TanDEM-X Hydro Technical Documentation

Last edited: January 31, 2023

Contents

1 Overview

Global and regional scale hydrologic models depend on the availability of stream networks with sufficient accuracy, attribution, and coverage. Existing global DEM-derived stream networks are limited by the resolution and accuracy of the source DEM, impacting the accuracy of stream placement and connectivity.

TanDEM-X Hydro (TDX-Hydro) is a new data suite which includes global streams, basins, and hydrologically-conditioned elevation data derived from 12m TanDEM-X data. The hydrologic conditioning process reduces the effects of vegetation and noise, uses existing datasets to encourage correct stream placement, and adjusts the surface to allow for continuous downslope flow. The accuracy of TDX-Hydro streams is expected to enable improvements for global-scale hydrologic modeling.

TDX-Hydro was produced by applying a custom hydrologic conditioning process to 12 m resolution TanDEM-X elevation data that had previously undergone an automated cleaning, coastline enforcement and void filling procedure. Global streams and basins were then derived from the conditioned DEM. The conditioning code is implemented in Python and was run on the Blue Waters supercomputer at University of Illinois Urbana Champagne.

The hydrologic conditioning process includes

1. vegetation bias removal
2. adaptive smoothing
3. river enforcement
4. depression handling

2 Contact Information

TDX-Hydro has been developed by the National Geospatial-Intelligence Agency, Office of Geomatics.

Contact email: SFNAGGeoscienceApplications@nga.mil

TDX-Hydro data are available for non-commercial and commercial use. For specific restrictions and use requirements see the License Agreement provided in Appendix A.

3 Data Access

The TDX-Hydro datasets can be accessed through NGA's Office of Geomatics website here: (`https://earth-info.nga.mil/`) under the "Geosciences" tab.

TDX-Hydro files are named with the following convention:

TDX_{dataset}_{hydrobasin ID number}_{version}.gpkg

4 Input datasets

The following datasets were used to create the hydrologically conditioned DEM, stream network and basins data sets:

1. TanDEM-X elevation data
2. Landsat Vegetation
3. Global Forest Canopy Height
4. ICESat-2
5. Surface Water Occurrence
6. Open Street Map (water layers)
7. HydroSHEDS HydroBASINS

4.1 TanDEM-X elevation data

The input DEM used as the "raw" elevation model in the hydro-conditioning process is a version of TanDEM-X that has previously undergone an automated cleaning, coastline enforcement and void filling procedure. The DEM is tiled consistently to one degree cells for the entire landmass independent of latitude.

Surface Type	Reflective
Resolution	nominally 12m
Horizontal Reference System	WGS84 Lat/Lon
Vertical Reference System	EGM08
Elevation Units	Meters
Null Value	-32767
Format	GEOTIFF
Bit depth	16 bit signed Integer
Classification	Unclassified//Limited Distribution

Table 1: Source TanDEM-X metadata

4.2 Tree Canopy Cover

The Landsat-derived Tree Canopy Cover (TCC) product was used to develop the global vegetation bias model for the TandDEM-X data set. The Landsat Vegetation Continuous Fields (VCF) tree cover layers contain estimates of the percentage of horizontal ground in each 30-m pixel covered by woody vegetation greater than 5 meters in height. The product is derived from all seven bands of Landsat-5 Thematic Mapper (TM) and/or Landsat-t Enhanced Thematic Mapper Plus (ETM+) [*Sexton et al.*, 2013].

More information can be found here:
https://lcluc.umd.edu/metadata/global-30m-landsat-tree-canopy-version-4

4.3 Global Forest Canopy Height

A 30-m spatial resolution global forest canopy height map was developed through the integration of the Global Ecosystem Dynamics Investigation (GEDI) lidar forest structure measurements and Landsat analysis-ready data time-series. The NASA GEDI is a space-borne lidar instrument operating onboard the International Space Station since April 2019. It provides footprint-based measurements of vegetation structure, including forest canopy height between 52° N and 52° S globally. [*Potapov et al.*, 2021]

More information can be found here: https://glad.umd.edu/dataset/gedi

4.4 ICESat-2

Two Ice, Cloud and land Elevation Satellite-2 (ICESat-2) datasets were used to create the database on which the global vegetation bias model was built:

1. ATLAS/ICESat-2 L3A Land and Vegetation Height (ATL08). The user guide can be found here: https://nsidc.org/data/ATL08
2. ATLAS/ICESat-2 L2A Global Geolocated Photon Data (ATL03). The user guide can be found here: https://nsidc.org/data/ATL03

4.5 Surface Water Occurrence

The Joint Research Centre (JRC) Global Surface Water Data was used to delineate water bodies within the DEM. The data, metadata, and users guide can be found here:
https://global-surface-water.appspot.com/download

4.6 Open Street Map

In combination with the JRC Surface Water Occurrence data, Open Street Map water layers were used to identify streams, rivers, lakes and other water features for the purposes of flattening large lakes and river segments, as well as burning in rivers and streams to improve the placement and connectivity of the DEM-derived stream network.

4.7 HydroBASINS

The HydroSHEDS project HydroBASINS dataset was used to both identify endorheic basins and to select appropriate processing regions in the hydrologic conditioning steps. [*Lehner and Grill*, 2013]. More information can be found here: https://www.hydrosheds.org/page/hydrobasins

5 Processing Workflow

Hydro-conditioning workflow:

1. Ingest and pre-process ancillary datasets
2. Vegetation bias removal
3. Adaptive smoothing (spurious noise removal)
4. River enforcement (stream burning)
5. Depression handling (pit removal)
6. Stream and basin generation

