Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2019-34-AC1, 2019

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "The importance of better mapping of stream networks using high resolution digital elevation models – upscaling from watershed scale to regional and national scales" *by* Anneli M. Ågren and William Lidberg

Anneli M. Ågren and William Lidberg

william.lidberg@slu.se

Received and published: 26 February 2019

We would like to thank reviewer 1 for comments and suggestions for the manuscript. We believe that the comments have identified important areas which required improvement. We have taken the comments on board to improve and clarify the novelty manuscript.

The scope of HESS is to encourage and support both fundamental and applied research. This study does not advance the fundamental research, but, it fills an important research gap in applied science. Which is to generate maps that can be used in practice, based on the best available science. It's it true that there are studies that has helped researchers understand the controls on stream head formations and perennial/non-perennial headwaters across larger regions. Jaeger et al. (2019) that the referee mentions, is an excellent example of such a study. However, such regional studies often aggregate the DEMs to larger pixels to minimize computational resources. Jaeger et al. (2019) use 30 m pixels, González-Ferreras & Barquín (2017) use 25 m, Julian et al. (2012) and Elmore et al. (2013) use 10 m. While this resolution is good for advancing the processes understanding it's important to work on a higher resolution when generating maps of headwaters that can be used in practice and that accurately places the modelled streams. In our case all field mapped streams had a width less than 6 m. For mapping purposes it therefore makes sense to use a scale <6 m. Russell et al (2015) worked on a 6.1 m resolution, stating that "The resolution of 6.1 m (20 ft) was selected to maximize the accuracy of terrain derivatives while minimizing computational resources." Jensen et al. (2018) use a resolution of 3 m. These studies are examples of better resolutions for mapping small streams, however that also meant that they worked on a catchment scale. Jensen worked on 3 catchments smaller than 45 ha and Russell worked on 7 catchments ranging 1.1-4.8 km2. In Lidberg et al. (2017) we showed that the accuracy of topographically derived stream networks increased with increasing DEM resolution. DEM derived stream networks were field validated against >30 000 field observations of stream/road intersections from 9 catchments ranging 68-1 993km2. Stream networks from the 2 m DEM intersected roughly twice the number of stream/road intersections as stream networks from the 16 m DEM. We also found that working with high resolution DEMs (2 m) requires special pre-processing methods as new challenges are introduced with increasing resolution because the effect of anthropogenic artefacts such as road embankments and bridges. These are problematic during the pre-processing step because they are elevated above the surrounding landscape and act as artificial dams (Figure 1). The effect of different pre-processing methods were studied in Lidberg et al (2017) and that knowledge was applied in this study.

This study is the first study where all headwater stream channels have been mapped throughout an entire country (Sweden is 450 295 km^2) at a high resolution (2 m). A drawback of working in a 2 m resolution on this large scale is the processing times for the modelling of stream channels. But, the rewards are that the modelled streams on the digital maps mostly follow the inundations of the small streams and ditches that can often be picked up by really high-resolution DEMs (Figure 2). We have now more specifically expressed the novelty of this study in the manuscript and also added a new figure in the manuscript, highlighting the benefits of working on a high resolution when mapping headwaters.

The development of LiDAR models have also changed forest management. It's now possible to segmenting individual trees from the LiDAR point clouds (Li et al. 2013) and to estimate tree height (Suárez et al. 2005) and biomass (Popescu 2007) of individual trees. Such models can be used to assess the vegetation structure (Morsdorf et al. 2004) and forest succession (Falkowski et al. 2009) and be used in, for example, fire risk assessment and fire behavior modeling as well as wildlife habitat assessment. As forest management work on higher resolutions (individual trees) in order to optimize different goal such as production or biodiversity, it's also important to develop tools for surface water protection that meets the demands of high-resolution maps by the future forest industry. Better mapping of headwater streams is a key part of this work.

When it comes to binary data in models of headwaters, studies like González-Ferreras & Barquín, (2017) and Jaeger et al. (2019) use machine learning (Random Forest) to model perennial/non-perennial streams. We explored the possibility of using machine learning algorithms in our study. However, our dataset did not contain stream heads and further it is hugely imbalanced with 3 323 streams and 616 444 non-streams.

To create a balanced dataset that could be analyzed further in the residual analysis we used the following approach: “Prior to analysis a balanced dataset was created by randomly selecting and equal number of FN and FP and all variables were scaled and centered."

As pointed out by the reviewer, logistic regression is a common method for classifications. However, all regression methods are based on a number of fundamental assumptions that needs to be met. Our dataset met all the assumptions for a logistic regression – except one, it violates the assumption of no multicollinearity. Meaning that several of our independent variables are highly correlated with each other. Therefore we cannot use a regression method. A way of dealing with multicollinearity is to use a multivariate approach where principal components are extracted from the dataset of X- variables. These component are extracted so they are orthogonal to each other, meaning that the components are not covariate. That is why the OPLS-DA was the correct statistical method to use on our dataset. In the article we wrote "First the X-matrix was tested for multicolinearity, using IBM SPSS Statistics 24 and since many of the explanatory variables in the residual analysis showed multicolinearity a multivariate approach was used."

