GEOMORPHOMETRY 2020 PERUGIA, ITALY JUNE 22 - 26 2020

A new and extendable global watershed and stream network delineation using GRASS-GIS

Giuseppe Amatulli[§]
School of Forestry and Environmental Studies - Center for Research Computing, Yale University, New Haven, CT-USA
Tushar Sethi, Longzhu Shen
Spatial Ecology, 35A, Hazlemere Road, Penn, Bucks HP10 8AD, U.K
Jaime Ricardo García-Márquez, Jens Kiesel, Sami Domisch
Leibniz-Institute of Freshwater Ecology and Inland Fisheries, Müggelseedamm 310, 12587 Berlin, Germany
[§] giuseppe.amatulli@gmail.com

***Abstract*—We used the hydrologically corrected digital elevation model (DEM) "MERIT Hydro" at a 3 arc-seconds (90 m) spatial resolution to derive a globally seamless, standardised stream network employing GRASS-GIS hydrological modules. Our main aim is to delineate a global stream network that extracts, in particular, small headwaters in greater spatial detail. We hence implement a low upstream area threshold (0.05 km^2) to initiate headwater streams, while using water bodies in the NHDPlusV2 dataset for calibration. The obtained hydrography layers (flow accumulation, flow direction, basins and stream network) are compatible with other r.stream.* modules within GRASS-GIS and hence the dataset can easily be customised and expanded by computing other hydrographical features.**

I. 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 Models (DEMs) and improvement in computational power have led to recent advances in terrain and hydrological analyses 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 [1-2-3-4]. 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 this study, we used the MERIT Hydro at 3 arc-seconds (90 m) spatial resolution and the NHDPlusV2 water bodies, which we processed using the hydrological modules within the Geographic Resources Analysis Support System (GRASS) open source GIS software [5] to derive a globally seamless and standardised hydrography. Unlike that of existing datasets, our aim is to represent headwaters at greater spatial detail, and create a product that can be re-ingested in GRASS-GIS to produce ancillary hydrographical features.

I. Data and method

A. Source layers

The MERIT Hydro dataset [6] was released in 2019 at a spatial grain of 90 m (3 arc-seconds) with the exception of Antarctica. It includes depression areas and hydrologically-adjusted elevation that incorporates, as an elevation correction factor, water occurrence datasets (G1WBM [7], GSWO [8], and OpenStreetMap [9]). The water occurrence datasets are used to modify the elevation satisfying the condition that "downstream is not higher than upstream". Here, it is important to consider that the G1WBM and GSWO are Landsat-derived, and therefore not useful for capturing small tributaries (smaller than 20-30m river width) or even large rivers under canopy cover. In contrast, OpenStreetMap is based on a direct survey of observed water bodies, therefore small tributaries are, depending on the region, depicted as a function of the survey effort. Currently, only a few countries represented in OpenStreetMap provide a high level of spatial accuracy for headwater streams. Therefore, our stream extraction algorithm has to be calibrated such that headwater streams can be delineated at maximum spatial accuracy.

In addition to MERIT Hydro, the NHDPlusV2 is a geo-spatial database of surface water features, developed by the US EPA Office of Water and by the US Geological Survey [10]. This dataset was derived from the US National Elevation Dataset (NED) in a 1 arc-second (~30 m) spatial resolution, and has about 3 million rivers at a 1:100,000-scale or higher [10] which has made it suitable to spatially validate the location of streams

Giuseppe Amatulli, Tushar Sethi, Longzhu Shen, Jaime Ricardo Garcia-Mírquez ,Jens Kiesel, and Sami Domisch (2020)

A new and extendable global watershed and stream network delineation using GRASS-GIS:

in Massimiliano Alvioli, Ivan Marchesini, Laura Melelli & Peter Guth, eds., Proceedings of the Geomorphometry 2020 Conference, doi:10.30437/GEOMORPHOMETRY2020_55.

that are derived from lower-resolution DEMs. The NHDPlusV2 is a reliable data source for locating water bodies, and can be therefore be used to calibrate the stream network delineation algorithm. For the calibration, we selected an area of 20 x 15 degrees in the Mississippi basin.