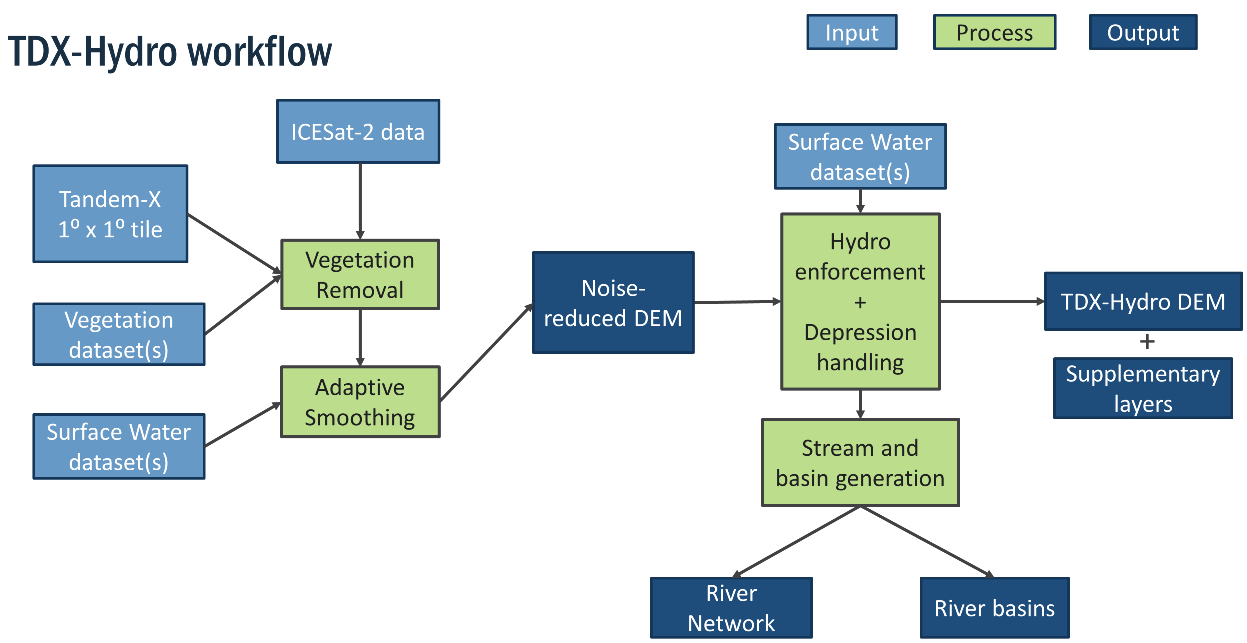

Figure 1: TDX-Hydro processing workflow

5.1 Pre-processing of input datasets

5.1.1 Landsat tree cover dataset

The tree cover dataset used here, described in 4.2, implemented the following void-filling workflow. The 2015 version of the TCC was chosen as the preferential version due to the temporal overlap with the collection of the TanDEM-X data. Gaps in the 2015 data were filled first with data from the 2010 dataset, then the 2005 dataset. Any remaining voids were spatially interpolated to create a continuous dataset to avoid propagating artifact edges through the vegetation model and into the DEM.

5.1.2 ICESat-2 Processing

We correlate ICESat-2's ATL03 and ATL08 datasets, tying the photon locations from ATL03 to their respective classifications (ground, canopy, etc) from ATL08. This allows us to leverage NASA's DRAGANN photon classification algorithm while re-sampling to higher along-track resolutions than the 100m segments of the ATL08 dataset. The derivation of the ATL08 dataset from the ATL03 dataset is described in *Neuenschwander and Pitts* [2019]. We used this process to create a database of over 5 billion co-located ICESat-2 ground elevations and TanDEM-X elevations. This database was used to build the vegetation bias models described below.

ICESat-2 pre-process workflow:

1. Import ATL03 and ATL08 for the same granule
2. Match photon locations from ATL03 to photon classification from ATL08
3. Create database of high confidence ground, canopy and top elevations and locations
4. Pull overlapping TanDEM-X points and interpolate ICESat-2 photons to create database of co-located ICESat-2, Landsat and TanDEM-X points

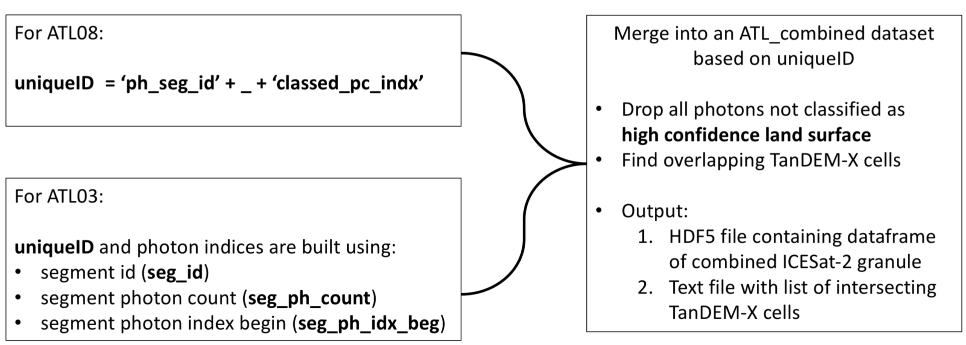

Figure 2: Creation of unique IDs on which to combine ATL03 and ATL08 ICESat-2 products

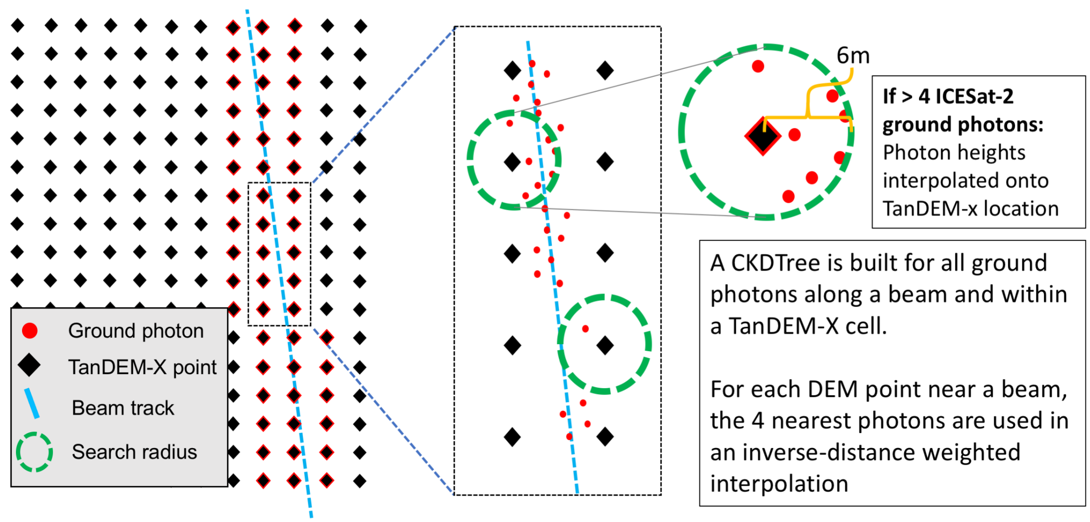

Figure 3: Interpolation of ICESat-2 ground photons

5.2 Vegetation Bias Removal

The vegetation bias model is built using ICESat-2 ground returns along with spatially continuous percent treecover, slope data, and for the 3-axis lookup table, treeheight.
Three methods were explored for a suitable global vegetation bias model:

1. An empirical polynomial equation
2. A 2-axis lookup table
3. A 3-axis lookup table *(currently implemented model)*

All methods use the parameters of local DEM slope and Landsat-derived treecover percent to estimate vegetation bias. The 3-axis table also uses the global tree height dataset.

5.2.1 Polynomial Model

The generalized form for our bivariate polynomial solution to compute vertical bias is of the form:

$$bias = C + s + t + s^2 + t^2 + I$$

where C is the intercept, s is local DEM slope in degrees, t is Landsat-derived treecover in percent, s^2 and t^2 are slope and treecover squared, and I is the interaction between s and t. To compute this polynomial solution, we used the correlated data described in 5.1.2.

5.2.2 Lookup Table Model

Both the 2-axis and 3-axis lookup tables were build using the same general method.

1. The bias measurements were clumped into each cell of the lookup table, and the mean bias value of any cell with more than 100 individual measurements was used directly.

2. For cells with fewer than 100 measurements, a value was interpolated.
3. A vegetation bias value of 0.0m was manually set for cells with a treecover or tree height value of zero.
4. A smoothing filter was applied along all axes of the lookup table to remove any artifact edges (especially in the interpolated values).

