https://doi.org/10.5194/essd-2020-291
Preprint. Discussion started: 21 November 2020

Open Access Earth System Science Data Discussions

A national topographic dataset for hydrological modeling over contiguous United States

Jun Zhang[1], Laura E. Condon[1], Hoang Tran[2] and Reed M. Maxwell[2]

[1]Department of Hydrology and Atmospheric Sciences, The University of Arizona, Tucson, AZ
[2]Department of Civil and Environmental Engineering, Princeton University, Princeton, NJ

Correspondence to: Jun Zhang (junzhang55@arizona.edu)

Abstract. Topography is a fundamental input to hydrologic models critical for generating realistic streamflow networks as well as infiltration and groundwater flow. Although there exist several national topographic datasets for the United States, they may not be compatible with gridded models that require hydrologically consistent Digital Elevation Models (DEMs). Here, we present a national topographic dataset developed support physically based hydrologic simulations at 1km and 250m spatial resolution over contiguous United States. The workflow is described step-by-step in two parts (a) DEM processing using a Priority Flood algorithm to ensure hydrologically consistent drainage networks and (b) slope calculation and smoothing to improve drainage performance. The accuracy of derived stream network is evaluated by comparing the derived drainage area to drainage areas reported by the national stream gage network. The slope smoothing steps are evaluated using the runoff simulations with an integrated hydrologic model. The processed DEM is designed to capture the topographic features and improve the runoff simulations for the models solving partial differential equations. The workflow uses an open-source R package and all output datasets and processing scripts are available and fully documented here. All of the output datasets and scripts for processing are published through Cyverse at 250m and 1km resolution. The DOI link for the dataset is https://doi.org/10.25739/e1ps-qy48 (Zhang and Condon, 2020).

1 Introduction

Topography is one of the most important inputs to hydrologic simulations; it defines watershed boundaries and shapes river networks. Surface flow travel times and runoff characteristics are very sensitive to hillslope characteristics (D'Odorico and Rigon, 2003; Freer et al., 2002; Frei et al., 2010; Gupta and Mesa, 1988). In addition to shaping surface flow networks, groundwater fluxes and residence times are also strongly driven by topographic gradients (e.g. Condon and Maxwell, 2015). While high resolution elevation data is not difficult to find; for example the National Elevation Dataset provides a 30m resolution Digital Elevation Model (DEM) across the US (Gesch et al., 2002). It is well established that topographic datasets require processing to be suitable for hydrologic simulation because flow networks and slopes can be sensitive to noise in the DEM and can be affected by the resolution and spatial gridding of a DEM (Habtezion et al., 2016; Sørensen and Seibert, 2007; Thompson et al., 2001; Vaze et al., 2010; Wolock and McCabe, 2000; Wu et al., 2008; Zhang and Montgomery, 1994). This processing generally consists of some combination of steps to (1) remove erroneous local minima that impeded flow (Kenny et al., 2008; Lindsay, 2016b, 2016a); (2) lower or 'burn' in the drainage network to ensure that flow occurs

https://doi.org/10.5194/essd-2020-291
Preprint. Discussion started: 21 November 2020

along identified stream segments (Lindsay, 2016b; Woodrow et al., 2016), (3) smooth the DEM to remove noise introduced by the sampling resolution (Gallant, 2011; Lindsay et al., 2019).

Although there are many tools and methods available to process topography for hydrologic applications, processing is generally site specific and we lack a national topographic dataset for the US that is designed for physical hydrology simulations. At the national scale there are three main topographic datasets for the US. First, as previously noted the National Elevation Dataset (NED), provides 30m national DEM primarily derived from U.S. Geological Survey (USGS) 10m and 30m DEMs (Gesch et al., 2002). This provides a high-resolution national topography dataset, however it includes no processing for hydrologic utilizations.

Second, the National Hydrography Dataset Plus (NHDplus) dataset provides a collection of geospatial data including various key features of stream network including elevation, drainage area and watershed boundary. It is derived from the 10m USGS elevation data and the national complete Watershed Boundary Dataset (U.S. Geological Survey, 2017). Because its public availability, high-resolution and the large spatial coverage, the dataset has been applied in many hydrologic studies considering runoff, river routing and flood inundation (David et al., 2011; Garousi-Nejad et al., 2019; Smith et al., 2013; Zhang et al., 2018). The NHDplus dataset provides the most complete hydrologic mapping of the US; however spatial inconsistences and noise in these datasets makes it difficult to this data directly for gridded hydrologic modeling. The NHDplus stream network is derived from various topographic map sources, leading to spatial inconsistencies and inaccuracies in river network (Moore et al., 2019). Spatial discrepancies have also been found between NHDplus and local higher-resolution light detection and ranging data (LiDAR) derived stream network (Samu, 2012). Discrepancies were also noted between NHDplus watersheds and Height Above Nearest Drainage (HAND) derived from 10m DEM (Garousi-Nejad et al., 2019).

Third, the National Water Model (NWM) is a national hydrologic modeling framework that simulates terrestrial hydrology at 250m spatial resolution. This model has its own DEM and stream mask that are derived from the 30m DEM in NED and stream network in NHDplus (Gochis et al., 2018). The TauDEM topographic processing tool (Tarboton, 2005) was applied to these inputs to generate a connected drainage network in D8 routing (Gochis et al., 2018). Although the NWM DEM does include topographic processing, it is still not directly suitable for gridded overland flow simulations because in the NWM (and many NHDPlus analyses) the stream reaches are processed as a network using 1-D Muskingum-Cunge routing method rather than following the actual DEM grid (Gochis et al., 2018; Johnson et al., 2019). As a result, these applications do not need to directly examine the DEM quality along the streams.

This is not the case for partial differential equations (PDE) based models that do gridded simulations. For these models topographic inputs are critical for shaping both subsurface (e.g. groundwater) and surface flow simultaneously. These models require a topographic dataset which represents these stream reaches accurately within the DEM itself. Also, PDE based solutions will often require D4 routing (as flux is transmitted across grid cell faces) rather than D8 routing used in processing such as for the NWM. We lack a national topographic dataset that is fully compatible with physically based hydrologic models requiring a hydrologically consistent DEM and a connected D4 routing everywhere.

https://doi.org/10.5194/essd-2020-291
Preprint. Discussion started: 21 November 2020

Here we present a high-resolution topographic dataset of the contiguous United States (CONUS) designed for PDE based hydrologic modeling applications that simulate surface and subsurface flows with a single, consistent topographic input. The workflow is discussed step-by-step in the following sections in two parts (a) DEM processing by Priority Flood algorithm (b) slope calculation and smoothing. Several cases are implemented for both parts with comparative analysis to illustrate the improvements of additional procedures. We compared the drainage area with measurements and applied slope into a hydrologic model to evaluate the surface flow simulation.

2 Data and Methods

This section covers the detailed information about input datasets including DEM and stream network maps (Section 2.1); topographic processing methodology for DEM adjustment and slope calculation (Section 2.2); and evaluation metrics used to assess the accuracy of drainage network and drainage performance (Section 2.3).

2.1 Domain and Input datasets

Our approach requires three primary inputs (1) a gridded Digital Elevation Model (DEM) of topography, (2) a mask of known stream locations (3) a mask of lakes and terminal points for endorheic basins (shown in Figure 1). All analysis covers the contiguous United States plus areas in Canada and Mexico that drain to the contiguous United States. This spatial coverage is consistent with the NOAA National Water Model (NWM) (2018). All analysis is completed and available at 1km and 250m spatial resolution, but for the purposes of discussion we focus on the 1km analysis throughout the results section. Our input datasets are primarily derived from the NWM version 1.2. The NWM inputs are derived from publicly available national datasets and have already been processed for hydrologic consistency within their modeling framework which provides an ideal starting point for this work. Additionally, by starting from NWM datasets we ensure that the resulting products are as consistent as possible with this growing framework of national hydrologic datasets.

We start from a 250m DEM that was developed for the NWM V1.2, which was developed from the 30m resolution National Elevation Dataset (Gesch et al., 2002) and processed using the TauDEM (Tarboton, 2005) to ensure drainage in a D8 routing scheme (i.e. with flow allowed out of every grid cell in 8 directions). We start from this DEM directly for 250m analysis, and for the 1km analysis we upscale the NWM DEM by selecting the minimum elevation within every 1km cell. The derived stream network is significantly affected by the upscaling approach (Moretti and Orlandini, 2018). The DEM is upscaled based on the minimum as opposed to the average because this is the appropriate in identifying the lowest points and capturing hydrologic features such as streams.

The stream network mask is a subset of the NWM stream mask based on Strahler Stream Order (Horton, 1945; Strahler, 1957). NWM stream mask is a rasterized version of the NHDplus V2.1 stream network lines (U.S. Geological Survey, 2017). The stream mask is used to guide the drainage patterns in the DEM for the topographic processing in this study. By experimenting with different stream densities, we found that the stream networks consisting of 5th order and higher streams

https://doi.org/10.5194/essd-2020-291
Preprint. Discussion started: 21 November 2020

Open Access Earth System Science Data Discussions

provides the best guidance in 1km resolution processing and a denser network consisting of 3rd order and above streams is ideal for 250m resolution processing. Figure 1 shows the 5th stream network that was used for the 1km DEM processing. For both the 3rd and 5th order stream masks additional manual adjustments were made to improve the resulting drainage network, which is discussed in more detail in Section 3.1.

The final input is a mask of lakes and terminal points for endorheic basins. Our topographic processing ensures that all grid cells in the domain drain to either an ocean cell, internal lake or a terminal point within endorheic basins, referred to as 'sinks'. The lakes and sinks presented in our domain mask provide end points for all internally draining (i.e. not draining to the ocean) cells. We selected 13 major lakes from the NWM lake shapefile which was developed based on NHDPlus. Only the major terminal lakes which are critical for watershed delineation were included in this topographic processing. The extents of these lakes are rasterized as shown in Figure 1 with the borders treated as outlets and the interior lake cells are excluded from topographic processing.

For the more arid endorheic basins such as are commonly found in the Great Basin of the western US intermittent rivers drain to much smaller water bodies which may not be perennially inundated. For these basins we designate single cells at the river outlet as sink cells which were added to the domain manually based on the 5th order stream mask. A total of 262 sinks were added to the domain as shown in Figure 1. The significance of these sinks for topographic processing is discussed further in Section 3.1.

