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.

Anonymous Referee #1

Received and published: 15 May 2016

The paper presents an interesting and potentially valuable approach to suggest parameters for stream extraction tools based on the past use cases. Although the manuscript is generally well structured there are issues with the use of terminology and description of the methods and use cases.

Specifically:

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.

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

p.2 l. 1 again ModelBuilder is not DTA-assisted tool - not clear what is meant here

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?

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

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

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

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

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

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

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

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

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?

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

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.

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?

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?

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

did all articles use SRTM or ASTER?

It is not clear why river density for evaluations - how is it computed? i 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).

Overall this is a promising approach, proposed in a highly simplified form in the manuscript. Careful revision of terminology and clarification of the workflow and al-

C3

gorithm issues is needed to avoid confusion and make the paper scientifically sound.