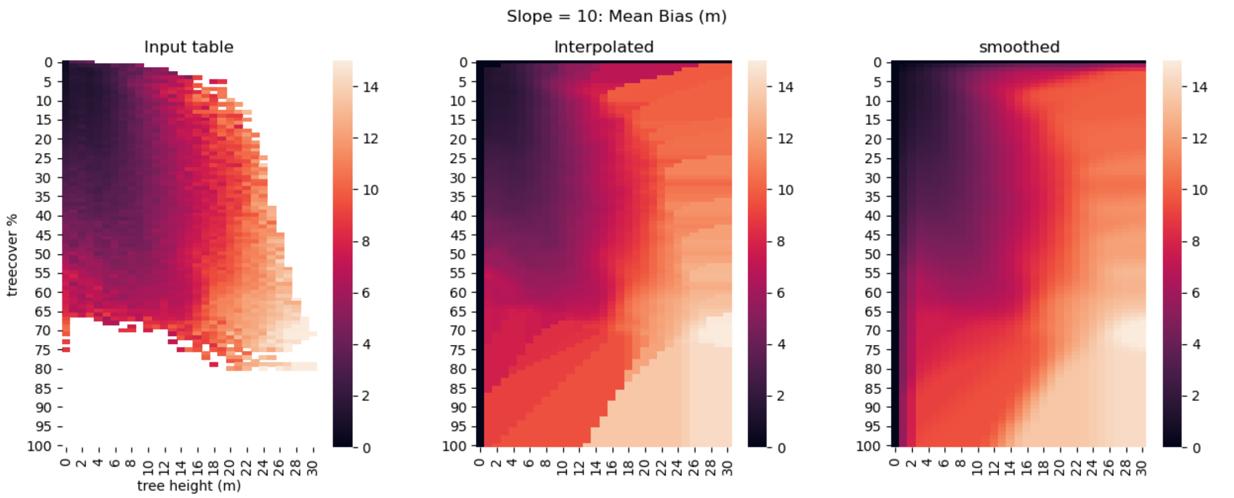

Figure 4: Slice of 3-axis lookup table. The raw mean value data (left panel) is interpolated (middle panel) and then smoothed (right panel) to create each slice of the lookup table.

5.2.3 Vegetation bias correction

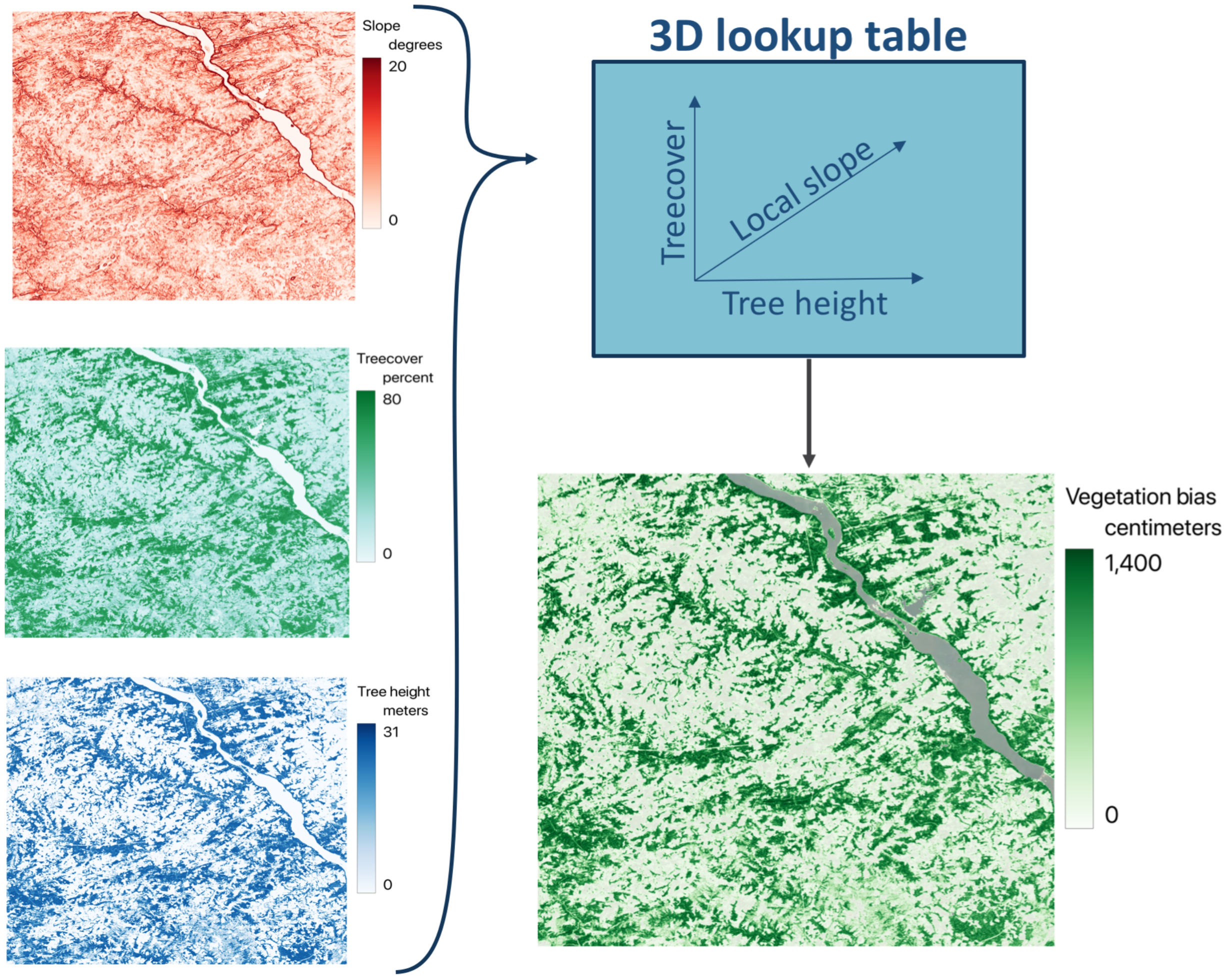

Figure 5: The three spatially continuous datasets are used to compute the modeled vegetation bias from the lookup table for each pixel within the DEM.

5.3 Adaptive Smoothing

The smoothing algorithm adapts the filter kernel size applied to each pixel depending on slope, curvature, and the presence of water. Low-slope pixels are smoothed with a wider kernel than high-slope pixels, and pixels that are identified as ridges, valleys, or water are held in place to preserve hydrologically significant features.

