Hydrology and
Earth System
Sciences

Discussions

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 #2

Received and published: 29 May 2016

Formalization of application context knowledge in digital terrain analysis I feel as relevant scientific problem also within the scope of HESS. The paper presents a novel concept of such formalization supported by experiment based on elaboration of results of 124 relevant papers. The overall presentation is well structured and clear. Amount and quality of supplementary material is appropriate. However I have a doubt about method used. Not that I contest the learning from previous studies, but procedure suggested in the paper contain too many debatable aspects.

Principal problems i) Authors suggest to replace a deep functional analysis of applica-

tion context by the method based on learning from various previous solutions regardless of their detailed knowledge. OK, deeper functional analysis can be too difficult and selection of only some elements of application context can be a solution. However selection of used attributes and similarity functions was reasoned only poorly and in no way verified. ii) Presumption that articles published in good journals are supposed to provide good solutions for their specific study areas based on experts' experience and knowledge of the target task can be justified in general, but it is probably too optimistic in some cases even considering that determination of drainage network is probably only marginal problem for a part of articles. So no every solution published in good journal have to be well. And therefore a method based on selection the only one 'exemplary' published solution I feel as problematic. iii) While the suggested computation of similarity of individual attributes between the new application and published one can be acceptable, the synthesis (computation of 'overall similarity') is more problematic. No (equal) weighting of used attributes is a basic problem. It is very improbable that similarity in name of target task, cell size, area, relief, slope distribution and hypsometric integral will have the same effect on determination of proper catchment area threshold for extracting drainage networks. iv) Evaluation of experimental results is very problematic. Authors write (23-25, p.13): "Four levels of E were established empirically to reflect the reasonableness level: reasonable (E∈[0,0.1]), acceptable (E∈(0.1,0.25]), questionable (E∈(0.25,0.5]), and unreasonable (E∈(0.5,+∞))." It is non committal for me and if authors do not specify this 'empirical establishment' I feel it as fully subjective division. Why the difference in drainage density is unreasonable only exceed 50 %?! It smell by purpose made establishment of intervals to show "that the proposed method performs satisfactorily" (9, p.14).

Some another problems / ambiguities - The title of the paper is too complex and not quite clear. A simplification is suitable (e.g. Case-based formalization and reasoning method for digital terrain analysis – determining the catchment area threshold for extracting drainage networks). - Because equal weight of all attributes the binary attribute 'the name of the target task' exclude (in final comparison) all cases with another name

of the target task. What is the reason of such hard limit? How was determined the attribute for particular cases? Names of types and their occurrence should be added for better understanding. - Attribute relief - is it one number for the whole area (then it very dependents on area size) or average value computed by what way? (moving window - the size and shape?) - Slope is scale dependent variable so distribution of slopes depend on grid size. Using of cumulative frequency distribution solve this problem only partially. - Similarity functions seem to be determined subjectively. Why difference in magnitude of cell size (and area) can better reflect the level of similarity between DTA applications than the numerical difference in cell size? Why is used natural log in one case and common in another? etc.

Final evaluation and suggestion In regard to aforementioned problems I cannot recommend the paper in the present form, (presented experiment is not enough documented to support the interpretations and conclusions). However, majority of problems could be eliminated by selection of more appropriate method of synthesis. I think, multidimensional regression is a way. This method provide for elimination of inappropriate possible influence of particular problematic published case studies (ii), reveal various weights (suitability) of used attributed and similarity functions (mainly if hierarchical partitioning will be used) (iii) and last but not least alternative results (using various attributes and methods of similarity computation) can be compared to find the most appropriate regression equation. Suitability of selected attributes and methods can be documented by this way (i) and it can partly also substitute problematic way of evaluation in this paper (iv).

Please also note the supplement to this comment:
http://www.hydrol-earth-syst-sci-discuss.net/hess-2015-539/hess-2015-539-RC2-supplement.pdf