Geosci. Model Dev. Discuss., 5, C311–C315, 2012
www.geosci-model-dev-discuss.net/5/C311/2012/


Geoscientific
Model Development
Discussions



# *Interactive comment on* "TopoSUB: a tool for efficient large area numerical modelling in complex topography at sub-grid scales" *by* J. Fiddes and S. Gruber

**E. Gutmann (Referee)**

ethan.gutmann@gmail.com

Received and published: 21 June 2012

Fiddes and Gruber present a new method of representing topographic spatial variability within a land surface model. This is a novel approach based on a clustering algorithm, and the results look very promising. In addition, the paper is well written with clear figures, albeit there needs to be more description in some sections. I think this paper should be accepted with some minor modifications. Primarily, the authors need to expand their methods section and discussion section.

Model description: There is only a limited description of the LSM they are using (GEOtop), while they do supply a reference, a paragraph or two describing the model

would be appropriate. e.g. Does it use a Temperature index snow model or a more physics based model? Does it have an explicit vegetation canopy? How thick is the soil and how many layers in the model, do the two covary?

A little more discussion of how the "informed clustering" is performed would be helpful. Do they run the training routine for multiple realizations of the K-means clustering algorithm or just run it once for training purposes?

Forcing data: They state the forcing data come from one met. station. While the quality of that station is not terribly relevant to this paper, it would be useful to know what the climate for this station and surrounding area is like. Mean annual precip, air temperature, peak snow depth/SWE, etc.

They allude to the fact that they adjust the forcing data spatially (e.g. temperature follows a mean lapse rate, and SWin varies, presumably purely as a function of the cosine of the angle of incoming solar radiation on topography?). How are these adjustments performed? Is this part of GEOtop or TopoSUB or both? Do they adjust precipitation as a function of elevation? Do they adjust relative humidity to match the air temperature adjustment or does their model use a mixing ratio or specific humidity input?

This is important because it is useful to understand what is driving the improvement in their model. Is improvement caused by a more realistic precipitation field? SW field? temperature field? All of the above? Such discussion of the importance of different variables would be useful as it would help guide future research. In addition, assuming that the fully distributed version of their model is using the same adjustments, then any improvement in their model is dependent on how these adjustments are treated. If, precipitation is not adjusted for elevation in this study, then another study (that did adjust precipitation) might get a different answer (ie the number of required iterations/samples/starts might change.)

These adjustments to the forcing data are part of if not the entire reason their model exhibits spatial variability, as such understanding how they are adjusted is a critical



GMDD

5, C311–C315, 2012

component of this paper.

Test simulations: Were the two years of spin-up run by looping over the single year of forcing data (2009-2010). Do they start on January 1? If so, how do they initialize snow?

Results (and Discussion?) The discussion describing what their results mean is very sparse, they should expand on every sub-section.

Ideally, I would also like to see some discussion of the temporal error characteristics. I think most of their discussion revolves around the mean errors (presumably over both space and time.) However, there is no discussion of the temporal errors. I realize their method is primarily focused on spatial characteristics, but it is likely to have implications for the temporal evolution of the model as well. For example, how does their model influence the timing of spring snow melt? They touch briefly on the spatial error characteristics with figure 10, but here too, more discussion of this figure would be helpful instead of just presenting it. Where/when does the TopoSUB model do better or worse? Why? Given that the characteristics of e.g. the GST maps are so similar, wouldn't a map of the differences be helpful?

section 5.4 Model Stability: The authors state that a significant increase is seen in model stability between 25-100 samples; however, the authors should put these errors in the context. While the model is indeed "more stable" I would suggest that the stability is within reasonable bounds already at 25 samples. At this point the errors are negligible (GST<0.15deg C,SWin<2W/mˆ2, SWE<4cm,airT<0.05deg C). One would see an order of magnitude more variation than those numbers using different land surface models, different parameters, or different forcing variables (measured or modeled). This is important because it says that their model only requires $\sim$25samples (and perhaps fewer) to achieve model stability. Although, figures 6 and 8 show that $\sim$200 samples are required to achieve minimal errors relative to the BASE case, as a result, with reasonable sample numbers to minimize errors, stability should almost



GMDD

5, C311–C315, 2012



never be an issue, so any work to improve the K-means clustering algorithm will have minimal payback at this point.

Minor comments:

p1045 l4: PREDS: The authors might want to add a few examples of their predictor variables here, instead of making the reader look in the table for them. l7: TVs: Same thing here, it would be useful to see what they are predicting, along with a description of how it is predicted, e.g. via the LSM or via a direct topographic interpolation (e.g. air temperature lapse rate adjustment and solar radiation cosine(?) adjustment.)

Throughout the paper they should select one naming system and stick with it, e.g. is it the "BASE" model, or the "Distributed"" model? Is it the "Lumped" model or the "SUB" model? Or is the lumped model the distributed model run at very low resolution?

Figure 9: This figure would be easier to interpret if the x and y axis had the same ranges, and/or if a 1:1 line was presented... I'm assuming the line on the graph is a regression line since they list r values (also, I would prefer to see rˆ2 values as rˆ2 has a more readily interpretable meaning as the percent of variance explained.) However, because the line is very close to 1:1, I can't tell if there is any bias. Do these data points come from a single time slice, or all are they data points in space and time? (also same comment as above, Is "DIST" the same as "BASE" and "LUMP" the same as "SUB"?).

Figures 9,10 need units on their axis / colorers.

An additional figure showing a map of some of the clusters would be useful though not necessary. I realize it would be difficult in the fuzzy case. Perhaps maps could be presented for "crisp" membership for some of the key points mentioned in e.g. figure 8 with n_samples=16, 64, 258 (and should that be 256?)

Thank you for a well thought out paper, it was enjoyable to read and I look forward to seeing the updated manuscript.

regards, Ethan Gutmann



GMDD

5, C311–C315, 2012

Interactive comment on Geosci. Model Dev. Discuss., 5, 1041, 2012.



GMDD

5, C311–C315, 2012