Having said that, we are also aware that the regression models often are a robust method, and we also tested a Logistic Regression using IBM SPSS Statistics v. 24. The Hosmer and Lemeshow test for significance showed that only three of the models were significant, step 2, 4 and 5. R2 for those models range 0.14-0.21, depending on the chosen method for R2 (Cox & Snell R2 or Nagelkerke R2). That's of the same order as the OPLS-DA which had R2Y(cum) = 0.16 and Q2(cum) = 0.15. The important variables in step 5 were Winter, StDev5, StDev10, X-coordinates and fine sediments. Similar to the findings in the OPLS-DA. So, applying a logistic regression does not change the major findings of this study, nor does it improve the explanatory power of the model (of the same order). However, bear in mind that it's also statistically wrong to use a regression method on this dataset.

It's true that we only create a single global model across all of Sweden and then use the residual analysis to infer how the models can be improved in the future by taking into account landscape variability. We have now expanded the discussion on the residual analysis in the manuscript. We believe that in order to further improve the modelling and adjusting the stream initiation thresholds across physiographic regions we need to create an entire new field dataset with stream heads mapped throughout the nation. While it's our aim to work towards that goal during the upcoming years, we have still learned a lot in this study which fills an important gap in the applied science community. We show that it's possible to generate stream networks on a high resolution (2 m) over an entire country. The plan is that the digital stream network (2 ha) will be made available through the Geodata Collaboration which give authorities, county councils / regions, municipalities, public-sector organizations and forest owners access to the maps. Wide access to these maps will improve the basis for landscape planning, including forestry, and can contribute to reduce the export of nutrients and hazardous substances to streams, lakes and the Baltic Sea.

Regarding the comments on English spelling and grammar errors. We will send the manuscript on languish editing to address these problems.

References

Elmore, A. J., J. P. Julian, S. M. Guinn, and M. C. Fitzpatrick. 2013. Potential Stream Density in Mid-Atlantic U.S. Watersheds. PLoS ONE. doi:10.1371/journal.pone.0074819.

Falkowski, M. J., J. S. Evans, S. Martinuzzi, P. E. Gessler, and A. T. Hudak. 2009. Characterizing forest succession with lidar data: An evaluation for the Inland Northwest, USA. Remote Sensing of Environment. doi:10.1016/j.rse.2009.01.003.

González-Ferreras, A. M., and J. Barquín. 2017. Mapping the temporary and perennial character of whole river networks. Water Resources Research. doi:10.1002/2017WR020390.

Jaeger, K. L., R. Sando, R. R. McShane, J. B. Dunham, D. P. Hockman-Wert, K. E.

Kaiser, K. Hafen, J. C. Risley, et al. 2019. Probability of Streamflow Permanence Model (PROSPER): A spatially continuous model of annual streamflow permanence throughout the Pacific Northwest. Journal of Hydrology X 2. The Authors: 100005. doi:10.1016/j.hydroa.2018.100005.

Jensen, C. K., K. J. McGuire, and P. S. Prince. 2017. Headwater stream length dynamics across four physiographic provinces of the Appalachian Highlands. Hydrological Processes 31: 3350–3363. doi:10.1002/hyp.11259.

Julian, J. P., A. J. Elmore, and S. M. Guinn. 2012. Channel head locations in forested watersheds across the mid-Atlantic United States: A physiographic analysis. Geomorphology. doi:10.1016/j.geomorph.2012.07.029.

Li, W., Q. Guo, M. K. Jakubowski, and M. Kelly. 2013. A New Method for Segmenting Individual Trees from the Lidar Point Cloud. Photogrammetric Engineering & Remote Sensing. doi:10.14358/pers.78.1.75.

Lidberg, W., M. Nilsson, T. Lundmark, and A. M. Ågren. 2017. Evaluating preprocessing methods of digital elevation models for hydrological modelling. Hydrological Processes 31: 4660–4668. doi:10.1002/hyp.11385.

Morsdorf, F., E. Meier, B. Kötz, K. I. Itten, M. Dobbertin, and B. Allgöwer. 2004. LIDAR-based geometric reconstruction of boreal type forest stands at single tree level for forest and wildland fire management. In Remote Sensing of Environment. doi:10.1016/j.rse.2004.05.013.

Popescu, S. C. 2007. Estimating biomass of individual pine trees using airborne lidar. Biomass and Bioenergy. doi:10.1016/j.biombioe.2007.06.022.