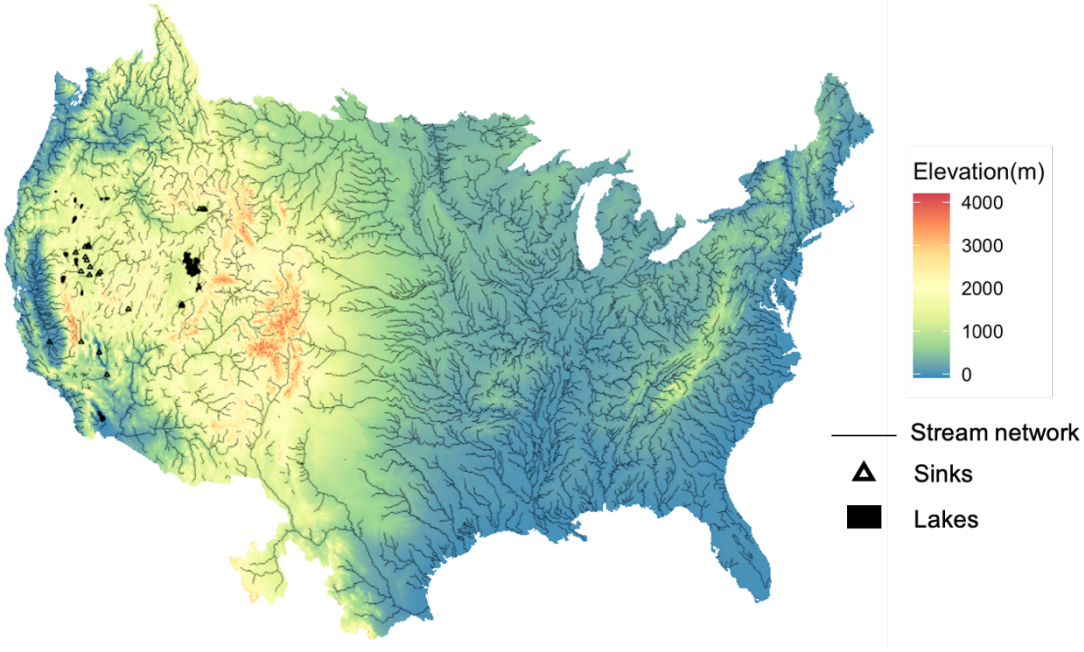


Figure 1: Map of the DEM datasets, stream network, lakes and sinks used for processing the 1km dataset.

https://doi.org/10.5194/essd-2020-291
Preprint. Discussion started: 21 November 2020

2.2 Topographic Processing Workflow

While some hydrologic processing has already been implemented in NWM DEM as described in the previous section, additional processing is needed for the purpose of our processing (1) ensure a fully connected D4 (as opposed to D8) drainage network and (2) to improve runoff performance for gridded overland flow simulations. Figure 2 summarizes the topographic processing workflow for the final dataset we present. All of the processing steps were completed using the Priority Flow toolbox (Condon and Maxwell, 2019) which is available on GitHub (https://github.com/lecondon/PriorityFlow) and can be installed as an R library. The parameters for each step of the workflow and their values are provided in Table 1. The primary output from the topographic processing workflow is a hydrologically consistent DEM processed to ensure realistic drainage networks and overland flow patterns. In addition to the DEM many other outputs are generated throughout the processing that define the drainage network, slopes and watersheds. These outputs are listed in Figure 2 and summarized in more detail in Section 4. The steps of the processing workflow are as follows:

1. *Minor Stream Burning*: First we apply a very small stream 'burn in' decreasing the elevation of all cells on the stream mask by one meter. Stream burning is an approach that has been used in previous studies to enforce flow along predefined channel segments (Lindsay, 2016b). Here a very minor elevation adjustment of one meter is to help locate streams in very flat parts of the domain where the channel may not be captured at all by the 250m or 1km DEM. A larger stream burn is unnecessary due to the later processing steps which ensure that the DEM processing reflects the channel network.
2. *Initial Topographic smoothing*: It is well established that the spatial resolution of a DEM can introduce artificial noise in the observed topography. Therefore, before processing we apply a smoothing filter to the DEM (Habtezion et al., 2016). Here we use a feature preserving smoothing approach from the *whitebox* library in R to remove the DEM roughness for the domain. The approach is a modified feature-preserving normal vector field smoothing technique so it is feasible for large DEM raster (Lindsay et al., 2019). To remove the nosie as well as preserve the topographic characteristics, the filter kernel size (*filter*) is set to 20 grid cells, the maximum difference in normal vectors (*norm_diff*) is 15 degree, Three elevation-update iterations (*num_iter*) are implemented with maximum allowable absolute elevation change (*max_diff*) of 5m for the whole domain.
3. *Priority Flood Algorithm to Ensure D4 drainage:* To be useful for Hydrologic simulation non-physical local minima must be removed from the domain. In this step we ensure that every cell in the domain has a drainage path out of the domain or to one of the predefined lake or sink internal boundaries. Here, we have an additional requirement that this drainage be accomplished in D4 (i.e. only in the primary North, South, East, West directions), this is useful for partial differential equation (PDE) based hydrologic simulations where fluxes occur across cell faces. The Priority Flood Algorithm is a well-established and mathematically optimal approach to remove erroneous local minima in a DEM and ensure complete drainage (Barnes et al., 2014; Liu et al., 2009; Soille and

https://doi.org/10.5194/essd-2020-291
Preprint. Discussion started: 21 November 2020

Gratin, 1994; Wang and Liu, 2006; Zhou et al., 2016). The standard priority flood approach starts from the exit points of a domain and works sequentially inward lifting cells as required to ensure that every cell is guaranteed a monotonically decreasing path to the exit. The Priority Flood algorithm is mathematically optimal in that it achieves a fully draining DEM with the least possible adjustments (or filling) to the original DEM, however it may not always yield the desired drainage network given a noisy or low-resolution DEM.

To address this limitation, Condon & Maxwell (2019) developed a new approach applying the priority flood algorithm in two steps: first along a pre-specified drainage network and then to the rest of the cells in the domain treating the processed stream network as the outlets. With this approach the advantages of the Priority Flood algorithm are maintained by the user and also can incorporate information about the stream network to prioritize drainage along this pre-specified network first. Additionally, the user can specify an optional epsilon value to be added to cells as they are processed by the algorithm. If an epsilon is used, then filling will be applied such that upstream cells are at least epsilon greater than their downstream neighbor. Here we use a *stream_epsilon* values of 0 for the first round of processing along the stream and a *global_epsilon* value of 0.1m for the rest of the domain. A small value of epsilon is to guarantee the drainage without large change to the original DEM. The results of this processing step are (1) a fully draining DEM and (2) a set of D4 flow directions indicating the primary flow direction for every grid cell in the domain.

4. *Stream elevation smoothing*: While the Priority Flood Algorithm ensures a monotonically decreasing path from any grid cell in the domain to an exit, it does not evaluate the smoothness of this path. Furthermore, at 250m or 1km resolution the DEM is more reflective of hillslope scale patterns than channel bathymetry. For physically based hydrologic simulations slopes along stream channels are an important variable controlling surface water drainage. Smoothing elevation gradients along stream reaches has been shown to decrease artificial ponding and improve the runoff simulations (Barnes et al., 2016).

   For the purposes of this smoothing step we first derive a new stream network using a drainage area threshold (*stream_th*). The flow direction files resulting from the Priority Flood processing can be used to calculate the drainage area for every grid cell in the domain. From this a stream network of any density can be derived based on the *stream_th*. This allows us to apply smoothing to a denser stream network than was used in the earlier Priority Flood processing script. The smaller the *stream_th* the denser the identified stream network will be, and the more cells will be smoothed. This stream network can then be divided into segments and associated sub-watersheds. Here we use a drainage area threshold of 100 km^2.

   Stream segments are smoothed by adjusting the elevations along the segment to maintain a constant slope from the starting to the ending point of each segment. After these adjustments are made to the stream reaches an additional elevation processing pass must be made to ensure that all of the bank cells along a stream reach still drain to the stream. Here we apply an epsilon threshold (*bank_epsilon*) to require that adjacent cells are at least *bank_epsilon* greater than their downstream neighbor. The bank adjustment step traverses up a stream segment checking that all

https://doi.org/10.5194/essd-2020-291
Preprint. Discussion started: 21 November 2020


Open Access Earth System Science Data Discussions

cells draining to every stream cell (based on flow direction) are at least *bank_epsilon* (0.1m in this dataset) higher. Anywhere this is not the case, neighboring cells are filled accordingly and processing continues up the bank until it reaches a point where all upstream neighboring cells are fully draining.

5. *Slope calculations and final slope smoothing:* Some hydrologic simulations require slopes rather than elevations as inputs. Therefore, the final step in processing is to convert the processed DEM to slopes in the x and y directions. Here we include three additional processing steps:
    a. *Applying maximum and minimum slope thresholds:* Maximum and minimum slope thresholds can be specified globally or for the hillslope (*slope_max, slope_min*) and stream cells (*slope_min_stream*) separately. Setting minimum slope thresholds has a similar effect to applying an epsilon in the Priority Flood algorithm, however it can be preferable because it does not propagate to upstream cells. For example, in very flat portions of the domain, even small epsilons can lead to large DEM adjustments as they are additive. The maximum slope is limited for the whole domain; *slope_min* and *slope_min_stream* are set to 10^{-4}.
    b. *Removing or dampening secondary slopes:* The priority Flood algorithm provides primary D4 flow directions for every grid cell (referred to as the primary flow direction). However, slopes are calculated for every cell in both x and y directions (i.e. in the primary flow direction and a secondary flow direction). Some previous approaches removed all slopes that transverse to the primary flow direction. For example, if a cell's primary flow direction is north then only its y slope would be used and its x slope would be set to zero (Barnes et al., 2016). However, previous research has shown that hydrologic performance is improved when secondary slopes are included (Barnes et al., 2014; Daniels et al., 2011). Additional details on slope calculations and secondary slope smoothing can be found in (Condon and Maxwell, 2019). Here we set no restrictions for the secondary slope outside the stream network, but remove the secondary slope along the stream network to improve the drainage inside the stream.
    c. *Fixing flat cells:* Finally, despite the fact that the Priority Flood algorithm ensures that all cells can drain and there are no local minima in the domain, anomalous water ponding points can still be created where the outlet slope of a grid cell is significantly less than the cells draining too it. Therefore, as a final smoothing step if the total slope out divided by the total slope into a given cell is less than a specified adjustment ratio (*adj_th,* set to 10^{-3} here*)*, the new outlet slope will be set to initial outlet slope times some scaler *(adj_ratio,* set to 10 here*)*.

https://doi.org/10.5194/essd-2020-291
Preprint. Discussion started: 21 November 2020