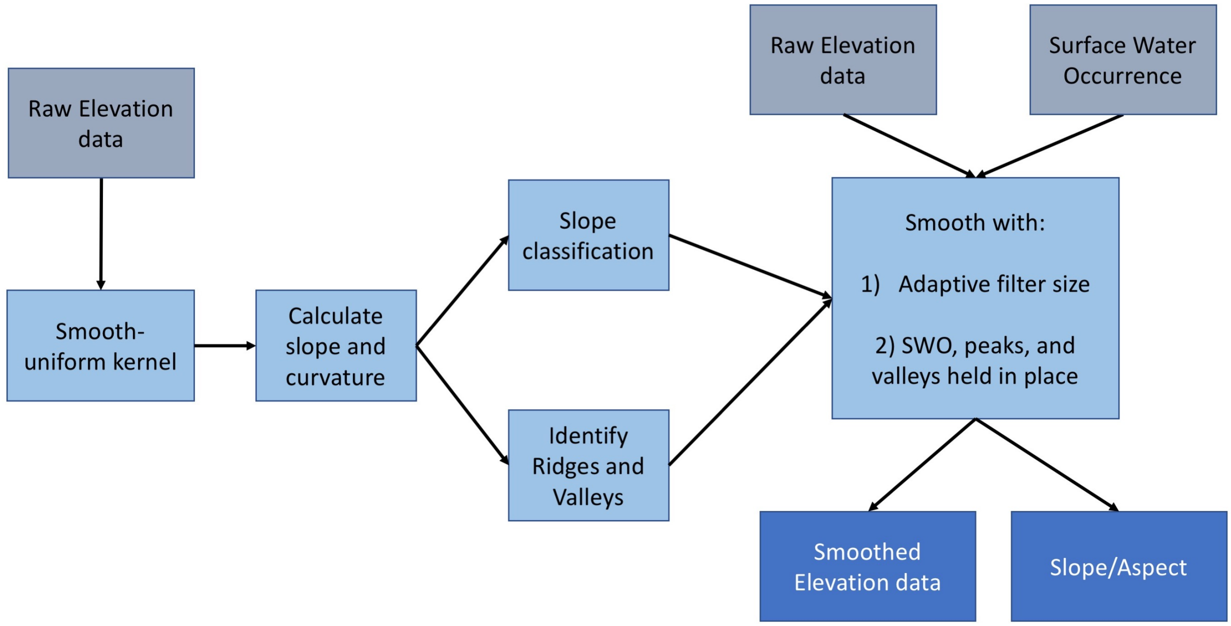

Figure 6: TDX-Hydro processing workflow for the smoothing portion of processing

Identification of valley and ridge features

A modified geometric curvature parameter is used to identify peaks and valleys within the DEM. This gradient-normalized curvature, compared to the Laplacian, tends to select narrower, more continuous features within the DEM and be well constrained to the center of ridges and drainage features [*Passalacqua et al.*, 2010]. However, this parameter becomes overly sensitive at very low slopes and involves handling division by zero errors. The modified version used here keeps the original terrain-adaptive capability while solving the latter two issues. For reference:

$$\text{Laplacian}: \nabla^2(z) = \nabla \cdot \nabla z$$

$$\text{Geometric}: \nabla_g^2(z) = \nabla \cdot \left(\frac{\nabla z}{|\nabla z|}\right)$$

$$\text{Mod. Geometric}: \nabla_{gmod}^2(z) = \nabla \cdot \left(\frac{\nabla z}{(|\nabla z| + c)}\right)$$

where z is the terrain height and c is a small constant (currently set to 0.1). Raising this constant lowers the sensitivity of the modified curvature at low slopes.

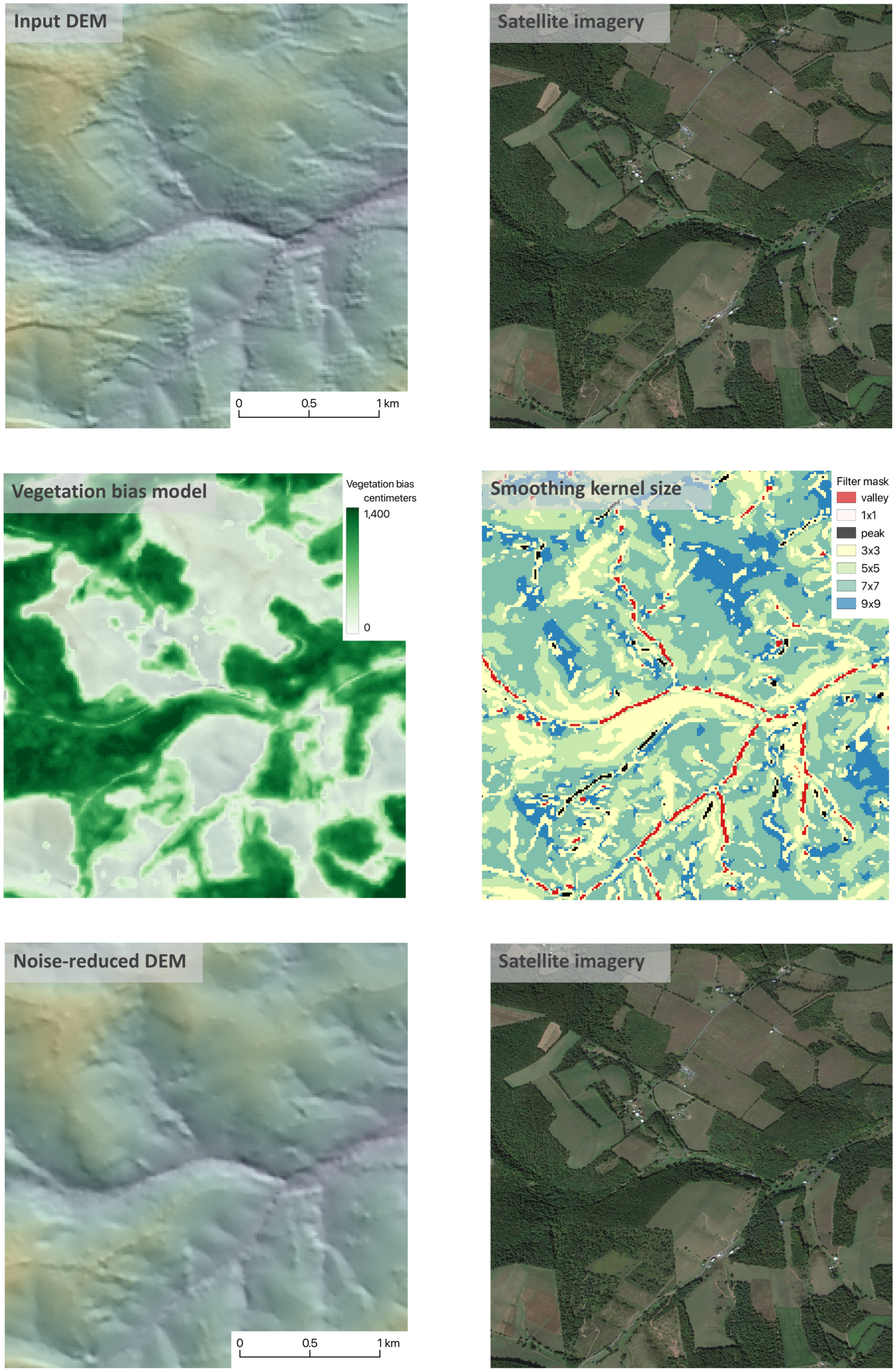

Figure 7: Example of noise and bias reductions applied. First the modeled vegetation bias is removed from the input DEM, and then an adaptive smoothing filter is applied.