B. Basin delineation and stream network extraction methodology

In order to address the large demand on computer memory posed by the calculation of the flow accumulation, we used the Level-2 basin polygons from HydroBASINS [11-12] to divide the MERIT Hydro DEM into computational tiles (see Figure 1). We manually re-located and enlarged the tiles across the maximum number of cells possible ($2^{31} - 1$, equivalent to ~2 bilion). This is the maximum number of cells that *r.watershed* can be processed by the RAM. We located the tiles in such a way that larger basins and 2.5 degrees of surrounding area were completely within the tiles. This allows maximum precision in basin delineation and stream network extraction. Using GRASS-GIS, we ran the following modules for each tile: *r.watershed, r.stream.extract, r.stream.basins*.

The command *r.watershed* requires a hydrologically-adjusted elevation as an input, which together with a depression raster map is able to compute flow accumulation and flow direction, using a multiple down-stream cells algorithm (-MFD flag) [13]. The *r.watershed* function also accepts the flow parameter that incorporates a raster map representing the amount of upstream area per cell. Therefore, we first calculated the surface-area of each cell using *r.cell.area* to compute the flow accumulation. Another parameter that influences the flow accumulation and ultimately the stream network location and length is convergence, which ranges from 1 to 10. Lower values result in higher divergence of flow direction, as opposed to higher values that indicate higher convergence.

After *r.watershed,* we ran *r.stream.extract* using flow accumulation and depression. In order to extract the stream we set the minimum upstream flow accumulation area (threshold) to 0.05 km^2 (~6 cells at the equator). This threshold is the minimum area necessary to initiate the channel head and hence results in higher stream density. The threshold typically differs by region [14], in according to climatic and geomorphological features. We note that this stream network delineation is part of a larger project aiming to calculate discharge at a monthly level. Hence, at a later stage we will shorten (prune) river sections and account for possible flow intermittency (flow > 0 m^3/sec) indicates stream presence and produces a dynamic representation that changes the stream dimension as a function of the monthly flow. The final output will be 12 monthly stream networks that differ in length and intermittencies.

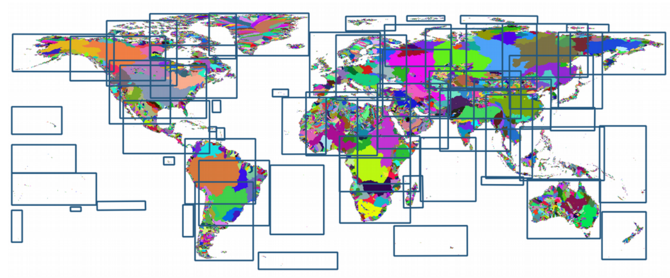

Figure 1: Global 1,608,114 drainage basins (random colours for illustrative purposes only) and computational tiles.

Transforming conventional static stream networks into a monthly dynamic representation overcomes the limitations of fixing *a priori* convergence parameters and the minimum upstream flow accumulation areas. The new computation will produce a stream network that is determined by the environmental characteristics found across each catchment. Stream discharge at a monthly level will be computed using a machine learning approach, which integrates meteorological factors, as well as geomorphometric features, soil types, land cover/use, among others. Hence, the stream length (stream presence if flow > 0 m^3/sec), will be modelled based on the interaction of these factors.

Subsequent to *r.stream.extract* we ran *r.stream.basins* to delineate basins using the flow direction and the previously calculated streams. For each tile, we retained only the complete basins and removed the ones adjacent to tile borders that were possibly truncated. We also used the same procedure (retain and remove) for the other hydrography layers, keeping the ones that belonged only to entire computed basins. Finally, we re-aggregated all of the hydrography layers to yield a complete global basin delineation.