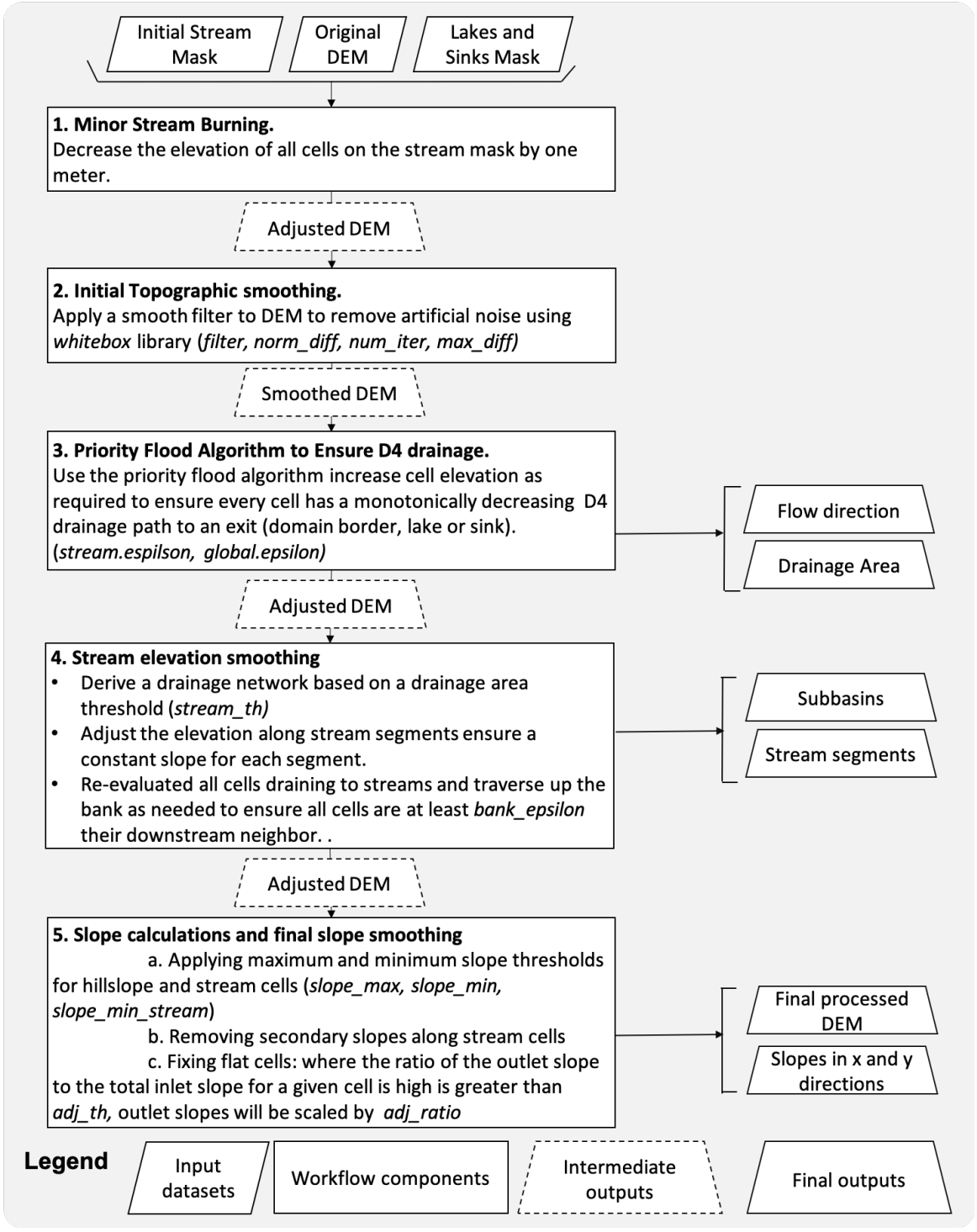


Figure 2: Workflow diagram outlining the major processing steps used and the outputs generated, input parameters are listed in italics and described in more detail in Table 1.

https://doi.org/10.5194/essd-2020-291
Preprint. Discussion started: 21 November 2020

Open Access Earth System Science Data Discussions

215 **Table 1: Parameters for the topographic processing and values in this dataset**

Parameters	Explanation	Value	Workflows step (*Name of function)*
filter	Size of the filter kernel.	20	(2) Initial Smoothing *(whitebox)*
norm_diff	Maximum difference in normal vectors, in degrees.	15	(2) Initial Smoothing *(whitebox)*
num_iter	Number of iterations.	3	(2) Initial Smoothing *(whitebox)*
max_diff	Maximum allowable absolute elevation change.	5	(2) Initial Smoothing *(whitebox)*
stream_epsilon	The minimum elevation difference required between upstream and downstream neigbhor cells when applying priority flood along the stream network	0	(3) Priority Flood (*StreamTraverse*)
global_epsilon	The minimum elevation difference required between upstream and downstream neigbhor cells when applying priority flood to the domain outside the stream network	0.1	(3) Priority Flood (*D4TraverseB*)
stream_th	Drainage area threshold used designate cells as stream cells	100	(4) Stream Smoothing (*CalcSubbasins*)
bank.epsilon	the minimum elevation difference between bank and stream cells	0.1	(4) Stream Smoothing (*RiverSmooth*)
slope_max	Maximum absolute value of slopes. If this is set to -1 the slopes will not be limited	-1	(5) Slope Calculation (*SlopeCalStan*)
slope_min	Minimum absolute slope value to apply to flat cells if needed	10^{-4}	(5) Slope Calculation (*SlopeCalStan*)
slope_min_stream	Minimum slope threshold to be applied only to the primary flow directions of stream cells	10^{-4}	(5) Slope Calculation (*RivSlope*)
adj_th	Threshold for slope adjustment. If the total slopes out of divided by the total slopes into a given cell is less than *adj_th* the outlet slope will be scaled by *adj_ratio.*	10^{-3}	(5) Slope Calculation (*FixFlat*)
adj_ratio	Scaler value for slope adjustment. New outlet slopes will be set to initial outlet slope times *adj_ratio*.	10	(5) Slope Calculation (*FixFlat*)

https://doi.org/10.5194/essd-2020-291
Preprint. Discussion started: 21 November 2020


Open Access Earth System Science Data Discussions

2.3 Evaluation Metrics

As noted in Section 2.1, there are already multiple national elevation datasets that are well established and have been previously validated. Our goal is to start from these datasets to develop a national topographic dataset that is suitable for gridded hydrologic simulation. Because we are starting from established elevation models (DEM for NWM V1.2) our evaluation is focused on the hydrologic improvements to the dataset. Specifically, we evaluate, (1) the extent to which the resulting drainage network matches observations and (2) the drainage characteristics of the resulting dataset. Details on the approach for evaluating each of these metrics follows:

1. *Accuracy of drainage network:* The location of drainage networks can be very sensitive to topographic processing especially when working with relatively low-resolution DEMs which do not resolve small stream channels. This is why we incorporate stream network information into the topographic processing with the initial stream burning step and in the priority flood algorithm. We evaluate the resulting stream network by calculating the drainage area for every grid cell in the domain based on the derived flow directions and comparing with the drainage areas reported through the USGS stream gage network. The Gages-II dataset includes 7,542 gages with drainage area ranging from 0.02 km^2 to 2,975,585 km^2 across the US (Falcone, 2011). When mapping the gages to the DEM we use the reported stream gage latitude and longitude to identify the closest grid cell. We provide a 3km buffer around each stream gage allowing the gage to move up to 3 grid cells in any direction for 1km resolution (12 grid cells for 250m resolution) if the percentage difference of resulting drainage area between calculated and USGS observation is less than 20%. This adjustment step corrects for locations where the grid resolution may result in a gage being mapped to a hillslope cell adjacent to the intended stream cell or where a gage may incorrectly fall on a tributary as opposed the main stem (or vice versa). After allowing for these minor adjustments we evaluate percentage difference in drainage area between the simulated drainage areas based on the topographically processed DEM and the stream gage network.
2. *Drainage performance*: In addition to the drainage network patterns we apply a hydrologic model to evaluate the hydrologic performance of the entire domain. Here we are concerned primarily with anomalies in the resulting DEM which inaccurately disrupt flow. To test this, we apply the hydrologic model ParFlow (Ashby and Falgout, 1996; Jones and Woodward, 2001; Kollet and Maxwell, 2006; Kuffour et al., 2020; Maxwell, 2013) which is an integrated hydrologic model that includes groundwater and surface water simulation, but for the purposes of this study we focus only on the overland flow simulation. ParFlow simulates overland flow using Manning's equation and the Kinematic Wave formulation where flow is driven by the bed slope and surface roughness (as specified by the Manning's Roughness coefficient). Here we evaluate the ponding depth of water over time to evaluate drainage patterns and identify stagnation points.

   The model is configured with a single shallow layer (0.1m) with an impervious surface (permeability in 10^{-6}m/hr) so that groundwater flow and infiltration processes are essentially ignored. A constant Manning's roughness

https://doi.org/10.5194/essd-2020-291
Preprint. Discussion started: 21 November 2020

coefficient to 4.4×10^{-4} is applied across the entire domain. To test the runoff performance, we simulate a single rainfall recession event. The runoff test consists of a 200 hours simulation with one-hour rainfall event at a rate of 50mm followed by 199 hours of recession. The intent with this idealized impervious runoff simulation is to isolate the impacts of topography on runoff processes. The runoff test provides an easy visual way to identify locations where the topography is contributing to erroneous runoff behavior and to quantitatively identify locations with poor drainage or anomalous ponding.

3 Topographic processing evaluation

We explored a wide range of parameter combinations in our testing to obtain a reliable topographic dataset for hydrologic modeling. Table 1 outlines all of the parameters that were used for the final topographic processing reflected in the datasets published here. For the purposes of this discussion we focus on the processing steps that had the largest impact on the accuracy of the resulting drainage network spatially (Section 3.1) and the runoff dynamics (Section 3.2). The goal here is not to present a sensitivity study of parameters in topographic processing. Rather this discussion is provided to illustrate the key topographic processing steps that can improve hydrologic performance. This section demonstrates the improvements of the published dataset over existing topographic products, specifically for hydrologic modeling applications. Furthermore, the workflow described above, and the results presented here can help guide others developing DEMs for different domains or resolutions.

