Hydrol. Earth Syst. Sci. Discuss., 7, C2477–C2483, 2010
www.hydrol-earth-syst-sci-discuss.net/7/C2477/2010/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Accurate stream extraction from large, radar-based elevation models" *by* M. Metz et al.

Anonymous Referee #2

Received and published: 27 September 2010

Review: HESS Discussion 7, 1–23, 2010 Manuscript: Accurate stream extraction from large, radar-based elevation models Authors: Metz, M., et al.

This manuscript describes an improved least-cost path search algorithm to enhance the computational capability and accuracy of automatically derived drainage networks from radar-based digital elevation models. The authors compare two such datasets at various levels of resampling for a study area in Panama, and compare their findings with reference points obtained from satellite-based and field-based GPS mapping. The authors contend that this improved technique facilitates efficient use of computational resources, while providing for better reliability in hydrological modeling especially for large data sets.

The manuscript is a technical contribution to the interdisciplinary field of geomorphometry, and fits the scope of the Special Issue of HESS in that it proposes a useful alternative to existing hydrological sink-filling algorithms. Language, grammar, and formal presentation of the study are mostly easy to follow, and well structured. Nevertheless, I have several remarks that the authors may want to consider if wishing to revise this work. Most of my major comments are concerned with a geomorphologist's, i.e. potential user, viewpoint. These are followed by a range of specific comments keyed to page and line numbers in the manuscript.

Overall, I recommend major revisions in light of the rather sparse documentation of the reference data, on which the whole point of the improved method's accuracy is based.

General Comments

1. The manuscript is rather technical and clearly intended for an audience with expertise in geomorphometry. With this scope, the algorithm that has been improved here could have been described in a more formal way. Moreover, to assist non-expert readers a schematic figure outlining how the improved least-cost path search algorithm works would be nice, while graphically highlighting the peculiarities of this method. Also, the IFSARE data could be briefly described, assuming that most readers will know what SRTM data are.

2. Despite the technical level of the manuscript and the frequent use of jargon, some of the wording comes across in a bulky fashion. Several statements would benefit from more frequent punctuations and less complicated phrases.

3. The validation of the stream extraction algorithms needs significant further elaboration, as it supports the argument about improved accuracy. It seems that distances between digital streams and satellite- and ground-based reference points are used to assess the accuracy of the drainage networks extracted. However, it is unclear which landforms or parts thereof these reference points record. I think a figure showing some of these points together with the extracted drainage networks is a minimum require-

ment.

4. Along these lines, the authors have digitized streams from Landsat imagery covering their study area, but fail to explain what types of streams were digitized and what criteria were used. In other words, which diagnostic criteria defined the stream channels or parts thereof? There are well known issues, for example, when trying to delineate braided or anabranching river channels. Which bands of the Landsat data were used at which resolution? How does this reverberate on the effect of growing channel width in a downstream direction?

5. The GPS ground truthing data require further specification in terms of acquisition and positional accuracy. How well could the rivers or their channels be accessed and to what degree of reliability have hydrologically meaningful boundaries between detected and recorded?

6. Have rivers of all orders been sampled equally or is there a bias towards a certain range of basin sizes? In this context I would suggest extending the testing scheme to include potential first-order effects of topography on the accuracy of drainage network extraction, assuming that rivers with low-gradient floodplains would be more difficult to capture during extraction.

7. The performance comparison (section 3.1) is impressive, but could be illustrated more by providing some benchmark data on the anticipated processing times for datasets of a given size. This would be quite informative and allow readers a more careful planning and management of resources for large dataset analyses.

8. The discussion is thin and starts off by recapitulating a lot of previous knowledge without any due referencing. The rest of the discussion seems to re-state some of the general results without digging any deeper by objectively discussing potential pitfalls, faulty assumptions, or the general robustness of the methodology. There is no mention of whether the results were expected or supporting some sort of hypothesis, or whether the authors think that these results are likely to be applicable elsewhere. I strongly

encourage the authors to be more critical with the work they present here.

Specific Comments (p/l)

Title: For reasons you spell out in the discussion section there cannot be an accurate extraction of drainage networks from DEMs without specifying a degree of reliability. Consider changing title appropriately.

2/18 “GPS”: Explain what type of GPS you used. Also describe the nature of “reference points”.

