Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2015-539-AC2, 2016

Hydrology and Earth System Sciences
Discussions
Open Access
HESSD

Interactive comment

Interactive comment on "Case-based formalization and reasoning method for knowledge in digital terrain analysis – Illustrated by determining the catchment area threshold for extracting drainage networks" *by* C.-Z. Qin et al.

C.-Z. Qin et al.

qincz@lreis.ac.cn

Received and published: 19 June 2016

The authors thank the anonymous referee for the constructive comments which are helpful for improving the final version of this manuscript. We answer these comments as below.

Comment 1: p. 1 l. 19 DTA-assisted tools (e.g., ArcGIS, GRASS, SAGA, White Box, TauDEM) ArcGIS and GRASS are large, general purpose GIS packages which include DTA tools, - reference to specific modules is needed here.

Response: DTA-assisted tools include general purpose GIS packages with DTA func-

Printer-friendly version

Discussion paper

tionality (e.g., “Spatial Analyst“ toolbar in ArcGIS, r.* modules in GRASS, “Terrain Analysis“ menu in SAGA, etc.) and domain-specific software (e.g., Whitebox, TauDEM, etc.) (Hengl and Reuter, 2009). We will revise the manuscript to clarify this point.

Comment 2: l. 25 I find the following sentence confusing algorithm knowledge, which is the metadata of a DTA algorithm - what do authors mean by this?

Response: The algorithm knowledge is the metadata of a DTA algorithm (including its parameters), such as the data type of input/output file, the number of parameters, and the valid range for each parameter. We will revise the manuscript to clarify this point.

Comment 3: p.2 l. 1 again ModelBuilder is not DTA-assisted tool - not clear what is meant here Response: ModelBuilder module in ArcGIS uses task knowledge and algorithm knowledge to aid connecting a set of DTA algorithms to be an executable DTA workflow in a interactive visual way. We will revise the manuscript to make it clear.

Comment 4: p.3, l. 6 this assumes that there is no validation data available - isn't the best way to find the optimal parameters running the tools with a set of parameters and find the best fit with the field data (or remotely sensed data if they provide sufficient information)? What if the case studies are inaccurate? Can this be taken into account?

Response: We agree with the reviewer that the best way to determine the optimal parameter-settings should be the evaluation based on the field data. However, at the beginning of the modeling, field data might be not easy to be obtained and the evaluation process is not easy to operate for those non-expert users. The method proposed in this study might automate the DTA modeling process, which makes it easy for users (especially non-expert users), and meanwhile the result model could be reasonable comparatively. We will revise the manuscript to discuss this point. The algorithm and parameter-settings presented in those journal papers might not be optimal, thus the corresponding cases might be inaccurate. In this study, we manually selected the peer-reviewed papers related to the drainage network extraction applications which were published in mainstream journals of related domains. By this means the cases used could be kept as accurate (or reliable) as possible. Additional research is needed to enhance the proposed method by taking the reliability of the case into account. We will revise the manuscript to discuss this point.

Comment 5: p. 7 l. 15 What is meant by aspect here?

Response: Here the “aspect“ means the kind of attributes designed to describe the terrain condition. We will revise the manuscript to use the term unambiguously.

Comment 6: l. 17 how do you compute relief - you refer to it as steep or gently sloping - isn’t that equivalent to slope? Relief in geomorphometry is a very specific metrics - specify here what you are using or use different term

Response: Here it means the total relief of the study area, which is the maximum minus minimum elevation within the study area. We will revise the manuscript to use the term properly.

Comment 7: l. 24 seven grades? did you meant seven classes or categories? It appears that you mix relief and slope - perhaps use equations to precisely define what you mean

Response: Yes, the slope gradient value was divided into seven classes. We will revise the manuscript to make it clear and also precisely define the calculation of the total relief used in this study.

Comment 8: l. 26-27 10 level x 7 grade - did you mean 10 elevation classes x 7 slope classes?

Response: Yes, we will revise the manuscript to make it clear and use the term “elevation-slope cumulative frequency distribution“ instead of the “relief-slope cumulative frequency distribution“ used in the original manuscript.

Comment 9: l. 30 relieves the DEM resolution effect ? what do you mean by relieves?

Response: DEM resolution has a strong influence on calculating the slope gradient and its frequency distribution (Chang and Tsai, 1991; Grohmann, 2015), while the DEM resolution has a comparatively weak influence on the cumulative frequency distribution of slope gradient. To relieve this DEM resolution effect and ensure the comparability of slope distributions from two cases with different DEM resolutions, we use the slope cumulative frequency in this study instead of the slope frequency distribution to describe the slope distribution. We will revise the manuscript to clarify this point.

Comment 10: p. 8 l. 20 comment - environmental conditions, especially the groundwater level could be more important than the topo parameters, so the case studies used should be evaluated for this and those where parameters other than the proposed ones play determining role should be excluded

Response: We agree with the reviewer that the groundwater level also plays important role on drainage network formation. However, the information of groundwater level is often difficult to be collected. Normal way of drainage network extraction by DTA is mainly based on topographic information. The method proposed in current study focuses on DTA domain and considers the area and the terrain condition for describing the study area characteristics of a DTA application case. Preliminary evaluation results show the reasonableness of the proposed method. The design of the attributes used to describe the problem part of a case could be improved to describe the domain-specific application-context information in an all-round and efficient manner, which needs additional research. We will revise the manuscript to discuss this issue.

Comment 11: p. 9 l. 17 - Doesn't the need to empirically adjust the shape of the bell curve beat the purpose of the proposed method?