3.1 Drainage Network Analysis

In this section we explore the impact of incorporating a prior stream network to guide the topographic processing. Figure 3 shows the difference in drainage area between the processed DEM and reported stream gage area when the processing is run without an initial stream mask included in the priority flow adjustment step. Figure 3a shows the percent difference in drainage area for all 7,542 gages used for comparison, while Figure 3b shows only 1,195 gages with drainage area over 5000 km^2. As illustrated by the green dots in the figure, in many locations the stream network matches well with stream gage observations even without incorporating the stream network. However, there are also a large number of gages with very large mismatches in drainage (1,661 gages overall and 204 large gages have more than a 20% difference in drainage area). Many of the poorly performing gages shown by red and dark blue dots are located in flat regions (e.g. downstream of Mississippi River and Great Lakes) and the southwest where many internally draining basins exist. Although, poor performance is not limited to these areas exclusively. The large number of gages with large drainage area differences demonstrate the need to incorporate some stream location information into the topographic processing.

https://doi.org/10.5194/essd-2020-291
Preprint. Discussion started: 21 November 2020

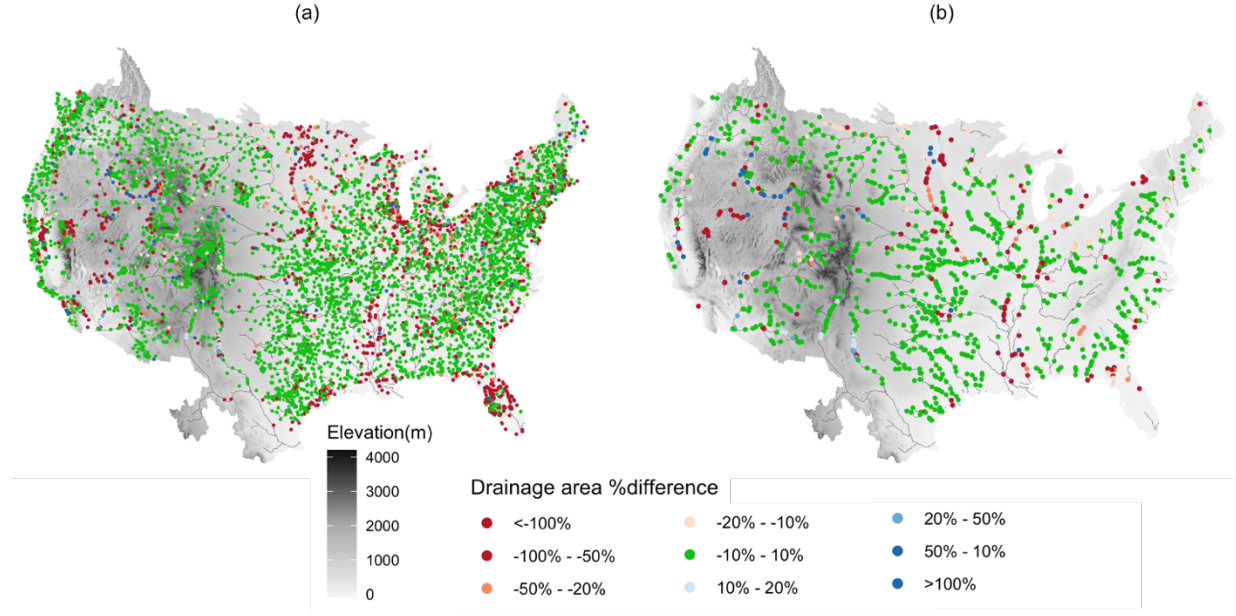


Figure 3: Drainage area comparison between processed with no initial stream network and USGS gage measurements. (a) all gages for comparison; (b) gages with drainage area over 5000 km^2

To address this, we use the 5th order stream mask to guide the topographic processing of the 1km domain (3rd order for the 250m domain). Figure 4 shows the stream gage performance maps by adding 5th order stream mask. Adding the stream mask to processing increases the number of gages with area matches within 10% from 5,066 in the original processing without any stream mask provided to 5,189 (880 to 937 for the large stream gages). Furthermore, it decreases the number of gages with area differences greater than 20% to 204 overall and 158 among the large gages. This is a clear improvement over the processing without a stream mask. However, there are still many locations with poor drainage area matches, including large gages.

https://doi.org/10.5194/essd-2020-291
Preprint. Discussion started: 21 November 2020

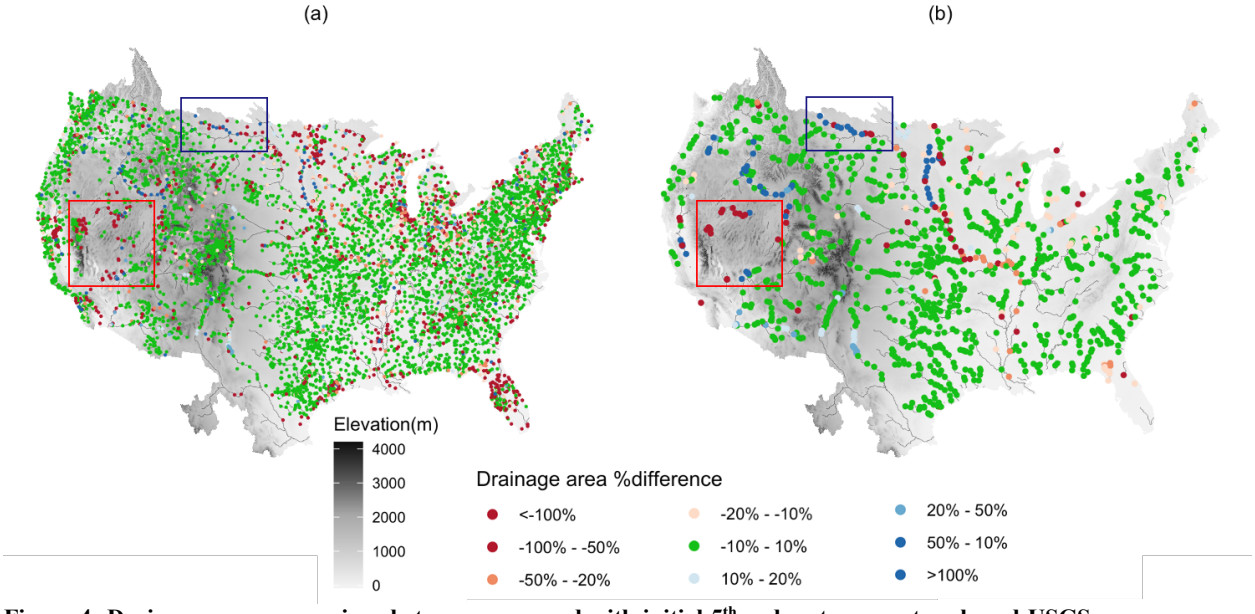


Figure 4: Drainage area comparison between processed with initial 5th order stream network and USGS gage measurements. (a) all gages for comparison; (b) gages with drainage area over 5000 km^2

Therefore, as a final step we applied a series of manual fixes to the stream network. Three primary types of fixes were applied: (1) adding sinks to the domain to increase the number of terminal points for internally draining basins, (2) manually modifying the stream mask to reflect higher resolution information, (3) removing erroneous stream outlets that were created where the headwater cells in the stream mask intersects the boundary of the domain. This last adjustment corrected some of the anomalously high drainage areas in the norther portion of the domain (blue box in Figure 4). For discussion purposes we will focus on the first two adjustments that impacted a much larger portion of the stream network.

Sinks were added primarily in the great basin portion of the domain (the red box in Figure 4), where endorheic basins are common. This arid portion of the domain is characterized by ephemeral streams (i.e. streams that do not flow year-round) that are often poorly mapped. Figure 5 shows a portion of the Great Basin before and after sinks were added. Figure 5a shows the input topography and cells with drainage area over 10^3 km^2 processed by Priority Flood approach without sinks. Before sinks are added to the stream network, the Priority Flood algorithm forces all of the cells within this area to drain out resulting in the linear stream network that does not align with the topography well and poor drainage area matching. The triangles in Figure 5b show the sinks that were adding according to the topography and 5th order stream network. Including these as terminal points in the topographic processing improves the resulting stream network and stream gage area matches significantly. Overall a total of 262 sinks were added to the domain.

https://doi.org/10.5194/essd-2020-291
Preprint. Discussion started: 21 November 2020

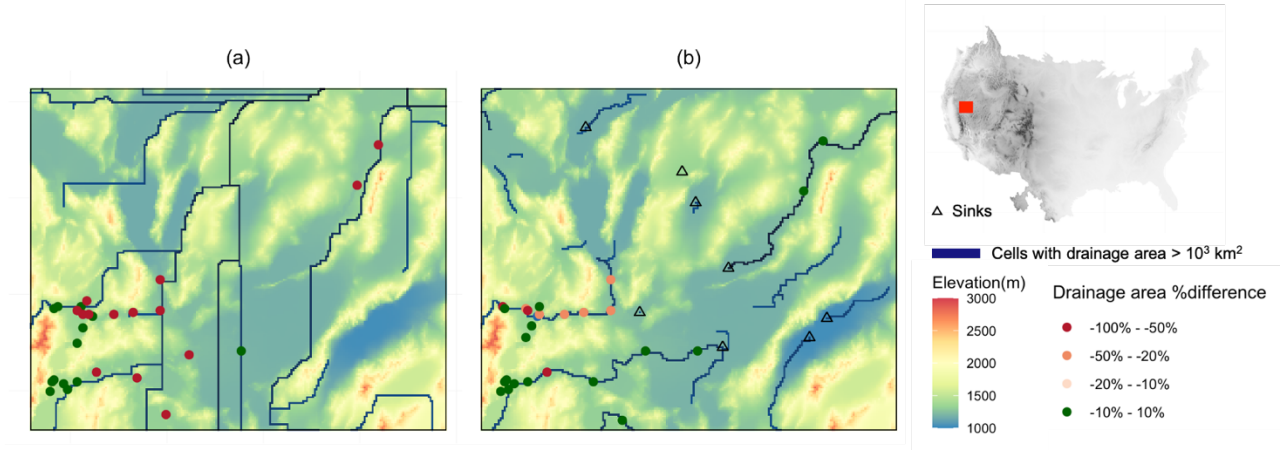


Figure 5: input topographic map (background of (a) and (b)) of a portion in the Great Basin shown as the red box in the upper-right CONUS map; the drainage network (blue cells in (a) and (b)); the percentage different of drainage area (colored dots) between processed and USGS observations (a) without sinks and (b) with pre-defined sinks

By far the most common manual update that was made in the processing step was to manually update the location of the stream mask. This was accomplished by visually comparing the 5th order stream network to the higher order streams and NHD streamlines and extending the 5th order stream mask as needed to reflect this higher resolution data. Note that simply applying a denser stream mask nationally (for example using the 3rd order stream mask for the 1km domain processing) is not a viable solution here. When translated to the 1km grid these denser stream masks result in too many grid cells being classified as stream cells and decreased quality of the stream network overall. Therefore, we work primarily from the 5th order stream mask but incorporate higher resolution information as is needed in an iterative process comparing our resulting drainage network to the stream gauge areas. Figure 6 gives an example of a location at the northern portion of the domain where the 5th order stream mask was expanded to improve performance of stream network derivation. Figure 6a shows the 5th order stream network in black line and 3rd order stream network in red line. Figure 6b shows the linear drainage network that results from the initial processing with 5th order stream. In this flat region, it is difficult to obtain the correct drainage path with only the topographic information. By adding 3rd order stream network, the drainage path presents a more reasonable pattern and an improved drainage area matching in Figure 6c. A similar process was completed for the 250m resolution starting from the 3rd order streams and extending using the 2nd order streams.