Russell, P. P., S. M. Gale, B. Muñoz, J. R. Dorney, and M. J. Rubino. 2015. A spatially explicit model for mapping headwater streams. Journal of the American Water Resources Association. doi:10.1111/jawr.12250.

Suárez, J. C., C. Ontiveros, S. Smith, and S. Snape. 2005. Use of airborne LiDAR and

Printer-friendly version

Discussion paper

aerial photography in the estimation of individual tree heights in forestry. Computers and Geosciences. doi:10.1016/j.cageo.2004.09.015.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-34, 2019.

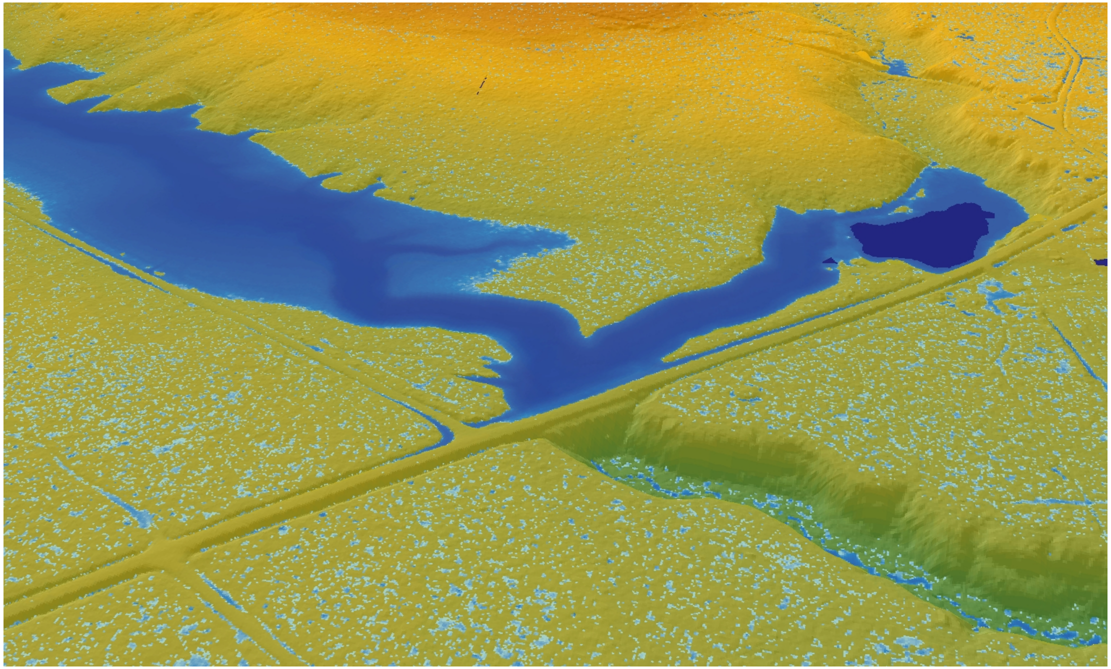

Fig. 1. Road banks are elevated above surrounding terrain and behaves like dams in the digital elevation model.

Printer-friendly version

Discussion paper

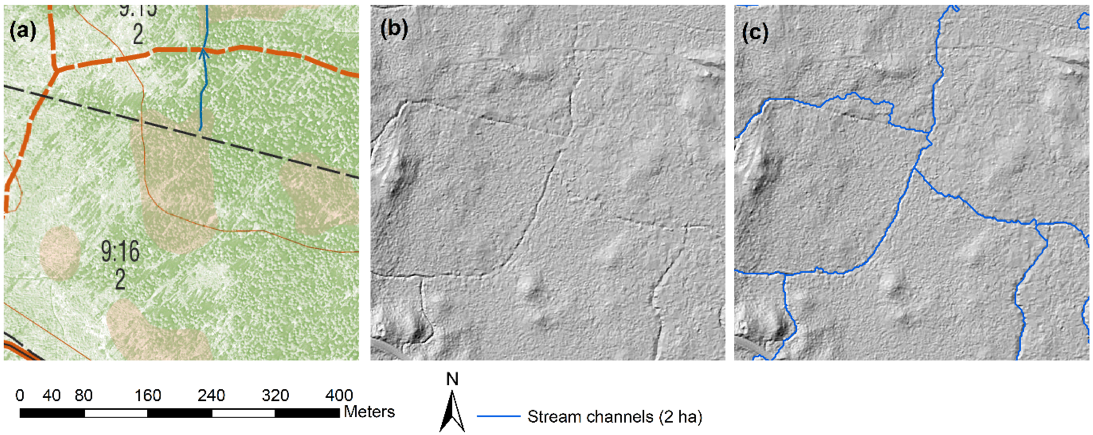

Fig. 2. a) Property map where streams are missing b) small stream channels and ditches in the DEM. c) modeled stream network inside DEM channels.

Printer-friendly version

Discussion paper