5.4 River enforcement (stream burning)

The river enforcement step uses two datasets to encourage correct stream placement: the JRC Global Surface Water Occurrence layer, and Open Street Map water layers.

	Large Rivers and Reservoirs	**Streams**
Identified by	• Polygon features within OSM • Surface Water Occurrence (SWO)	• Line features within OSM
Categorized as	• Large river/reservoir/lake • Wetland (specifically excluded)	• Small river • Stream • Canal/drainage
Processed using	• Adjustment below adjacent land • Burning based on SWO value • Soft monotonicity enforcement	• Moving window minimum filter • Small in place burn (1-2 m) • Burn depth scaled to water type

Table 2: River enforcement methods

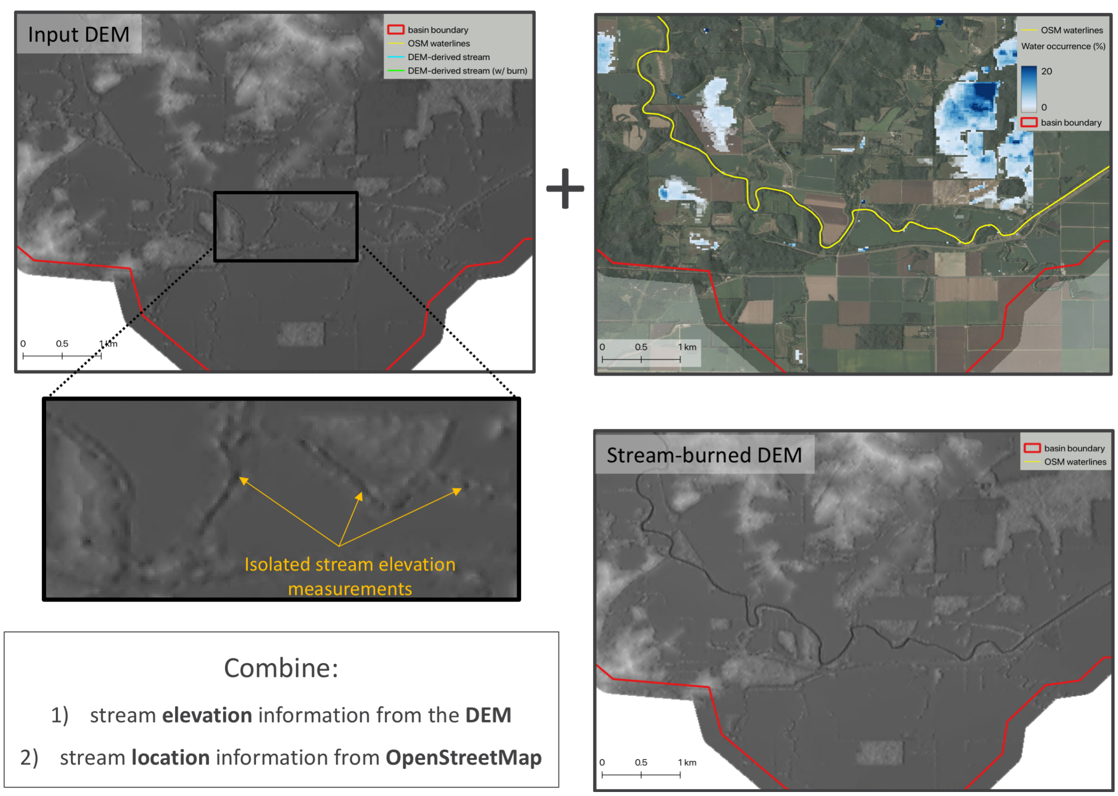

Figure 8: Example of stream being burned into the DEM by combining the disjointed elevation measurements within the DEM with the location information from Open Street Map. A moving minimum filter is applied along the stream channel to connect the elevations and create a more continuous drainage feature.

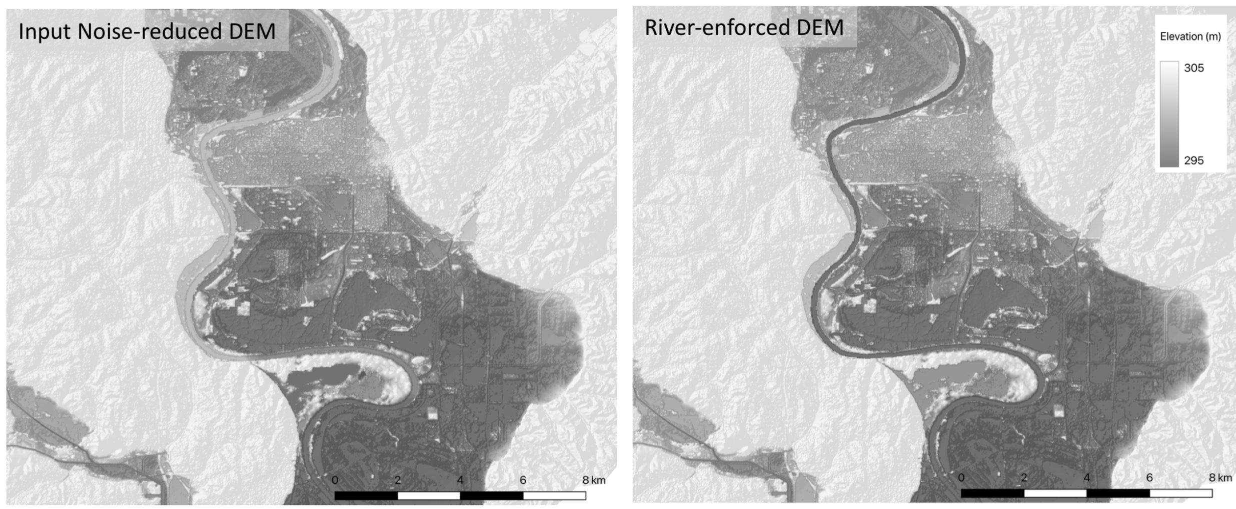

Figure 9: Example of a large river being lowered below surrounding land to ensure derived stream remains within the channel

5.5 Depression Handling

In the final conditioning step, each depression within the DEM is assigned a fill, carve, or intermediate solution based on a decision tree that considers solution geometries, interactions with downstream pits, and other factors. The depression handling algorithm is designed to solve depressions with minimal modification of the DEM.

To simplify the process of selecting solutions for each depression, pre-processing steps are applied to the DEM before analyzing the potential fill and carve solutions for all remaining depressions. Using WhiteboxTools, pits that can be solved by filling or carving a single pixel are resolved. For each remaining depression, a potential fill and carve solution is calculated using WhiteboxTools. Intermediate fill-carve solutions are also considered for large depressions. Details of each potential solution are catalogued, and the preferred solution is selected using the decision tree. For more information on depression handling parameters, see appendix A.1.

5.6 Endroheic Basin Handling

To ensure appropriate processing of endorheic basins, sink points are identified and set to nodata in the conditioned DEM. Any depression containing a sink point is left unsolved during the custom depression handling process described above. These sink-containing depressions are solved during the final pit filling step as part of the TauDEM workflow.