In order to calibrate the best convergence value, we repeated the full processing chain for each of the 10 convergence factors available in GRASS-GIS, and spatially overlaid the newly-developed stream network on top of the one from NHDPlusV2. The convergence value that produced the maximum overlap between our computed stream network and the NHDPlusV2 network was considered the most appropriate value, and was used for the remainder of the global-scale implementation.

II. RESULTS

With the described procedure we were able to complete global layers for flow accumulation, flow direction, basins and stream networks. In total, we delineated 1,608,114 basins (see Figure 1), and for each we extracted constituent stream networks. These stream networks overlapped cleanly against the

NHDPlusV2, although some spatial overestimation in MERIT-DEM-derived streams is known. A final validation procedure will be conducted at the end of the project for accurate discharge computation. The final optimal convergence parameter equals 10 (see Figure 2). Once delineated, each basin can be used as a computational unit to calculate other hydro-geomorphological features, such as stream length, stream order or residence time. For context, flow accumulation can reach a maximum value of 5,509,644 km^2 (Amazon River), followed by 3,671,834 km^2 (Nile River) and 3,197,652 km 2 (Mississippi-Missouri River).

III. Conclusion

We demonstrated the processing of MERIT Hydro by means of r.watershed and *r.stream.** modules in GRASS-GIS to derive a stream network that more closely matches the observed spatial representation of streams and rivers. As demonstrated, the NHDPlusV2 can be used effectively to calibrate the convergence parameters.

Other regional and national stream networks can also be used to validate the final product. From a 5computational perspective, GRASS-GIS provides fast and flexible functions for hydrological modelling with automated scripting workflows, and allows the processing of very large datasets using efficient algorithms and memory management. We are in the process of accurately calibrating and validating stream network extraction and basin delineation worldwide. For the calibration and validation procedures, we will use detailed stream networks that are freely available from governmental institutions. The final products will be stored as hydrography layers that can be re-ingested into GRASS-GIS to produce ancillary hydrographical features.

IV. Acknowledgements

This study was supported in part by the facilities and staff at the Yale Center for Research Computing (YCRC) and by the Leibniz Competition (J45/2018).

References

[1] D. G. Tarboton, R. L. Bras, and I. Rodriguez-Iturbe, 1992. “A physical basis for drainage density,” Geomorphology, vol. 5, no. 1, pp. 59–76.

[2] D. R. Montgomery and E. Foufoula-Georgiou, 1993. “Channel network source representation using digital elevation models,” Water Resources Research, vol. 29, no. 12, pp. 3925–3934.

[3] R. A. Heine, C. L. Lant, and R. R. Sengupta, 2004. “Development and comparison of approaches for automated mapping of stream channel networks,” Annals of the Association of American Geographers, vol. 94, no. 3, pp. 477–490.

[4] J. D. Pelletier, 2013. “A robust, two-parameter method for the extraction of drainage networks from high-resolution digital elevation models (dems): Evaluation using synthetic and real-world dems,” Water Resources Research, vol. 49, no. 1, pp. 75–89.