Response: Currently we empirically set the shape-adjusting parameter (w) with fixed values for two attributes with bell-shaped similarity function. Preliminary evaluation results show that the proposed method with these settings performs well. The way of setting the shape-adjusting parameter will be explored as a part of future research. For example, if case base is with a large size, a machine learning algorithm might be available for calibrating the shape-adjusting parameter automatically. We will revise the manuscript to include the discussion on this issue.

Comment 12: eq. 1 ln(0.5) is a constant - why ln and not the constant value directly?

Response: We accept this advice and will revise Eq. (1) accordingly.

Comment 13: p. 10 l. 4 and 5 magnitude of cell size - did you mean absolute value? Magnitude does not make sense here. If it is indeed absolute value (as indicated in Table 2), this treats cell size larger the same as cell size smaller - there is a fundamental difference between downscaling and upscaling or going to higher level of detail versus lower level of detail in terms of stream extraction - how do you account for this issue?

Response: In this study, we try to keep the similarity function on each attribute as a simpler form before more detailed research could be conducted to improve it. Current design of the similarity function for cell size is mainly based on two reasons. First, the numerical difference in cell size does not work. Taking an application with 10-m resolution as example, another application with a coarser resolution of 25 m is comparable to it from a cell size perspective, while on the other hand the resolution cannot be less than or equal to 0 m. Secondly, a bell-shaped similarity function for a logarithmic transformation of cell size could balance the decrease of similarity value for those situations with a coarser resolution or a finer resolution. Note that the similarity value on cell size will rapidly decrease to be about 0.58 when the resolution is coarsened to be double the resolution of a case or is refined to be a half of the case's resolution. The lower similarity value will deny the corresponding case to be a credible solution provider for the new application problem. This means that the current method proposed does not suggest a large-step downscaling and upscaling application of existing cases. We will revise the manuscript to state this point.

Comment 15: p. 10 l. 15 - what is meant by area - total area of the study site? magnitude here probably should be again the absolute value

Response: Yes, the area attribute is the total area of the study site. In this study we design a bell-shaped similarity function for a logarithmic transformation of area based on the idea similar to the design for the cell size attribute. Please also see our response above to the 14th item of comments from anonymous referee #1. We will revise the manuscript to make this point clear.

Comment 16: p. 10 l. 22 it is not clear what is meant by relief here - providing an equation or more precise definition is necessary, is it the difference between the minimum and maximum elevation in the study area? If yes, please check how the term relief is used in literature and what should you use here.

Response: We will revise the manuscript to use the term "total relief" and also precisely define the calculation of the total relief, i.e., the maximum minus minimum elevation within the study area.

Comment 17: p. 12 l. 1 the presented workflow applies to only the older algorithms and is highly simplified - this needs to be mentioned. For example, filling of pits (many are often real) and flat areas is not necessary if least cost path algorithm is used - see e.g. Metz et al. 2011, doi:10.5194/hess-15-667-2011r the second step also is not quite accurate - spatial distribution of catchment area sounds confusing - perhaps you meant flow accumulation or contributing areas for each grid cell?

Response: We accept this advice and will revise the manuscript accordingly. A new reference (Metz, M., Mitasova, H., and Harmon, R. S.: Efficient extraction of drainage networks from massive, radar-based elevation models with least cost path search, Hydrol. Earth Syst. Sci., 15, 667-678, 2011) will be cited in the revised manuscript.

Comment 18: l. 15 it is apparent that the proposed experiment applies only to ArcGIS-based workflow which is highly limited and somewhat obsolete, but it can still be used as a case study, given the large number of users who would use this tool. Were all the articles used as case base using the same algorithm?

Response: In most of articles used for case preparation a single flow direction algorithm (such as D8 algorithm) was adopted, when a few articles did not state clearly the flow direction algorithm used. Note that the experiment in this study was designed to focus on the determination of CA threshold for drainage network extraction, not the flow direction algorithm used. We will revise the manuscript to state this point.

Comment 19: p. 12 l. 29, 30 - what is meant by extracting here? perhaps identifying?

Response: Yes, we will revise the manuscript accordingly.

Comment 20: did all articles use SRTM or ASTER?

Response: Some articles used for case preparation in this study used DEM with a finer resolution than that of SRTM or ASTER DEM. However, those DEM are often not easy to collect by us. Therefore, we used these open DEM data to derive the case attributes such as area, total relief, elevation-slope cumulative frequency distribution, and hypsometric curve. And this process also makes each of these attributes comparable between a case and a new application problem. We will state this point in the revised manuscript.

Comment 21: It is not clear why river density for evaluations - how is it computed? Why not the total length of the river network? How many validation cases lead to shorter streams and how many were longer (see Fig. 4).

Response: The river density was calculated by the total length of the extracted drainage network divided by the area of the study site. In current manuscript, the relative deviation of river density was used as an index for quantitative evaluation of the proposed method. Based on Eq. (2) in the manuscript, which defines this index, the index value will be same if the total length of river network is used instead of the river density. Compared with the length of drainage network, the river density can also be used to make comparison between the results for different application problems, although this comparison has not been made for discussion in current manuscript. The counts of validation cases which got shorter and longer drainage networks from the proposed method are 16 and 28, respectively. We will revise the manuscript to provide this information.

References

Chang, K. and Tsai, B.: The effect of DEM resolution on slope and aspect mapping, Cartogr. Geogr. Inf. Syst., 18, 69-77, 1991.

Grohmann, C. H.: Effects of spatial resolution on slope and aspect derivation for regional-scale analysis, Comput. Geosci., 77, 111-117, 2015.

Hengl, T. and Reuter, H. I.: Geomorphometry: Concepts, Software, Applications, Elsevier, Amsterdam, 2009.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2015-539, 2016.