1038/s41592-019-0686-2, 2020.

Wang, Y. and Zheng, T.: Comparison of light detection and ranging and national elevation dataset digital elevation model on floodplains of North Carolina, Natural Hazards Review, 6, 34–40, 2005.

Wijkman, A. and Timberlake, L.: Natural disasters: acts of God or acts of man?, Routledge, 2021.

Wing, O. E., Bates, P. D., Sampson, C. C., Smith, A. M., Johnson, K. A., and Erickson, T. A.: Validation of a 30 m resolution flood hazard model of the conterminous U nited S tates, Water Resources Research, 53, 7968–7986, 2017.

Wing, O. E., Bates, P. D., Smith, A. M., Sampson, C. C., Johnson, K. A., Fargione, J., and Morefield, P.: Estimates of present and future flood risk in the conterminous United States, Environmental Research Letters, 13, 034 023, 2018.

Witt III, E. C.: Evaluation of the US Geological Survey standard elevation products in a two-dimensional hydraulic modeling application for a low relief coastal floodplain, Journal of Hydrology, 531, 759–767, 2015.

Xu, J., Wang, Z., Shen, F., Ouyang, C., and Tu, Y.: Natural disasters and social conflict: A systematic literature review, International journal of disaster risk reduction, 17, 38–48, 2016.

Yang, L., Jin, S., Danielson, P., Homer, C., Gass, L., Bender, S. M., Case, A., Costello, C., Dewitz, J., Fry, J., et al.: A new generation of the United States National Land Cover Database: Requirements, research priorities, design, and implementation strategies, ISPRS journal of photogrammetry and remote sensing, 146, 108–123, 2018.

Zahran, S., Shelley, T. O., Peek, L., and Brody, S. D.: Natural disasters and social order: Modeling crime outcomes in Florida, International Journal of Mass Emergencies and Disasters, 27, 26–52, 2009.

Zhang, J., Huang, Y.-F., Munasinghe, D., Fang, Z., Tsang, Y.-P., and Cohen, S.: Comparative analysis of inundation mapping approaches for the 2016 flood in the Brazos River, Texas, JAWRA Journal of the American Water Resources Association, 54, 820–833, 2018.

Zheng, X., Maidment, D. R., Tarboton, D. G., Liu, Y. Y., and Passalacqua, P.: GeoFlood: Large-Scale Flood Inundation Mapping Based on High-Resolution Terrain Analysis, Water Resources Research, 54, 10–013, 2018a.

Zheng, X., Tarboton, D. G., Maidment, D. R., Liu, Y. Y., and Passalacqua, P.: River channel geometry and rating curve estimation using height above the nearest drainage, JAWRA Journal of the American Water Resources Association, 54, 785–806, 2018b.

Zheng, X., Godbout, L., Zheng, J., McCormick, C., and Passalacqua, P.: An automatic and objective approach to hydro-flatten high resolution topographic data, Environmental Modelling & Software, 116, 72–86, 2019.

Zheng, X., D'Angelo, C., Maidment, D. R., and Passalacqua, P.: Application of a Large-Scale Terrain-Analysis-Based Flood Mapping System to Hurricane Harvey, JAWRA Journal of the American Water Resources Association, 58, 149–163, 2022.