Hydrology and
Earth System
Sciences

Discussions

EGU
Open Access

Interactive comment on "Harnessing Big Data to Rethink Land Heterogeneity in Earth System Models" *by* Nathaniel W. Chaney et al.

Anonymous Referee #2

Received and published: 6 December 2017

Authors have implemented a hierarchical multivariate clustering approach to represent land surface heterogeneity in an Earth System Model Grid cell. The approach is taking advantage of fine resolution datasets from various sources in addition to a Digital Elevation Model (DEM) to identify a set of characteristics hillslopes through a multivariate clustering algorithm. Each characteristic hillslope is then discretized into height bands, and a set of tiles are delineated for each band to represent heterogeneity of soil and land cover types.

While the topic of this paper is of great interest for the global hydrologic modeling community, no attempt has been made to assess the performance of these simulations against observations. Indeed, fine resolution model simulations are used as the truth. I suggest authors to take advantage of various satellite products to define what scale of

heterogeneity needs to be incorporated in each ESM grid cell to better represent land surface states and fluxes.

Authors have performed a limited sets of sensitivity analysis to assess the impacts of height band length and number of clusters on simulated fluxes. However, no insights have been provided in order to define these parameters in various landscapes. Further discussion will be really helpful to inform the modelers.

No physically based approach are presented for delineating the height bands. Can authors implement a more physically based approach for defining the height bands? For example, the methodology of Khan et al. 2013 (Environmental Modelling and Software) for delineating landforms can be informative in this case.

Another issue is regarding the number of variables that are required for a multivariate cluster analysis. I suggest authors to perform a sensitivity analysis to identify factors that are most influential in defining these clusters.

Further descriptions about the discretization methodology is required particularly those that relate to Figure 4. I suggest authors to prepare a flowchart that explains every step from discretization to mapping back the results to the fine resolution.

Further clarification regarding the simulation elements are required.

Another major point is how does the surface and subsurface connectivity between the tiles within a single band and between bands are maintained? What rules do you implement?

How tiles are represented for model simulations? Does each tile represented by a point scale simulation?

P6-L20: The authors indicate that attributes of each characteristic hillslope is obtained through arithmetic averaging. How do you handle categorical data like soil type and vegetation types in this case?

Section 3.2.1. What is a typical size of a characteristic hillslope?

Section 3.2.2. L 30: Needs further explanation.

Can authors provide further insights and general recommendations for implementation of their approach in other geographic regions?

Can authors provide further information about the enhancements made in this approach compared to the earlier work of delineating HydroBlocks?