https://doi.org/10.5194/essd-2020-291
Preprint. Discussion started: 21 November 2020

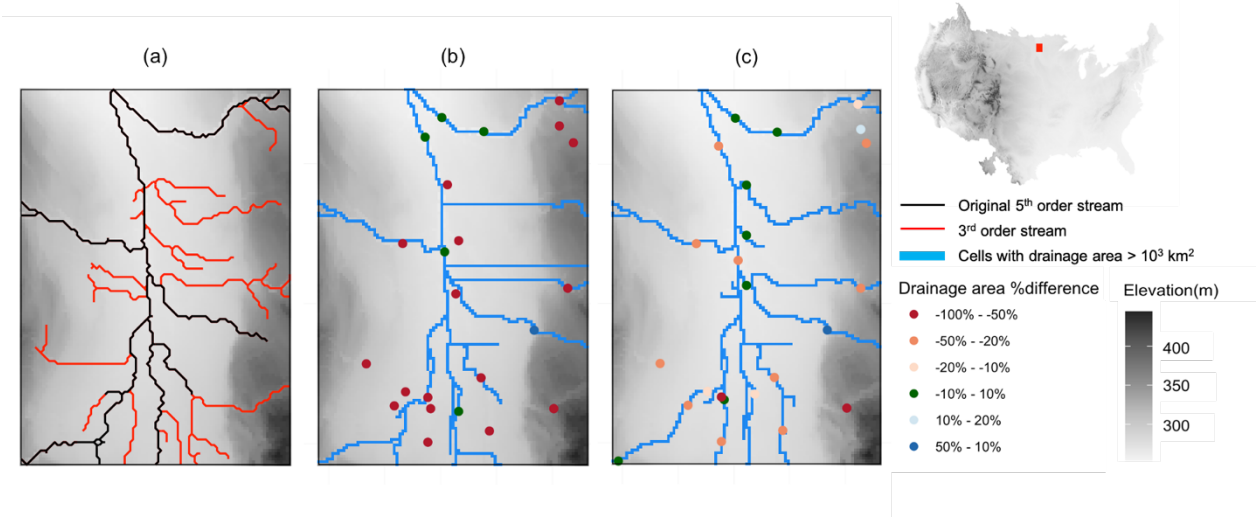


Figure 6: input topographic map (background of (a), (b) and (c)) of a valley in the northern portion of CONUS shown as the red box in the upper-right CONUS map. (a) the 3rd (red) and 5th (black) order stream network; the drainage network (blue cells in (b) and (c)); the percentage different of drainage area (colored dots) between topographic processed and USGS measurement (b) with 5th order stream network and (c) with stream networks expanded by 3rd order stream network

Figure 7 shows the final drainage area evaluation after all of the adjustments to the stream network are incorporated. Overall this demonstrates a strong improvement in the drainage area network relative to both the case with no stream network at all and the original 5th order stream mask. Nearly all (90.0%) of the large drainage area gages (from 880 with no stream mask to 1,075 gages with the final stream mask) now have area agreements within 10% and 74.2% of the gages overall (from 5,066 with no stream mask to 5,595 gages with the final stream mask). Figure 7(c) summarizes the performance across the three test cases grouped by drainage area. As can be seen here the final channel network improves performance gages across all drainage areas. Performance remains most variable in the smallest drainage basins (< 100 km^2). This makes sense as these are the locations where the DEM resolution will have the largest impact and where the stream mask will have the least impact (because 5th order streams do not extend into small headwater basins generally). Certainly, the performance of small drainage basins could be further improved with additional manual corrections to the stream network. The focus here though is on larger basins which will be most relevant for 1km resolution hydrologic simulations.

The 250m domain is not plotted here but is included in the processed datasets. Similar analysis of the 250m DEM shows that, as would be expected, a higher resolution DEM provides better information for drainage network derivation. There are 6,048 gages have area agreements within 10% using the unmodified 3rd order stream mask directly from NWM. For reference this is 982 more gages than the 1km DEM with the unmodified stream mask. Similar improvements have been found by applying all the processes to 250m resolution. Here too the processing does not significantly improve performance for gages less than 100km^2. Area agreement is enhanced vastly for gages over 500km^2 that the number of gages with area agreements over 50% decreases from 377 to 255.

https://doi.org/10.5194/essd-2020-291
Preprint. Discussion started: 21 November 2020

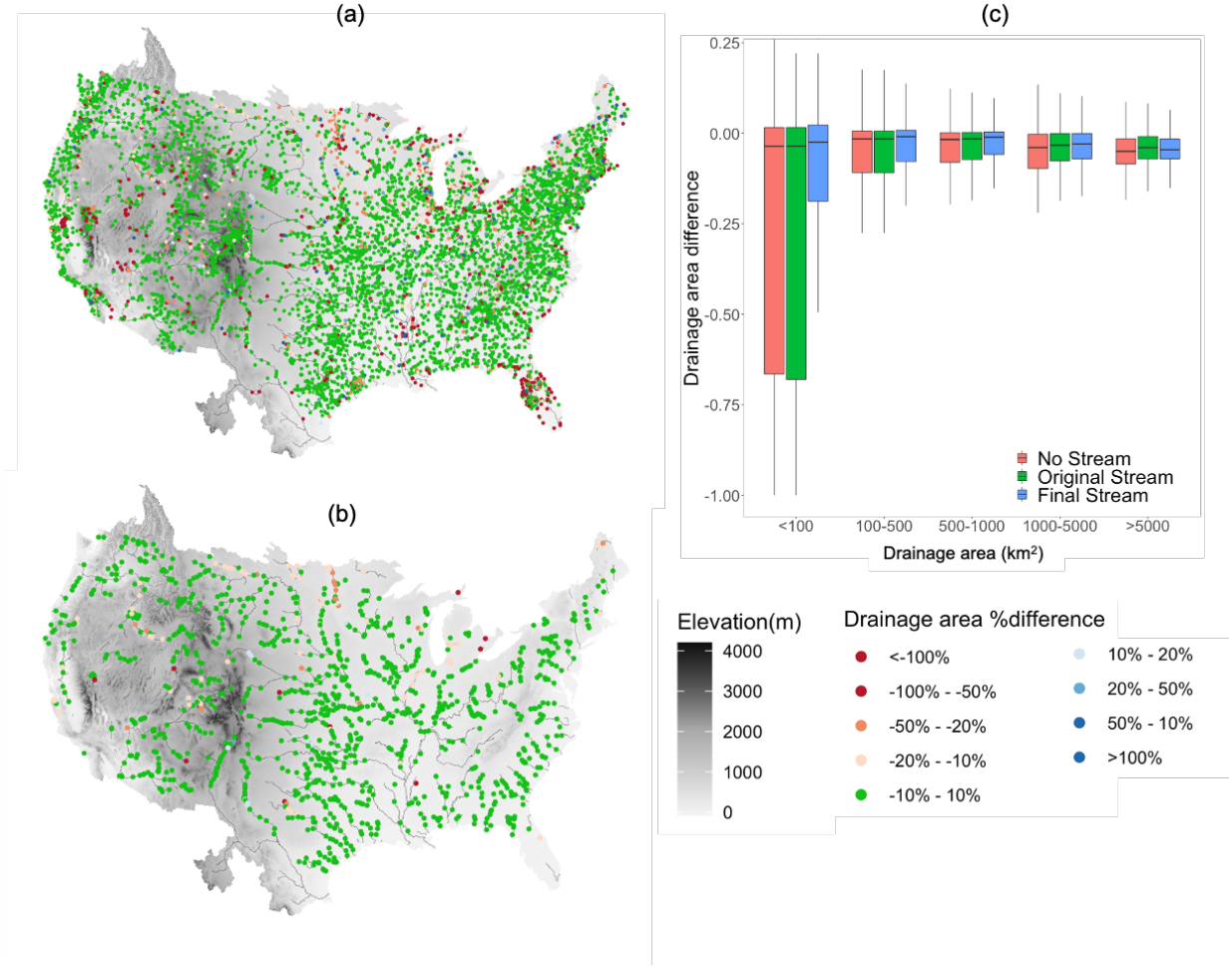


Figure 7: Drainage area comparison between processed with final order stream network and USGS gage measurements. (a) all gages for comparison; (b) gages with drainage area over 5000 km². (c) the boxplot of drainage area difference with different stream networks grouped by drainage area size.

3.2 Drainage and overland flow performance

In addition to the drainage network location we evaluate how the topographic processing influences the runoff characteristics of the domain. As described in Section 2.3, this behavior is evaluated using runoff tests and assessing anomalously high ponding depths. Here we consider the impact of smoothing along the stream (Step 4 in the topographic processing), the flat fix step applied to the rest of the domain (Step 5c in the topographic processing) and removing the secondary slope along the stream cells (Step 5b in the topographic processing). For reference we compare four cases listed below with progressively more processing steps applied,

(a) *No Smooth:* slope calculated by adjusted DEM from Priority Flood algorithm without any stream smoothing or flat fixing;

https://doi.org/10.5194/essd-2020-291
Preprint. Discussion started: 21 November 2020

(b) *Add Stream Smooth: No Smooth* case with *stream* smoothing added;

(c) *Add Fix Flat: Add Stream Smooth* case with the flat fix applied to the domain;