4/8 Can you give examples of the “new generation of watershed analysis tasks”?

4/15 “I/O” needs an explanation.

4/26 “open” vs. “closed list” may want to be explained for non-expert readers.

5/26 “SRTM DEM”: You may want to acknowledge the data source.

6/2-4 “The elevation data were not bare earth representations and over large regions include elevation surface defined by triple canopy tropical forest environment with tree heights of more than 30m above the bare earth surface.”: This reads a bit awkward. How about simply stating that the DEM data represent three different layers of forest canopy?

6/5 “gaps in the dataset filled”: How sensitive is your proposed improved least-cost path search algorithm to the gap filling procedure? You have investigated the effects of resampling, but not any effects of the original gap filling method.

6/10 “SRTM DEM was then resampled to 30m resolution”: This was surely done for reasons of comparability only, right? Otherwise I do not see any benefits of artificially increasing the nominal grid-size resolution of these data.

7/13 “Stream segments digitized from Landsat imagery”: This needs some further elaboration. What type of streams did you digitize how from which of the Landsat bands?

9/7 Delete “this”.

9/10 “median distances of reference points” may be meaningless if it is not fully clear which channel or floodplain elements you had measured as reference points, and at which effective resolution.

9/12-14 “While the stream locations for the SRTM 90m DEM were about 100m away from reference points, the median distances improved to down to 34.4m for the IFSARE 10m DEM.”: This should not be much of a surprise, should it?

9/15 “Streams extracted from LCP flow accumulation were closer to the GPS field points”: You may want to comment on how reliable you judge these proximity measures. Does it matter if, for example, a meander bend is positionally not that accurate, but any given automatically derived stream segment close to a given reference point? In other words, what sort of criteria did you postulate for channel planform or pattern matching?

10/3-4 “with vastly different levels of detail captured by their original resolutions of 10m and 90m respectively”: This is rather obvious.

10/9-10 “the accuracy of stream locations was considerably higher for IFSARE 30m than for SRTM 30m”: Again, this seems to be quite obvious, with the SRTM data having a 90-m resolution originally. Artificially increasing the resolution will not add more detail to those coarse data, so making this point is a bit superfluous.

10/11-12 “smaller rivers in mountainous regions”: You should define “small” in this context. Also, one could easily imagine that the methods have a more similar accuracy in terrain with some relief, whereas most of the errors would occur in low-gradient floodplain areas. Consider extending your testing scheme for these first-order effects of topography.

11/5-6 “bare earth models” could also read “digital terrain models”.

11/2-14: Many of the points mentioned here have been known before, and you may want to acknowledge some of the key references that had outlined this before.

11/19 “the ground truth” -> “reality”.

11/27 “relatively flat landscape”: Relative to what? Such potential bias towards topography may weigh heavily on your assessment of accuracy. You should try to make the assessment as representative as possible, i.e. evenly spread over the size range of drainage basins and relief in a given study area.

12/1-2 “no method was able to accurately delineate streams because the information was simply not present in SRTM 90 m”: Consider rephrasing. The information to extract drainage networks is there, but whether it is of sufficient accuracy for a given task is a different matter.

12/6-7 “SRTM does not provide sufficient vertical detail to determine river courses in flat areas.”: But this has surely been known before?

12/8-10 “Resampling to a higher horizontal (and for SRTM also higher vertical) resolution seems to be helpful in improving the results of watershed analysis from low-resolution DEMs”: I would expect from a discussion to hear your interpretation of this observation. Did it fit your expectations or hypothesis (if any)? Also, one might wonder whether this improvement through resampling was just a lucky shot given the values you chose. In other words, such a general statement should be supported by more than what you show here or, alternatively, be toned down accordingly.

Table 2 features spurious figures. Also make sure document you the positional accuracy of your reference points due to measurement method and accessibility.

Fig. 1. The red points showing the satellite-derived river points are next to useless at this resolution. Consider showing a detail of a satellite image with some of the points digitized.

Fig. 2. Explain box-and-whisker signatures here and in other figures.

Fig. 6. Some more context for this figure would be nice. Do you show the rule or the exception here? How was the river course determined in the first place?

Fig. 7. State in the caption what you intend to show here exactly. Would it not be more instructive to zoom in further and also show the relation with reference points?

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 7, 3213, 2010.