Watershed and stream network delineation using digital elevation models and spectral satellite information

G. Amatulli, S. Domisch, and P. Raymond

1Yale School of Forestry & Environmental Studies, New Haven, 06511, CT 06511, USA.
2Leibniz-Institute of Freshwater Ecology and Inland Fisheries, Berlin, D-12489, Germany

Abstract

We use the Global Multi-resolution Terrain Elevation Data 2010 (GMTED, 7.5 arc-seconds) and Landsat Maps Global Surface Water (GSW, 1 arc-seconds) processed by GRASS-GIS hydrological modules to derive a fully standardized stream network. The GSW water occurrence was burned into the GMTED surface to force the drainage direction, thereby directing the stream network to pass through the observed water body. We run an iterative process to identify the best carving profile and relative depth by combining water occurrence and elevation roughness. The obtained results are intended to be implemented at the global level.

Keywords: Stream network, hydrology, digital elevation model.

1 Introduction

The location and structure of streams and rivers underpin a myriad of patterns and processes in hydrology, geomorphology, geography and ecology. The wide availability of digital elevation data and improvement in computational power have led to recent advances in terrain and hydrologic analyses using digital elevation models (DEMs) on the local and global scales. Extracting a stream network from DEMs is based on the computation of the upstream flow accumulation. It yields a potential analysis of geophysical features, but does not account for stream hydraulics or water availability. For the delineation of watershed and drainage networks, a large number of techniques and algorithms have been implemented (Tarboton et al., 1992; Montgomery and Foufoula-Georgiou, 1993; Heine et al., 2004; Pelletier, 2013). Current algorithms are based on the natural phenomena that water follows the steepest and shortest direction along a relief, and accumulates along valleys, lowlands, flat areas and depressions. In addition to DEMs, satellite images provide spatial high-resolution and global coverage of water presence (Pekel et al., 2016; Feng et al., 2016) that can be used to extract the locations of large water bodies, as well as the temporal variability in water presence (i.e., seasonality). These maps of water surface, even if fragmented in space, due to the minimum mapping unit of the satellite images, can be used to improve the location of stream networks derived from DEMs.

In this study, we use the Global Multi-resolution Terrain Elevation Data 2010 dataset (GMTED;
7.5 arc-seconds) (Danielson and Gesch, 2011) and Landsat Maps Global Surface Water (GSW; 1 arc-
seconds) (Pekel et al., 2016) processed by Geographic Resources Analysis Support System (GRASS)
open source software (GRASS Development Team, 2017) hydrological modules to derive a fully
standardized stream network for the conterminous United States. We selected the conterminous
United States due to the wide variation of landscapes, geomorphology and climate, as well as the
availability of the NHDplus dataset (David et al., 2011), which can be used for further validation
purposes.

2 Data and method

2.1 Source layers

The GMTED dataset was released in 2010 by the United States Geological Survey and National
Geospatial-Intelligence Agency globally at three spatial grains of 1 km (30 arc-seconds), 500m (15
arc-seconds) and 250m (7.5 arc-seconds), with the exception of Greenland and Antarctica, where
only 1-km data sets are available. The original GMTED product is a composite product based
on several gridded elevation data sources with spatial grains ranging from approximately 30m to
2 km (1 to 60 arc-seconds) (Danielson and Gesch, 2011). These single DEMs were merged using
various aggregation methods: minimum elevation, maximum elevation, mean elevation, median
elevation, standard deviation of elevation, systematic subsample, and breakline emphasis. In par-
ticular, breakline emphasis is especially useful for deriving hydrologic features because it retains
topographic breaklines (ridge lines and stream channels) as depicted underneath the full resolution
elevation data (Danielson and Gesch, 2008; Gesch et al., 1999).

The GSW dataset describes water presence/variation over the past 32 years, using more than 3
million Landsat satellite images (Pekel et al., 2016). The unique spectral signature of surface water
in the infrared and visible range allows for the differentiation between water and dry land surfaces
using image classification techniques, enabling the generation of a time series that documents water
intra/inter annual variability of any water body that can be detected with the 30m Landsat pixel
(Pekel et al., 2016). Apart from providing a promising data source to locate water bodies that
meets this threshold, GSW can also be used to delineate DEM-derived stream networks using water
presence as an indicator of valleys and depressions, and forcing the DEM-derived stream network
to pass through the observed water body. This method provides a valid alternative, especially in
flat areas where the DEM-derived stream network could be substantially different than the actual
stream network.

2.2 Stream network delineation methodology

The quality of DEMs influences the derived stream network, and even small errors in accuracy can
greatly affect the geographic location of the stream (Wilson and Gallant, 2000). This phenomenon
is more evident in naturally flat surfaces where such errors in accuracy are larger than the actual
relief variation. One way to address this limitation related to errors in the location of the DEM-
derived stream network is to use the "stream burning" (or carving) approach (Saunders, 1999).
This approach, introduced by Hutchinson (1989), proposes the use of ancillary information such as
already existing stream network products, to "carve" the DEM and force the flow to pass through
those cells that correspond to the actual stream network.
The GSW dataset depicts water occurrence in terms of a percentage of presence during the 32 years of observation. This information was burned into the GMTED surface to force the drainage direction, thereby directing the stream network to pass through the observed water bodies. The water bodies can be seen as the location where the DEM should converge the actual stream network in order to maximize the similarity between the actual and the DEM-derived location of the stream network. The burning operation can be modeled by the variation of three parameters: i) depth of burning, ii) water occurrence and, iii) elevation roughness. The three parameters can be combined in an index that incorporates the spatial variation as a function of the DEM variation and water occurrence. The burning index can be expressed by the following formula:

$$DBI_i = WO_i * D_i * NE\sigma_i$$  \hspace{1cm} \text{Equation 1}

In particular: $WO_i$ identifies the water occurrence parameter, ranging from 0 to 100; $D_i$ identifies the depth, ranging from 0m to 200m with 10m increments; $NE\sigma_i$ identifies the elevation roughness and is expressed at each pixel by calculating the normalized standard deviation of elevation in a circular mowing window of a 41 pixel radius, and ranges from 0 to 1.

In order to check the linear and non-linear influence of the $WO_i$ and the $NE\sigma_i$, their values were altered within the constant data range. The following combinations were tested:

$$DBI_i = (WO_i > 0,1) * D_i$$  \hspace{1cm} \text{Equation 2}

$$DBI_i = \log(WO_i + 1)/4.615121 * D_i$$  \hspace{1cm} \text{Equation 3}

$$DBI_i = \log(WO_i + 1)/4.615121 * D_i * NE\sigma_i$$  \hspace{1cm} \text{Equation 4}

$$DBI_i = \log(WO_i + 1)/4.615121 * D_i * (NE\sigma_i)^2$$  \hspace{1cm} \text{Equation 5}

$$DBI_i = \log(WO_i + 1)/4.615121 * D_i * (1 - NE\sigma_i)$$  \hspace{1cm} \text{Equation 6}

$$DBI_i = \log(WO_i + 1)/4.615121 * D_i * (1 - NE\sigma_i)^2$$  \hspace{1cm} \text{Equation 7}

The depth-burned index was then subtracted from the original GMTED ($GMTED_i - DBI_i$), creating "canyons" where the observed GWS water bodies were present. This carved GMTED was then further corrected by removing coarse sinks and peaks using the GRASS r.hydrodem-function. Finally, a new flow direction layer was created using the GRASS r.watershed-function, enabling water to flow into multiple down-stream cells (-MFD flag)(Quinn et al., 1991), forcing a positive flow accumulation for potential underestimates (-a flag) and emphasizing the flow accumulation in flat areas (-b flag). To identify the best carving option, three steps (carving; DEM correction; stream network delineation) were run iteratively until the maximum overlap between the new DEM-derived stream network and the GWS observed water bodies was reached.

## 3 Results

The six carving options produced different "canyons" with different depths and profiles (Fig. 1 shows a transect profile of the GMTED (top), and also displays results from the different carving equations (bottom)). The six carving equations produced distinct streams networks with varying degrees of overlap with the GWS water occurrence (Figure 1). The overall trend is reported in Fig. 2 and the maximum level is reached by Equation 4 at a 50m carving depth. Thus, the best option
was a direct function of the water occurrence and an indirect function of the standard deviation of the elevation. This also highlights that a constant carving (Equation 2), which is usually used, produces less reliable results than a dynamic carving that follows the water occurrence and the topography variation.

4 Conclusion

We demonstrated the use of spectral derived water occurrence layer to carve the GMTED, in order to derive a stream network that more closely matches the observed stream networks. The 30m GWS resolution can be also used for carving DEMs on a higher resolution than used here (e.g. SRTM 90m), thus emphasizing a more accurate stream delineation. However, a thorough calibration procedure is needed to test and implement the best carving option at this spatial resolution. From a computational prospective, GRASS provides fast and flexible functions for hydrological modeling with automated scripting workflows, and allows for the processing of very large data sets using efficient computational algorithms and memory management. Other carving options can be tested.
Figure 2: On the left side: number of DEM-derived stream network pixels that overlap with the GWS bodies versus the carving depth. The maximum value identifies the best carving option, in this case obtained by the carved surface delineation with Equation 4. On the right side: a) GWS occurrence; b) stream network, labeled with red pixels obtained with carving options described by Equation 2; c) stream network obtained with carving options described by Equation 4. The maximization of the overlapping is evident in the c) panel.

and a global implementation of this procedure is in progress.

5 Acknowledgements

This study was supported in part by the facilities and staff at the Yale Center for Research Computing (YCRC). The study was supported by NASA grant NNX17AI74G and Yale Institute for Biospheric Studies (YIBS).

6 References

Danielson, J. J. and D. B. Gesch

Danielson, J. J. and D. B. Gesch

Feng, M., J. O. Sexton, S. Channan, and J. R. Townshend


GRASS Development Team

Heine, R. A., C. L. Lant, and R. R. Sengupta

Hutchinson, M.

Montgomery, D. R. and E. Foufoula-Georgiou

Pekel, J.-F., A. Cottam, N. Gorelick, and A. S. Belward

Pelletier, J. D.

Quinn, P., K. Beven, P. Chevallier, and O. Planchon

Saunders, W.

Tarboton, D. G., R. L. Bras, and I. Rodriguez-Iturbe

Wilson, J. P. and J. C. Gallant