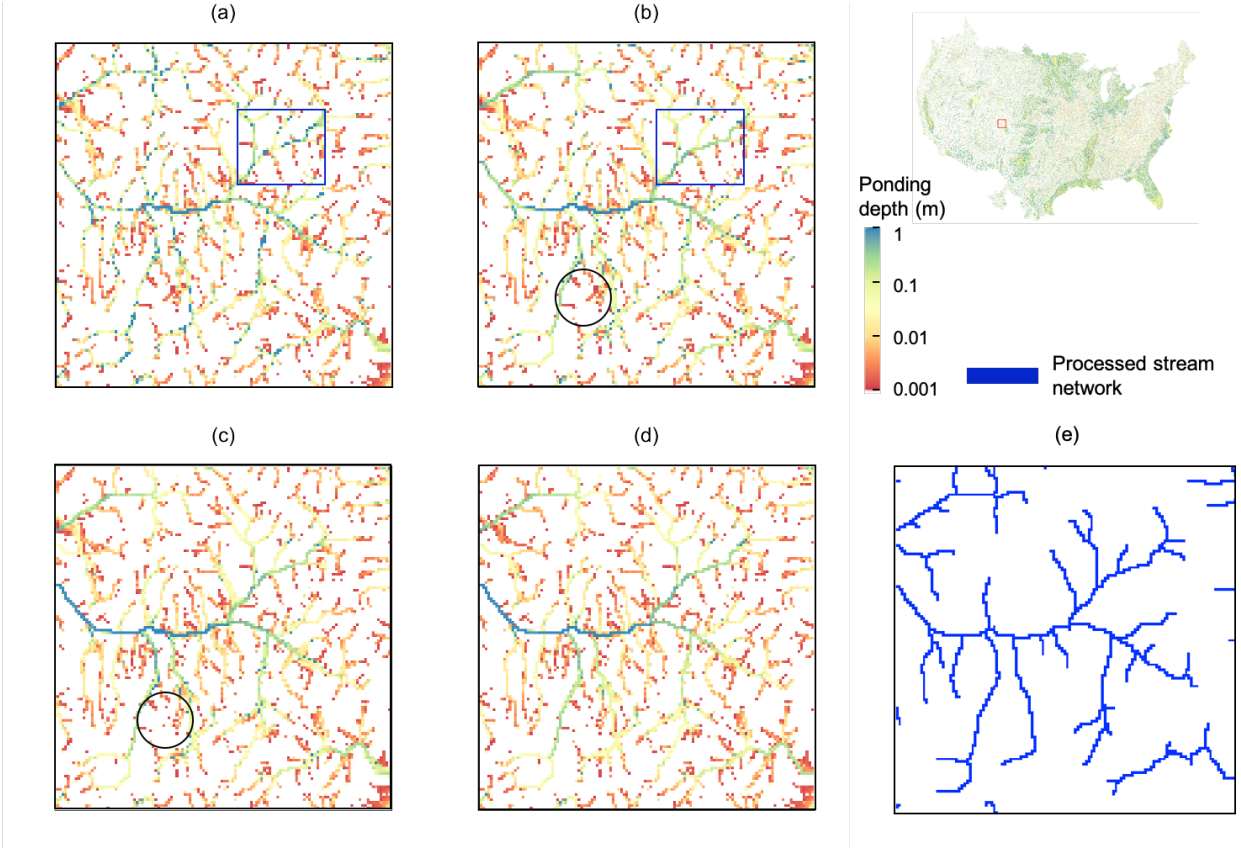
(d) *Remove Secondary: Add Fix Flat* case with removed secondary slope along the stream cells.

Note that in all of these cases processing steps 1-3 in the topographic processing remain the same, therefore the flow directions and location of the drainage network remain unchanged. The smoothing steps evaluated here do not alter the primary flow directions, they simply adjust the gradients along flow paths.

Ponding depths from a small domain located in Colorado at hour 20 of the runoff test are displayed in Figure 8 to illustrate the local impact of smoothing. The domain is 16,254 km^2 (126km by 129km) and the elevation ranges from 1,601 m to 4,054 m. The expected stream network based on drainage area thresholds is plotted Figure 8e and additional stream smoothing area applied along the stream network.

As would be expected the resulting stream network shape is the same, the largest ponding depths occur along the main stem of the drainage network in all cases. What is notable in this figure though are the spatial differences in how this ponding occurs. In the *No Smoothing* case there are discontinuities along the main stem of the stream as well as and many localized high ponding points across the domain (Figure 8a). Applying the stream smoothing in the three subsequent cases smooths these discontinuities along the main drainage network where smoothing was applied. An example of this improvements is show in the blue box of the domain in Figure 8a and b. Adding the flat-fix step decreases the number of isolated ponding points outside the stream network (an example is shown in the black circle in Figure 8b and c). Note that even with this smoothing step applied there are still isolated ponding locations. This is acceptable as all cells still have the ability to drain even it is slow and because our goal is not to achieve uniformly fast drainage everywhere. Rather the flat fix step is intended only to address those locations where large anomalies in the processed DEM result in significant ponding which may not be physically realistic. Finally, the *removed secondary slope along stream step* does not change the smoothing properties of the domain it simply removes the secondary slope in the stream to force water in the stream drain only in the primary direction, which increases the drainage speed of stream cells.

https://doi.org/10.5194/essd-2020-291
Preprint. Discussion started: 21 November 2020



385 **Figure 8: the spatial distribution of ponding depth from runoff tests from four slope cases (a) No Smooth (b) Add Stream Smooth (c) Add Flat Fix (d)Remove secondary. (e) stream network from Priority Flood approach with drainage area over 100km^2**

To further illustrate the impact of the stream smoothing step Figure 9a plots the elevation along the main stem shown in Figure 9c before and after smoothing is applied. Segments are plotted in different colors in Figure 9c and separated by dashed lines in Figure 9a. As can be seen here, in both cases the stream path is monotonically decreasing, however before
390 stream smoothing is applied there are many large steps along the stream path. With smoothing constant slopes are applied along each stream segment by adjusting the DEM along the stream. Note that the same elevation smoothing along the stream is carried out in the three smoothed test cases. Figure 9b compares the ponding depth along the stream for all four cases. Comparing first the *No Smooth* to the *Add Stream Smooth* it can be see that this step has the largest impact in the stream, significantly reducing the ponding depth variability from cell to cell. Adding the flat-fix increases the ponding depth
395 especially shown at the downstream. This is because the flat-fix step increases the drainage speed of the domain as a whole resulting in larger ponding depth in the stream. Finally, removing the secondary slope along the stream results in the smoothest pressure variation along the stream network by draining all water only to the primary direction.

https://doi.org/10.5194/essd-2020-291
Preprint. Discussion started: 21 November 2020

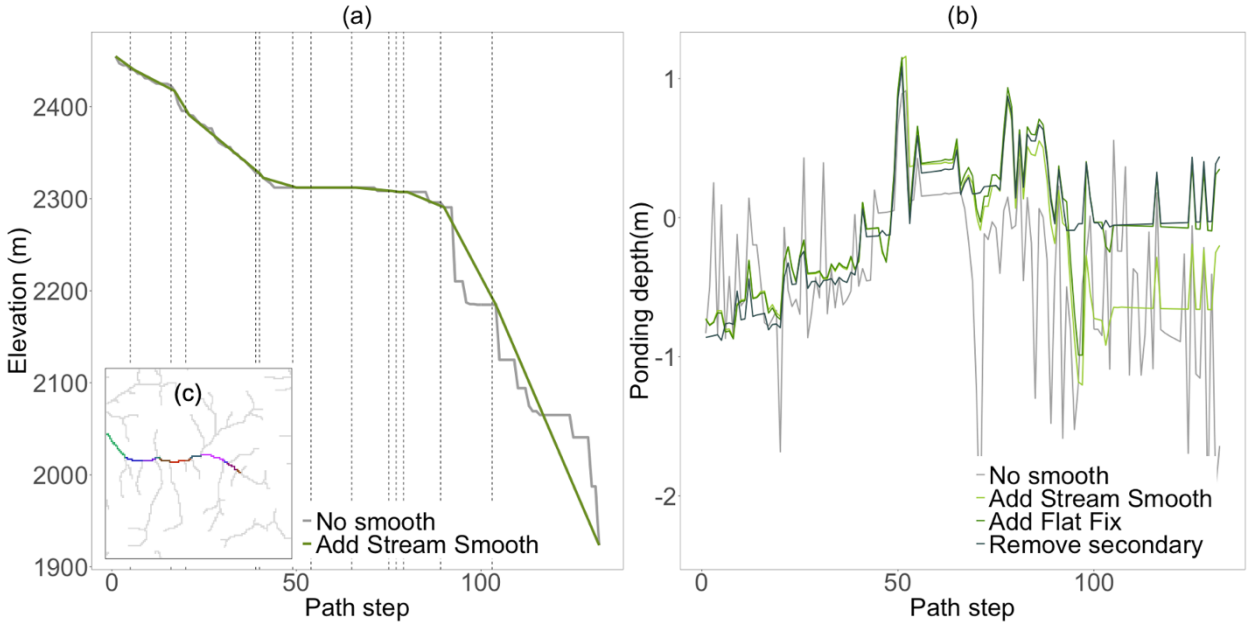


Figure 9: (a) the elevation along an example stream segment (shown in c) before and after smoothing applied along the stream divided into segments by dashed lines; (b) the ponding depth along the main stem from the runoff tests with four slope cases; (c) the main stem with segments plotted by different colors

The small domain shown in Figure 8 and Figure 9 illustrates the major behaviors seen across the domain in response to the smoothing steps. Figure 10 summarizes the results across the CONUS simulation. The background of Figure 10 shows the ponding depth from the runoff test with the final slope at hour 20 and the barplots summarize the percentage decrease in cells with ponding greater than 0.1 m outside the main stream network after 20 hours of the runoff simulation relative to the baseline, *No Smooth* case. The stream cells are excluded from this analysis as larger ponding depth is expected in the stream network. Results are summarized by major watershed outlined in black on the figure. Still, smoothing the stream increases the drainage rate across the domain so the *Add Stream Smooth* case does significantly decrease the number of ponding points outside the domain. The *Add Flat Fix* impacts a small number of cells in most basins, which is also to be expected given that it is targeting isolated discontinuities in the DEM. The flat fix step has the largest relative impact in steeper domains such as the upper Colorado due to the large variability of slopes between neighbor cells here. Removing the secondary slope has the largest impact in flatter portions of the domain such as the Mississippi Embayment.

https://doi.org/10.5194/essd-2020-291
Preprint. Discussion started: 21 November 2020