Sink points are identified for each level 12 HydroBASIN with the attribute ENDO = 2, denoting a terminal endorheic basin. Pixels with elevations less than or equal to the first percentile within the basin are selected and separated into contiguous areas. The largest contiguous area is isolated, and if it contains water pixels, only the area overlapping the water pixels is considered. Within this search area, the lowest elevation pixel is assigned as the sink. In the case of multiple pixels tied for the lowest elevation, the centroid of the minimum pixels within the search area is assigned as the sink.

5.7 Stream and Basin Delineation

The hydrologically conditioned DEM is processed using TauDEM software to generate stream networks with an upstream area threshold of 5 km^2, as well as their associated catchments. Stream networks and associated files were produced for each level 2 HydroBASINS polygon.

5.8 Quality Control Checks and Manual Edits

Each stream network was visually checked to identify potential issues. The TDX-Hydro streams were compared with HydroSHEDS streams and a Google Satellite imagery basemap. Major issues were resolved using semi-manual edit techniques:

1. **Noflow Walls**
   This optional subroutine allows a user to input a polygon shapefile designating areas that flow should not be allowed to pass through. This tool was generally applied in distributary environments when a river diverts to a minor channel rather than continuing along the more major channel. Pixels within the noflow wall polygon are raised during the depression handling step to block flow, and then are lowered to just above their original elevations (or a user-defined value) prior to TauDEM processing.

2. **Manual Burn**
   In cases where stream connectivity is interrupted by a feature such as canyon too narrow to be captured by TanDEM-X, it may be necessary to digitize the missing stream section and burn it in. This feature was rarely used during global TDX-Hydro processing, but was occasionally necessary.

3. **Parameter Adjustments**
   Some issues could be resolved by simply adjusting some of the processing parameters. Examples include increasing the maximum burn depth allowed in specific areas, reducing the buffer size used in basins bordering the date line, and determining whether to use Open Street Map data in distinguishing between ocean and land pixels.

6 Output datasets

6.1 Conditioned TanDEM-X DEM

The conditioned DEM has been processed through vegetation bias removal, smoothing, river enforcement, and depression handling. At its original 12 m resolution, the DEM is restricted by a Limited Distribution (LIMDIS) caveat.

6.2 Hydrographic datasets

A suite of hydrographic datasets are associated with the DEM, including flow direction and flow accumulation rasters, vector stream networks, and hydrologic catchments associated with each stream segment. These datasets are produced using TauDEM software; more information about their attribution is availalbe in the TauDEM documentation. A shapefile containing any endorheic sink points is available as well.

7 Known issues

1. **Water Elevations**
   The TDX-Hydro datasets have been built from TanDEM-X data that have undergone an automated process to remove spikes and wells, fill voids, flatten large water bodies, and enforce coastlines. While this process results in cleaner starting dataset, the water-flattening process can occasionally set elevations within a reservoir to that of the reservoir outlet. This moves the large jump in water elevation upstream, often to a bridge or other natural break.

2. **Basin Boundaries**
   Conditioned DEMs were processed through TauDEM for each level 2 basin polygon from the HydroBASINS dataset, with a 1 km buffer applied. Because HydroBASINS was built using a different, lower-resolution DEM, the boundaries do not perfectly align with TanDEM-X. This leads to some streams missing portions of their upstream area beyond the level 2 basin boundary, and others becoming truncated at the boundary. In most cases, only minor streams are affected. However, basins with particularly problematic borders have been processed in combination with neighboring basins to reduce these issues. Major basin boundaries will be redefined based on TanDEM-X for future versions of TDX-Hydro.

3. **Coastal Land Mask**
   TDX-Hydro relies on three datasets to distinguish between land and ocean pixels along coasts: 1) the Edit Data Mask (EDM) which was automatically generated during TDX pre-processing, 2) HydroBASINS polygons which were dissolved over the AOI and had any interior holes closed, and 3) Open Street Map water layers. A pixel was designated as ocean if it was both marked as ocean in the EDM and was outside the HydroBASINS polygon. In some cases, a third condition was applied in defining coasts: pixels which were marked as rivers in OSM were not designated as ocean. This condition prevented

premature truncation of streams in some cases, but often led to under-classification of ocean pixels, so it was not often applied.

This method prevented rivers from being clipped too far inland, but generally allowed streams to extend slightly past the true coast line. Some artifacts exist where streams form slightly offshore, and problematic coastal clipping can occur when the dissolved HydroBASINS polygon only partially covers a coastal bay.

4. **Date Line Crossings**
   The TDX-Hydro code is not currently able to handle processing basins that cross the date line. For the current version of TDX-Hydro, basins that cross the date line were given custom treatment. They were clipped back from the date line, and streams were allowed to flow to the date line as if it were coast. Affected level 2 HydroBASINS include: 5020055870, 3020009320, and 8020020760.

5. **Re-sampling artifacts**
   The TanDEM-X dataset is tiled to 1x1 degree tiles, and has jumps in longitudinal resolution at 50, 60, 70, and 80 degrees latitude to maintain a nominal 12m pixel size. Basins that cross one or more of these boundaries are re-sampled to eliminate any resolution jumps within a basin to allow hydrography datasets to be derived from the DEM. This re-sampling has two side effects:

   (a) Different resolutions among adjacent basins at higher latitudes. The resolution jumps are moved from the latitude boundaries to the hydrobasin boundaries.

   (b) Artifacts in the low-latitude region of a basin. The re-sampling method available for the level 2 basin mosaics (nearest neighbor) results in columns of pixels being 'eliminated' to achieve the uniform basin resolution. This is most notable in high latitude basins that span a large latitude range (i.e Japan, Iceland plus surrounding islands, Chile).

   More sophisticated re-sampling techniques and better selection of processing regions will be used for future versions.

6. **Artifact in source DEM** The TanDEM-X DEM contains an artifact running N-S along the western side of the prime meridian.

A Data license

B Parameter selection and sensitivity analysis

B.1 Depression Handling parameters

1. Basin Levels:

   The HydroBASINS dataset provides nested, hierarchical polygons of global hydrologic basins ranging from level 1 (coarse) to level 12 (detailed), all derived from 15 arc-second resolution raster data [*Lehner and Grill*, 2013]. The depression handling code first processes at a subbasin level (6-9 depending on polygon area), then mosaics these watersheds together and passes the mosaic through TauDEM processing. The basin level parameters can be changed by the user, depending on the scale of analysis being performed.

2. Fill Method:

   Within the WhiteboxTools software package utilized within the depression handling code, there are two separate fill methods that can be designated:

(a) The Standard Fill (SF)

(b) Wang and Liu (WL)

The Standard Fill method first identifies interior grid cells with no lower elevation neighbor, ultimately creating a pit catalog. Each pit is then addressed from highest to lowest elevation to help resolve interconnected depressions. For each individual pit, the algorithm systematically moves outward in a spiral direction through each of the surrounding neighbor cells until it eventually finds a low-lying neighbor cell that it labels as an outlet. The entire depression is then back filled to the same elevation as the outlet. A final flat fix can be applied across these uniformly flat areas to ensure continuous flow from each grid cell to the designated outlet.