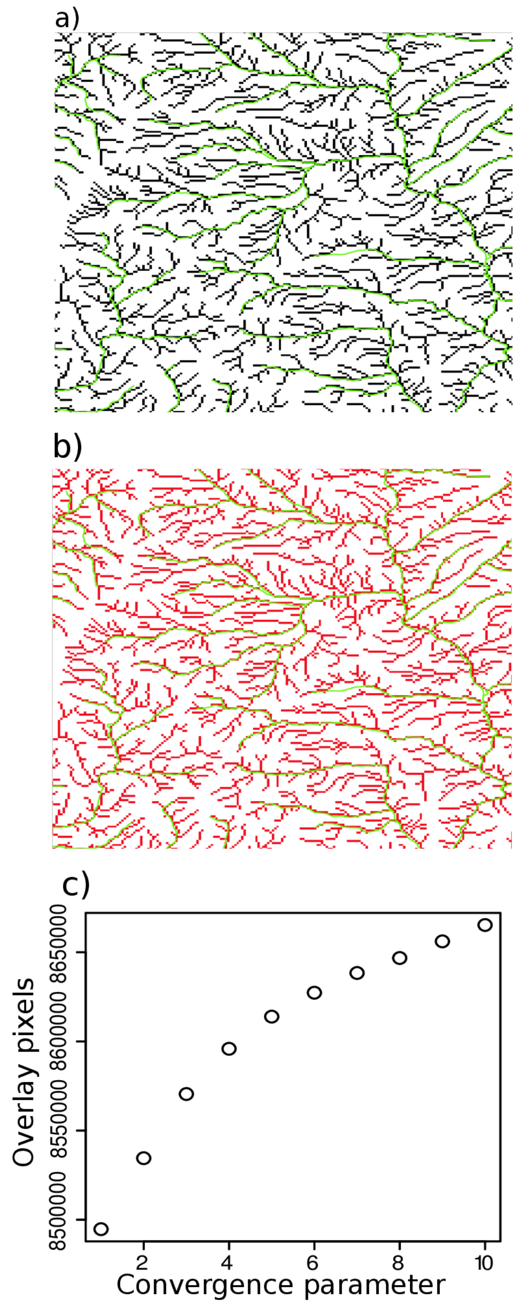

Figure 2. Convergence calibration. (A) The green lines represent the NHDPlusV2 stream network; the black lines show the newly-developed stream network generated using a convergence value of 1, while (B) the red lines represent the newly-developed stream network with a convergence value of 10. (C) The NHDPlusV2 and the newly-developed stream network pixels overlap as a function of the convergence.

[5] GRASS Development Team, Geographic Resources Analysis Support System (GRASS GIS) Software, Version 7.2, 2017. Open Source Geospatial Foundation.

[6] D. Yamazaki, D. Ikeshima, J. Sosa, P. D. Bates, G. H. Allen, and T. M. Pavelsky, 2019. “Merit hydro: a high-resolution global hydrography map based on latest topography dataset,” Water Resources Research, vol. 55, no. 6, pp. 5053–5073,.

[7] D. Yamazaki, M. A. Trigg, and D. Ikeshima, “Development of a global˜ 90 m water body map using multi-temporal landsat images,” Remote Sensing of Environment, vol. 171, pp. 337–351, 2015.

[8] [8] J.-F. Pekel, A. Cottam, N. Gorelick, and A. S. Belward, 2016. “High-resolution mapping of global surface water and its long-term changes,” Nature, vol. 540, no. 7633, pp. 418–422.

[9] OpenStreetMap contributors, 2017. “Planet dump retrieved from https://planet.osm.org .” https://www.openstreetmap.org.

[10] C. H. David, D. R. Maidment, G.-Y. Niu, Z.-L. Yang, F. Habets, and V. Eijkhout, , 2011. “River network routing on the nhdplus dataset,” Journal of Hydrometeorology, vol. 12, no. 5, pp. 913–934.

[11] B. Lehner, K. Verdin, and A. Jarvis, “New global hydrography derived from spaceborne el-evation data, 2008. “Eos, Transactions American Geophysical Union, vol. 89, no. 10, pp. 93–94.

[12] B. Lehner 2014, “Hydrobasins technical documentation, version 1. c,” HydroSHEDS.

[13] P. Quinn, K. Beven, P. Chevallier, and O. Planchon, 1991. “The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models,” Hydrological processes, vol. 5, no. 1, pp. 59–79.

[14] J. V. Vogt, R. Colombo, and F. Bertolo, 2003. “Deriving drainage networks and catchment bound-aries: a new methodology combining digital elevation data and environmental characteristics,” Geomorphology, vol. 53, no. 3-4, pp. 281–298.