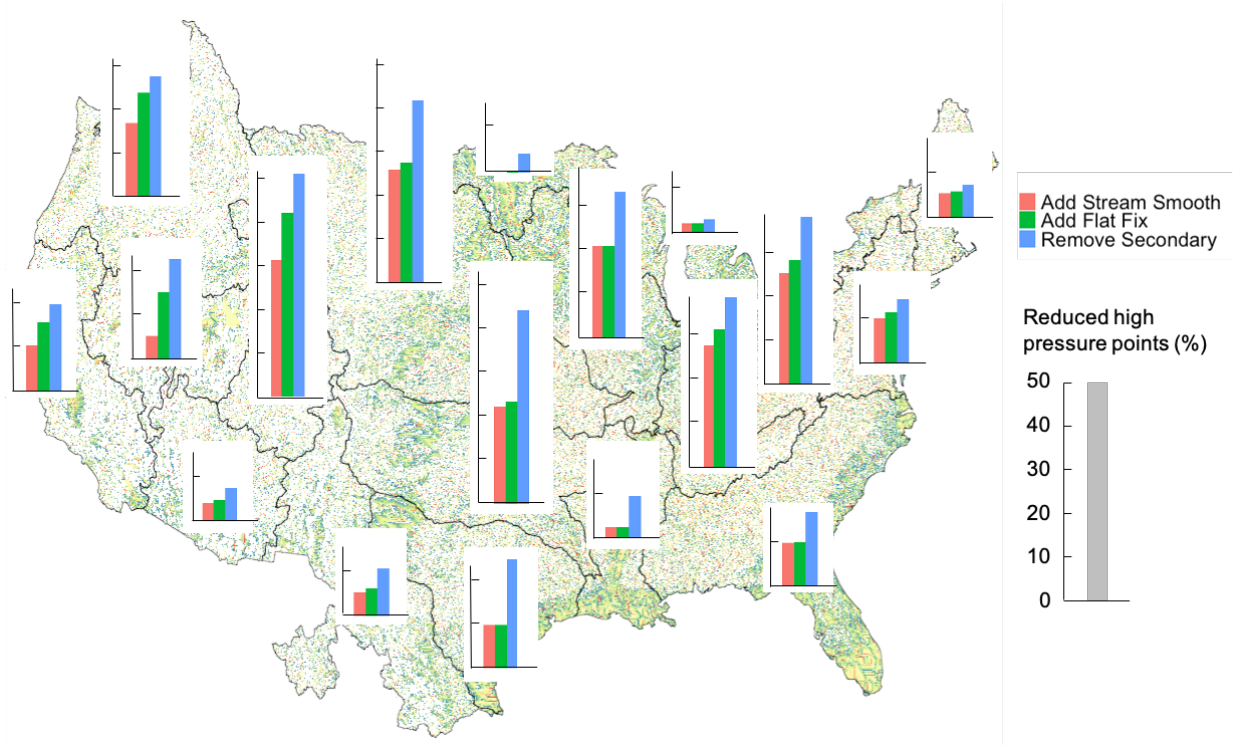


Figure 10: the background map is the ponding depth from runoff test at hour 20 with the final slope; barplots are the percentage decrease in cells with ponding depth over 0.1m over the main stream network relative to the baseline, No Smooth case

4 Final topographic dataset

Figure 11 shows the final elevation map, slope and drainage area that result from the topographic processing. These maps plot the 1km outputs, but all results are also available at 250 m resolution. Elevations and slopes are the primary inputs for the hydrologic modeling, but additional datasets defining the drainage networks and watersheds are also generated through the processing step (as shown by the Final outputs in Figure 2). While there are other national datasets that that derive drainage networks and watersheds, we include these outputs here too because they are on the same grid and derived through the same processing steps as the DEM, and therefore provide spatial information which is perfectly aligned with the final processed topography. The dataset includes: gridded map of elevation, slope in the x and y directions, primary flow directions, drainage area, stream segments, drainage basins associated with stream segments and distance to streams. Details on each of these output datasets are listed in Table 2.

https://doi.org/10.5194/essd-2020-291
Preprint. Discussion started: 21 November 2020

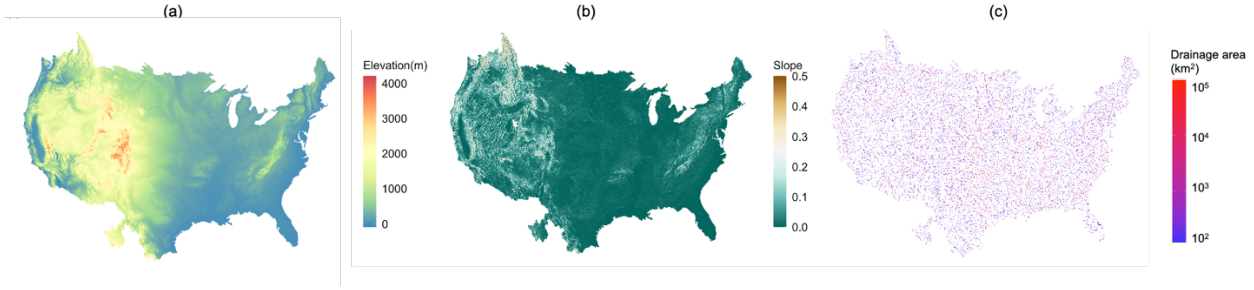


Figure 11: demonstration of output datasets over CONUS (a) processed DEM (b) Slope (c) Drainage area

Table 2: Summary of gridded outputs included in the dataset

Name	Unit	Description
DEM	m	Final DEM after all processing steps described here.
Flow direction	-	Primary flow direction for every grid cell in D4 directions (1=down, 2=left, 3=up, 4=right).
Drainage area	km^2	Drainage area calculated based on the primary flow directions.
Slope	-	Final slopes in the x and y directions calculated at cell faces.
Stream Mask	-	Mask with 1 for cells designated as 'streams' and 0 for everything else. Stream cells were specified as all cells with a drainage area greater than 100km^2.
Stream segments	-	Mask of stream segments with their segment IDs.
Subbasins	-	Map of the subbasins indicating the drainage area for each stream segment.
Stream Order		Strahler stream order calculated from the Stream Mask.
Distance to Stream	km	Map of distance to the stream mask for each cell.

Figure 12 summarizes the differences between the final resulting DEM presented here and the original DEM, including both the fraction of grid cells where the elevation changes and how much elevations were adjusted by. Adjustments split according to the point in the processing where they occurred (1) the priority flood adjustment step where elevations were adjusted to ensure complete drainage and (2) the subsequent smoothing steps. Overall the priority flood algorithm changed the elevation in 11.6% of the grid cells while the stream smoothing steps impacted 17.4% of the cells. In both cases the median slope adjustment magnitude is small (1.74m for Priority Flood and 3.12m for Smoothing). The Priority Flood algorithm frequently results in larger elevation adjustments than the stream smoothing step. Note also in Figure 12b that the priority flood algorithm only increases elevations from the original to the processed DEMs whereas the stream smoothing step can both increase and decrease elevations. Spatially the largest fraction of grid cells is adjusted in the flatter basins, for example the Mississippi Embayment (HUC 8), while the largest elevation adjustments occur in the steeper western portions of the domain such as Colorado and California (HUCS 13-18).

https://doi.org/10.5194/essd-2020-291
Preprint. Discussion started: 21 November 2020

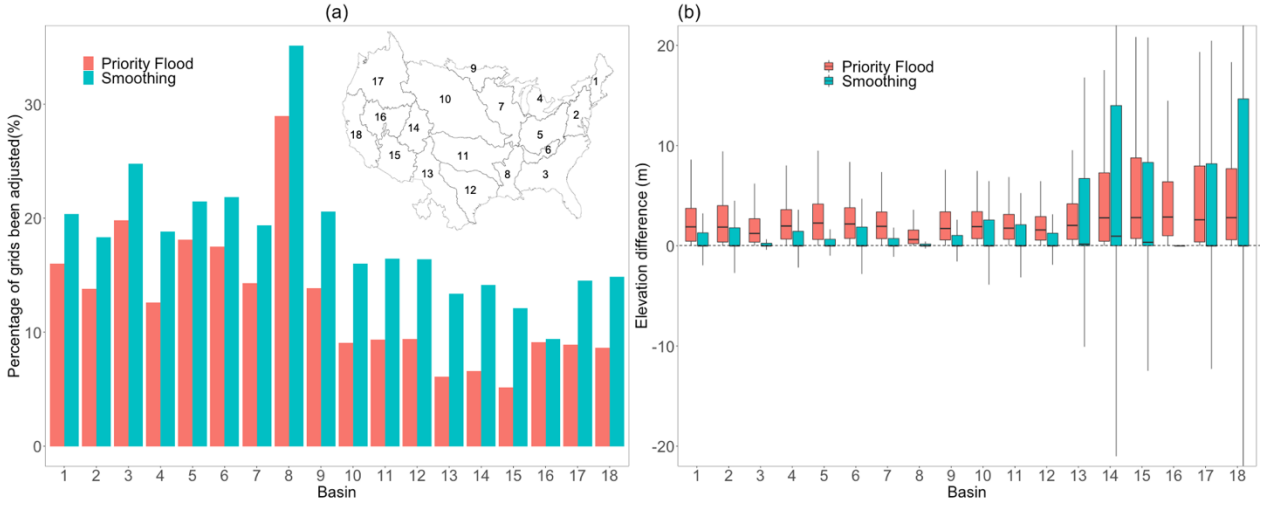


Figure 12: (a) the percentage of grids been processed in two processing parts in HUC2 basins; (b) the elevation difference after the two processing parts

5 Code and data Availability

All of the topographic processing was completed with the PriorityFlow R library which is available on GitHub (https://github.com/lecondon/PriorityFlow). All of the output datasets listed in Table 2 at 250m and 1km resolution are published through Cyverse (DOI link: https://doi.org/10.25739/e1ps-qy48) (Zhang and Condon, 2020). Datasets are available in tif, text and pfb formats. Along with these outputs, the input datasets they were generated from and the R scripts used for processing are also available. This will allow others to reproduce our work as well as generate their own versions with different processing settings if desired.

6 Conclusions

This study presents a high-resolution topographic dataset developed for physically based hydrologic simulations. We combine a Priority Flood topographic processing algorithm with a-priori stream network mask and USGS stream gage drainage areas to develop DEM and flow direction rasters which will closely match observed drainage networks. We also apply a series of slope smoothing steps along stream reaches and globally to improve surface runoff performance. These smoothing steps are unique from other global topographic smoothing approaches because we directly consider flow direction and stream locations to provide different smoothing along the stream reaches. The resulting DEM is designed to capture hydrologic features and improve runoff simulations for PDE based hydrologic models which require gridded topography. This performance is evaluated nationally using reported drainage areas from the USGS stream gage network and runoff simulations. All outputs from the processing are available across the contiguous US at both 1km and 250m resolution. The

https://doi.org/10.5194/essd-2020-291
Preprint. Discussion started: 21 November 2020

processing workflow developed here uses the open source R package PriorityFlow and is fully documented with the published datasets so that other can modify processing for different resolutions or specific sub domains as desired.

Acknowledgements

This work was supported by the U.S. Department of Energy Office of Science, Offices of Advanced Scientific Computing Research and Biological and Environmental Sciences IDEAS project and the Sustainable Systems Scientific Focus Area under Award Number DE-AC02-05CH11231 and the US National Science Foundation Office of Advanced Cyberinfrastructure under Award number OAC-1835855.