The Wang and Liu algorithm examines each cell within the DEM based on its subsequent spill elevation, starting with the edge cells to locate an outlet, and utilizing a least cost search technique for optimal flow paths.

More detailed information can be found within the WhiteboxTools technical documentation [*Lindsay*, 2018].

The standard fill method with no gradient applied to flat regions was used for TDX-Hydro

3. Carve Method:

Within the WhiteboxTools software package utilized within the depression handling code, there are two separate carve methods that can be designated: The Standard Carve (SC) or Least Cost (LC).

The Standard Carve utilizes a breach first and fill behind approach to resolve continuous flow paths. As the name implies, the Least Cost approach uses a least-cost path analysis to identify the breach channel that connects pit cells to some distant lower cell. The cost of the various potential breach paths is determined by the amount of elevation lowering needed to cut the breach channel through the surrounding topography.

Although the Least Cost algorithm offers an optimized version of the Standard Carve, there are several limiting parameters that must be provided (see the following section). The Standard Carve also provides increased computational efficiency when processing such large regions.

More detailed information can be found within the WhiteboxTools technical documentation [*Lindsay*, 2018].

The Standard Carve option was utilized for TDX-Hydro

4. Least Cost Breach Parameters (If Applicable):

If the user decides to utilize the Least Cost breaching algorithm, then a max cost and search radius must be provided, along with an option to minimize the breach distances or not. Using the same Palouse watershed in Washington (which spans 6 total TDX cells), another sensitivity test was performed to determine the most appropriate values for each of these parameters. As shown in the figures below, the maximum cost parameter had little impact on both the number of pits detected and the number of incomplete carves remaining in the catalog. It is important to note that there is a trade off on computation efficiency for greater cost thresholds. Thus, the user should weigh those options before deciding on an appropriate max cost parameter. The radius, however, had a substantial impact on the same results, appearing to plateau at radius of approximately 25m.

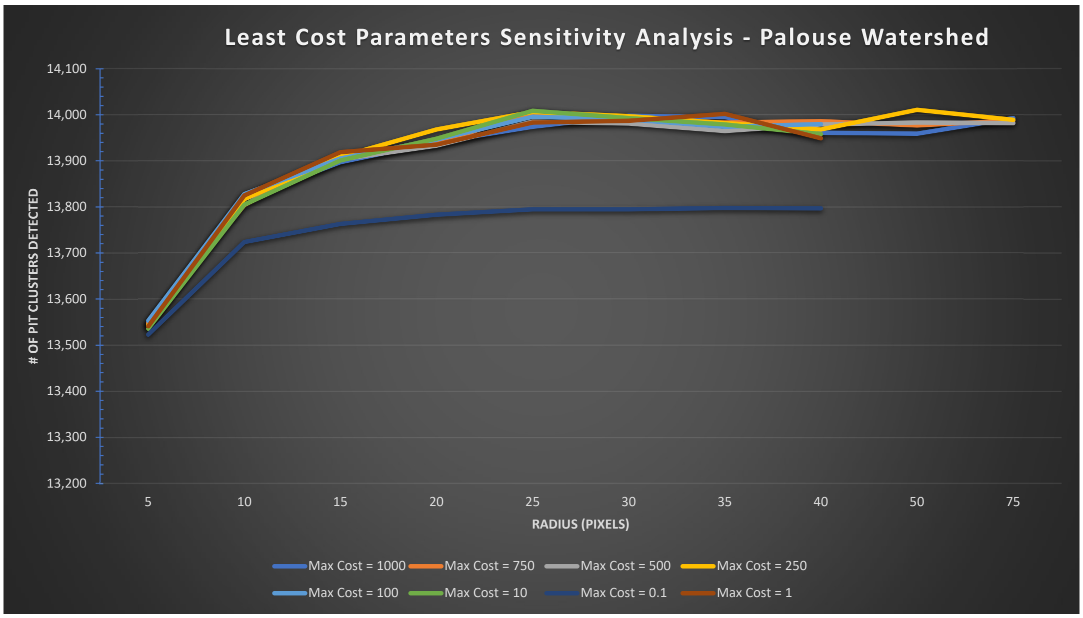

Figure 10: Least Cost Parameters Sensitivity Testing - Number of Pits Detected

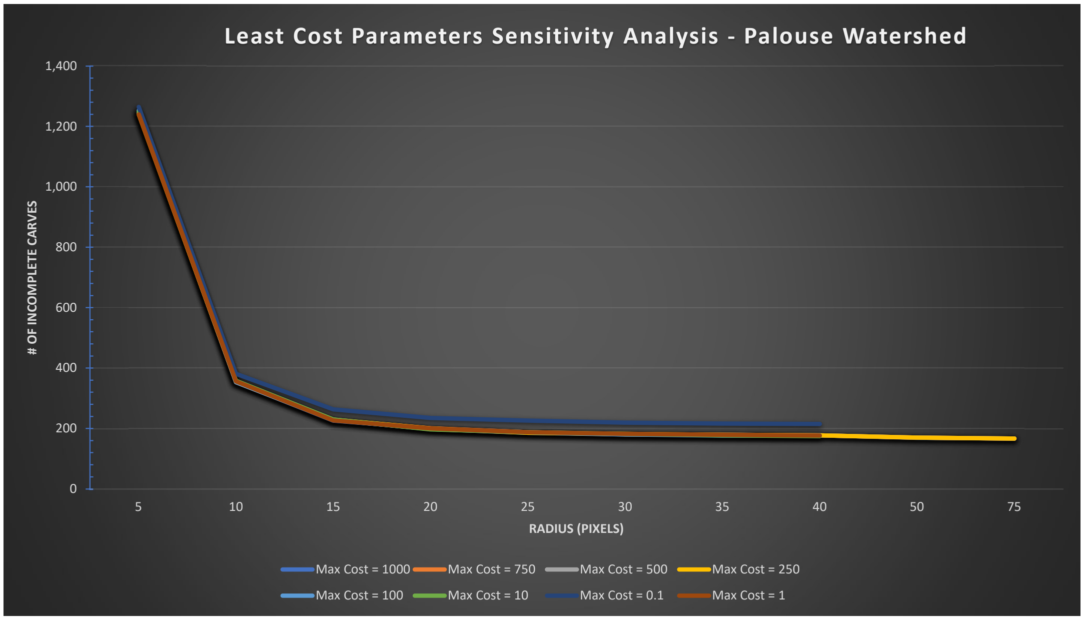

Figure 11: Least Cost Parameters Sensitivity Testing - Number of Incomplete Carves Remaining

If the minimize breach distance option is left unselected, the default setting, then the algorithm will rely on cost as opposed to distance, which is the original intent.

5. Solution Decision Tree Thresholds:

The depression handling algorithm includes a custom catalog for combined fill/carve solutions. This combination of approaches often results in more realistic hydrological solutions and less modification to the original DEM. Thus, a decision tree was incorporated with various fill and carve thresholds, a maximum carve depth, and a combined fill interval for possible solutions before an optimal one is selected.

Using the same case study as before (the Palouse watershed in Washington), both the various fill and carve thresholds are compared to the number of detected depressions within the basin that are solved by fill, carve, or combination therein.

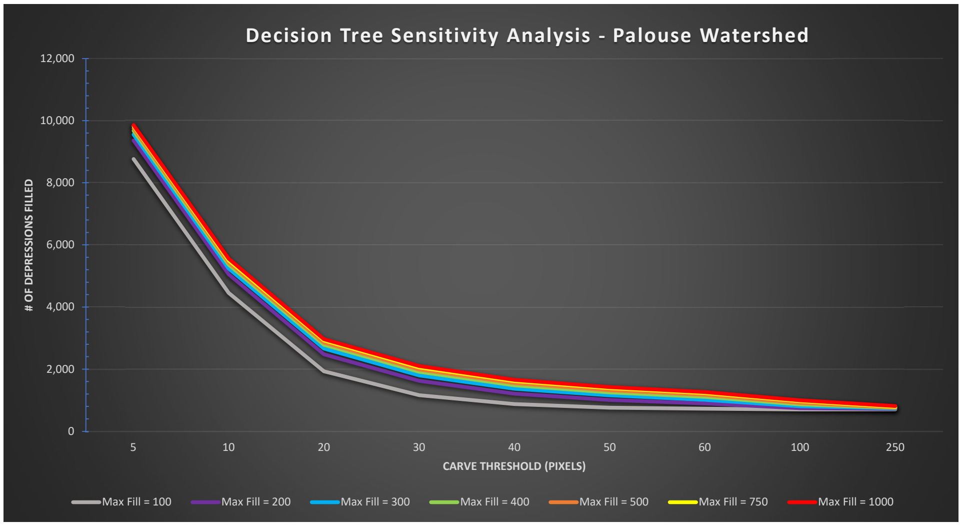

Figure 12: Decision Tree Sensitivity Testing - Fill Only Solutions

(a) Maximum Fill Volume:
The higher the fill volume threshold the more depressions will be solved by a fill-only approach. The lower thresholds of both fill and carve result in far more combined solutions therefore requiring far more computational time.

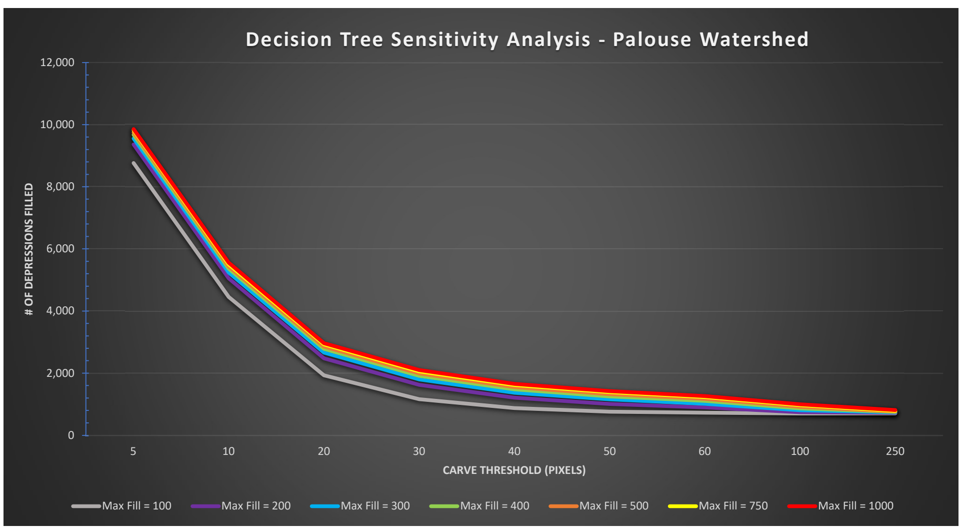

Figure 13: Decision Tree Sensitivity Testing - Maximum Fill Threshold

(b) Maximum Carve Depth:

The maximum carve depth was determined based on the Global Reservoir and Dam Database (GRanD v1.3) developed from a consortium of contributing institutions and routinely updated [*Lehner et al.*, 2011]. The database represents large scale reservoirs and associated dams (cumulative storage capacity of 6,197 km3), currently totaling 7,320 across the globe. After removing the database entries with inadequate or erroneous attribution, the following table was created to determine the appropriate carve threshold.

The maximum carve depth needs to be large enough to ensure that the downstream flow can carve through these infrastructure features while not allowing too much freedom to inaccurately carve through hills, cliffs, or other topographic features throughout the DEM. **A maximum carve depth of 100m was selected to capture nearly 98 percent of the dam database.** Any remaining dams will most likely be caught with the other built in catches developed as part of the depression handling code.

Height Above Reservoir (m)	Percentage of Database
25	56.65
50	85.49
75	94.63
100	97.92
125	99.05

Table 3: Dam Heights in GRanD Database

(c) Combined Fill Interval:

The combined fill interval represents the fractional step for partial solutions considered. The interval is determined for each individual pit within the custom catalog by taking the maximum elevation of an entirely fill solution minus the lowest elevation in the original DEM, and then multiplying by the user-supplied interval. Then, the algorithm subtracts the first interval from that entirely fill solution to preserve the flat fix already applied at the surface of the depression, and then resolves the carve option at that level. The algorithm continues to decrease each partial solution by the same interval, cataloguing the resultant fill and carve volumes for each potential solution. At the end, an optimal solution is selected that will result in the least modification to the original DEM.

A fill interval of 0.2 was selected to allow for comprehensive comparison of possible combined solutions while remaining computationally reasonable.

References

Lehner, B., and G. Grill, Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems, *Hydrological Processes*, *27*(15), 2171–2186, 2013.

Lehner, B., et al., Global reservoir and dam (grand) database, *Technical Documentation, Version*, *1*, 1–14, 2011.

Lindsay, J. B., Whiteboxtools user manual, *Geomorphometry and Hydrogeomatics Research Group, University of Guelph, Guelph, Canada*, *20*, 2018.

Neuenschwander, A., and K. Pitts, The atl08 land and vegetation product for the icesat-2 mission, *Remote sensing of environment*, *221*, 247–259, 2019.

Passalacqua, P., T. Do Trung, E. Foufoula-Georgiou, G. Sapiro, and W. E. Dietrich, A geometric framework for channel network extraction from lidar: Nonlinear diffusion and geodesic paths, *Journal of Geophysical Research: Earth Surface*, *115*(F1), 2010.

Potapov, P., et al., Mapping global forest canopy height through integration of gedi and landsat data, *Remote Sensing of Environment*, *253*, 112,165, doi:https://doi.org/10.1016/j.rse.2020.112165, 2021.

Sexton, J. O., et al., Global, 30-m resolution continuous fields of tree cover: Landsat-based rescaling of modis vegetation continuous fields with lidar-based estimates of error, *International Journal of Digital Earth*, *6*(5), 427–448, doi:10.1080/17538947.2013.786146, 2013.