Author contribution

J.Z. and L.C. were in charge of processing the dataset including processing DEM and conducting the runoff simulations. J.Z. prepared the datasets, wrote the manuscript, plotted the figures and did the analysis. L.C. provided the PriorityFlow R package, documented the dataset and helped with the analysis and figures plotting. H.T. helped in evaluating the runoff simulations and reviewed the manuscript and documentation. R.M. provided overall supervision in processing the datasets and reviewed the manuscript.

Competing interests

The authors declare no competing interests.

References

Ashby, S. F. and Falgout, R. D.: A parallel multigrid preconditioned conjugate gradient algorithm for groundwater flow simulations, Nucl. Sci. Eng., 124(1), 145–159, 1996.

Barnes, M. L., Welty, C. and Miller, A. J.: Global topographic slope enforcement to ensure connectivity and drainage in an urban terrain, J. Hydrol. Eng., 21(4), 6015017, 2016.

Barnes, R., Lehman, C. and Mulla, D.: Priority-flood: An optimal depression-filling and watershed-labeling algorithm for digital elevation models, Comput. Geosci., 62, 117–127, 2014.

Condon, L. E. and Maxwell, R. M.: Evaluating the relationship between topography and groundwater using outputs from a continental-scale integrated hydrology model, Water Resour. Res., 51, 6602–6621, doi:10.1002/2015WR017200.A, 2015.

Condon, L. E. and Maxwell, R. M.: Modified priority flood and global slope enforcement algorithm for topographic processing in physically based hydrologic modeling applications, Comput. Geosci., 126(January), 73–83, doi:10.1016/J.CAGEO.2019.01.020, 2019.

https://doi.org/10.5194/essd-2020-291
Preprint. Discussion started: 21 November 2020

D'Odorico, P. and Rigon, R.: Hillslope and channel contributions to the hydrologic response, Water Resour. Res., 39(5), n/a-n/a, doi:10.1029/2002WR001708, 2003.

Daniels, M. H., Maxwell, R. M. and Chow, F. K.: Algorithm for flow direction enforcement using subgrid-scale stream location data, J. Hydrol. Eng., 16(8), 677–683, 2011.

David, C. H., Maidment, D. R., Niu, G. Y., Yang, Z. L., Habets, F. and Eijkhout, V.: River network routing on the NHDPlus dataset, J. Hydrometeorol., 12(5), 913–934, doi:10.1175/2011JHM1345.1, 2011.

Falcone, J. A.: GAGES-II: Geospatial attributes of gages for evaluating streamflow, US Geological Survey., 2011.

Freer, J., McDonnell, J. J., Beven, K. J., Peters, N. E., Burns, D. A., Hooper, R. P., Aulenbach, B. and Kendall, C.: The role of bedrock topography on subsurface storm flow, Water Resour. Res., 38(12), 1–5, 2002.

Frei, S., Lischeid, G. and Fleckenstein, J. H.: Effects of micro-topography on surface–subsurface exchange and runoff generation in a virtual riparian wetland—A modeling study, Adv. Water Resour., 33(11), 1388–1401, 2010.

Gallant, J.: Adaptive smoothing for noisy DEMs, Geomorphometry 2011, 7–9, 2011.

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 Resour. Res., 55(10), 7983–8009, doi:10.1029/2019WR024837, 2019.

Gesch, D., Oimoen, M., Greenlee, S., Nelson, C., Steuck, M. and Tyler, D.: The national elevation dataset, Photogramm. Eng. Remote Sensing, 68(1), 5–32, 2002.

Gochis, D. J., Dugger, A., Barlage, M., Fitzgerald, K., Karsten, L., McAllister, M., McCreight, J., Mills, J., Rafieeinasab, A. and Read, L.: The NCAR WRF-Hydro Modeling System Technical Description, 2018.

Gupta, V. K. and Mesa, O. J.: Runoff generation and hydrologic response via channel network geomorphology - Recent progress and open problems, J. Hydrol., 102(1–4), 3–28, doi:10.1016/0022-1694(88)90089-3, 1988.

Habtezion, N., Tahmasebi Nasab, M. and Chu, X.: How does DEM resolution affect microtopographic characteristics, hydrologic connectivity, and modelling of hydrologic processes?, Hydrol. Process., 30(25), 4870–4892, 2016.

Horton, R. E.: Erosional development of streams and their drainage basins; hydrophysical approach to quantitative morphology, Geol. Soc. Am. Bull., 56(3), 275–370, 1945.

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, Nat. Hazards Earth Syst. Sci., 19(11), 2405–2420, 2019.

Jones, J. E. and Woodward, C. S.: Newton–Krylov-multigrid solvers for large-scale, highly heterogeneous, variably saturated flow problems, Adv. Water Resour., 24(7), 763–774, 2001.

Kenny, F., Matthews, B. and Todd, K.: Routing overland flow through sinks and flats in interpolated raster terrain surfaces, Comput. Geosci., 34(11), 1417–1430, 2008.

Kollet, S. J. and Maxwell, R. M.: Integrated surface–groundwater flow modeling: A free-surface overland flow boundary condition in a parallel groundwater flow model, Adv. Water Resour., 29(7), 945–958, 2006.

https://doi.org/10.5194/essd-2020-291
Preprint. Discussion started: 21 November 2020

Kuffour, B. N. O., Engdahl, N. B., Woodward, C. S., Condon, L. E., Kollet, S. and Maxwell, R. M.: Simulating coupled surface–subsurface flows with ParFlow v3. 5.0: capabilities, applications, and ongoing development of an open-source, massively parallel, integrated hydrologic model, Geosci. Model Dev., 13(3), 2020.

Lindsay, J. B.: Efficient hybrid breaching-filling sink removal methods for flow path enforcement in digital elevation models, Hydrol. Process., 30(6), 846–857, 2016a.

Lindsay, J. B.: The practice of DEM stream burning revisited, Earth Surf. Process. Landforms, 41(5), 658–668, doi:10.1002/esp.3888, 2016b.

Lindsay, J. B., Francioni, A. and Cockburn, J. M. H.: LiDAR DEM smoothing and the preservation of drainage features, Remote Sens., 11(16), 17–19, doi:10.3390/rs11161926, 2019.

Liu, Y.-H., Zhang, W.-C. and Xu, J.-W.: Another fast and simple dem depression-filling algorithm based on priority queue structure, Atmos. Ocean. Sci. Lett., 2(4), 214–219, 2009.

Maxwell, R. M.: A terrain-following grid transform and preconditioner for parallel, large-scale, integrated hydrologic modeling, Adv. Water Resour., 53, 109–117, 2013.

Model v1.2, N. W.: Fulldom_hires_netcdf_250m.nc.Gz. National Oceanic (NOAA)., and Atmospheric Administration, 2018.

Moore, R. B., McKay, L. D., Rea, A. H., Bondelid, T. R., Price, C. V, Dewald, T. G. and Johnston, C. M.: User's guide for the national hydrography dataset plus (NHDPlus) high resolution, Open-File Rep., 80, doi:10.3133/ofr20191096, 2019.

Moretti, G. and Orlandini, S.: Hydrography-driven coarsening of grid digital elevation models, Water Resour. Res., 54(5), 3654–3672, 2018.

Samu, N. M.: Spatial Discrepancies between NHDPlus and LIDAR-Derived Stream Networks, University of Tennessee. [online] Available from: http://trace.tennessee.edu/utk_gradthes/1202, 2012.

Smith, V. B., David, C. H., Cardenas, M. B. and Yang, Z. L.: Climate, river network, and vegetation cover relationships across a climate gradient and their potential for predicting effects of decadal-scale climate change, J. Hydrol., 488, 101–109, doi:10.1016/j.jhydrol.2013.02.050, 2013.

Soille, P. and Gratin, C.: An efficient algorithm for drainage network extraction on DEMs, J. Vis. Commun. Image Represent., 5(2), 181–189, 1994.

Sørensen, R. and Seibert, J.: Effects of DEM resolution on the calculation of topographical indices: TWI and its components, J. Hydrol., 347(1–2), 79–89, doi:10.1016/j.jhydrol.2007.09.001, 2007.

Strahler, A. N.: Quantitative analysis of watershed geomorphology, Eos, Trans. Am. Geophys. Union, 38(6), 913–920, 1957.

Survey, U. S. G.: National Hydrography Dataset Plus High Resolution (NHDPlus HR) - USGS National Map Downloadable Data Collection., 2017.

Tarboton, D. G.: Terrain analysis using digital elevation models (TauDEM), Utah State Univ. Logan, 2005.

Thompson, J. A., Bell, J. C. and Butler, C. A.: Digital elevation model resolution: Effects on terrain attribute calculation and quantitative soil-landscape modeling, Geoderma, 100(1–2), 67–89, doi:10.1016/S0016-7061(00)00081-1, 2001.

https://doi.org/10.5194/essd-2020-291
Preprint. Discussion started: 21 November 2020

Vaze, J., Teng, J. and Spencer, G.: Impact of DEM accuracy and resolution on topographic indices, Environ. Model. Softw., 25(10), 1086–1098, doi:10.1016/j.envsoft.2010.03.014, 2010.

Wang, L. and Liu, H.: An efficient method for identifying and filling surface depressions in digital elevation models for hydrologic analysis and modelling, Int. J. Geogr. Inf. Sci., 20(2), 193–213, 2006.

Wolock, D. M. and McCabe, G. J.: Differences in topographic characteristics computed from 100- and 1000-m resolution digital elevation model data, Hydrol. Process., 14(6), 987–1002, doi:10.1002/(SICI)1099-1085(20000430)14:6<987::AID-HYP980>3.0.CO;2-A, 2000.

Woodrow, K., Lindsay, J. B. and Berg, A. A.: Evaluating DEM conditioning techniques, elevation source data, and grid resolution for field-scale hydrological parameter extraction, J. Hydrol., 540, 1022–1029, doi:10.1016/j.jhydrol.2016.07.018, 2016.

Wu, S., Li, J. and Huang, G. H.: A study on DEM-derived primary topographic attributes for hydrologic applications: Sensitivity to elevation data resolution, Appl. Geogr., 28(3), 210–223, doi:10.1016/j.apgeog.2008.02.006, 2008.

Zhang, J. and Condon, L. E.: JZhang_LCondon_CONUS_Topography_Sep2020, , doi:https://doi.org/10.25739/e1ps-qy48, 2020.

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, J. Am. Water Resour. Assoc., 54(4), 820–833, doi:10.1111/1752-1688.12623, 2018.

Zhang, W. and Montgomery, D. R.: Digital elevation model grid size, landscape representation, and hydrologic simulations, Water Resour. Res., 30(4), 1019–1028, 1994.

Zhou, G., Sun, Z. and Fu, S.: An efficient variant of the priority-flood algorithm for filling depressions in raster digital elevation models, Comput. Geosci., 90, 87–96, 2016.