Author reply

Joe Melton, Reinel Sospedra-Alfonso, Kelly McCusker

April 24th 2017

Dear Editor,

We thank H. Zheng, E. Blyth and the two anonymous reviewers for their time and care in reviewing our paper. We believe that we can address their concerns in our revision. The reviewer comments are repeated below in their entirety with our replies in italic font. We have appended a marked up version of our manuscript showing each change made to make it easy to find the revisions that we have mentioned in the following text.

Reviewer 1 : H. Zheng

This paper represents valuable efforts of incorporating sub-grid soil heterogeneity in a coupled land surface–ecosystem model, CLASS-CTEM. Instead of a mean value in each large CLASS-CTEM grid cell, several soil texture tiles are identified from a fineresolution soil texture dataset using a clustering algorithm. The effects of tiling soil texture on modeled soil moisture, net ecosystem exchange, respiration, and carbon storage were examined. Although the authors state that "soil textures appear to be reasonably represented for global scale simulations using a simple grid-mean value," the results could still benefit the modeling community. The paper is well-written and clearly aligned with the goals of the Geoscientific Model Development Journal. I recommend its publication with a major revision.

We thank Dr. Zheng for the positive comments.

General comments:

1. I am willing to see the modeling results when minPts is set to 1%. The purpose of using soil tiles is to represent spatial heterogeneity. However, by setting minPts to 5%, the most heterogeneous regions (e.g. U.S. and China in Figure 6b) turn out to be the regions with the least numbers of soil tiles (Figure 6a). As a result, I don't think 5% is a reasonable threshold. On the other hand, in your ideal case, there is strong assumption that the

fine-scale soil texture distribution is symmetrical. This assumption may not hold in the reality when the soil is heterogeneous.

Reply: A run of the clustering algorithm with a 1% threshold was already included in our original manuscript(MS) (discussed on page 8 line 13 and Figure A4). As can be seen in the figure A4, the number of clusters found in the USA and China does increase as proposed. However the number of clusters also increases broadly across the entire domain.

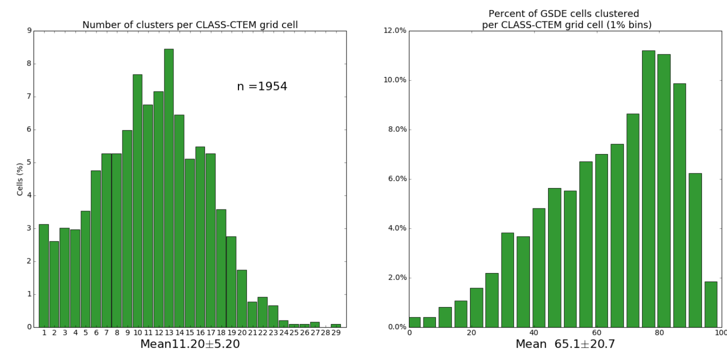

Figure 1: Number of clusters determined for each CLASS-CTEM gridcell using a 1% threshold. Compare to Fig A3 in the MS.

The mean number of clusters found using a 1% limit increases to over 11 (compared to around 3 using our previous 5% threshold) while the mean percent of GSDE cells clustered (became part of a tile) increased to 65.1 ± 20.7 % compared to our clustering using 5% threshold (57.0 ± 20.1 %). Generally, clustering using a minPts lower limit of 1% does not result in a gain in information for the model. Clustering with the 1% limit causes a large increase in the number of CLASS-CTEM gridcells with >8 tiles (Fig 1 above and Fig A4 in MS) and, as demonstrated in the MS Fig 1 (and also discussed below in reply to Reviewer 2), the model is generally insensitive to >8 tiles per gridcell. As a result a 5% threshold appears to be reasonable for most locations

For the comment regarding our ideal test case, yes the assumption of the soils being sand and clay mixtures is not likely to be very realistic but the point of this test was twofold: first, to ensure the model was sensitive to soil textural changes, and second, to determine if the model threshold to soil texture changes has some limiting value. Indeed as we can see in the MS's Fig 1 after around 7-8 tiles the model ceases to be sensitive to increased number of tiles. However, in addressing a comment for Reviewer 2 below, we also tested the impact of the grid cells chosen for this test and found that the model sensitivity limit probably differs slightly between gridcells with the driest regions showing sensitivity up to around 12 tiles while wetter tropical and temperate regions are around 7-8. The *arid cells with higher sensitivity are arguably less important globally in terms of water and carbon fluxes.*

Some specific comments:

1. P4L4, NCEP = National Centers for Environmental Prediction.

Reply: Thanks - corrected.

2. P7, L16 and L18. What is “pore content”? Air?

Reply: Yes, now changed to ‘pore (air) content’

Reviewer 2: Anonymous

In this study, the authors investigate a tiling approach for soil textures in CLASS-CTEM with a clustering analysis to identify representing soil textures. It is an interesting study and provides useful information to land modelers although the effect on the terrestrial carbon cycle is relatively small. There are some minor issues that I would like the authors to explain before the manuscript can be accepted for publication.

Reply: Thanks, we are glad you found the work interesting.

1) Can you explain more about the treatment of surface heterogeneity in your model? Is the subgrid variability of PFTs represented by a single tile?

Reply: We took the PFT distribution at the grid cell level and gave each tile the same distribution (see line 34 page 5). For example, if the CLASS-CTEM grid cell had 30% needleleaf evergreen tree, 50% C3 grass and 20% bare ground coverage, each tile would have that same PFT distribution applied to it. We have added an example like this to the text to make sure it is clear. While it is outside the scope of this present study, it would be interesting in the future to try and distribute the PFT coverage to the same tiles as its underlying soil textures, as Reviewer 3 (below) suggests.

2) Why is the sum of the weight percent soil textures less than 100 %?

Reply: The total soil texture is taken to be sand + clay + silt = 100%, thus any remainder is the silt content of the soil. We added this explanation to the revised MS.

3) It seems that the effects on C cycle from the sensitivity test (section 3.1) is larger than the realistic simulations (section 3.2, 3.3). Why?

Reply: The sensitivity test was taken to be an extreme case to help us understand the upper limit for model sensitivity to the tiling. The chosen mineral soil textures - either sand or clay - are representative of the most extreme textural changes the model could experience and thus should produce a large change in the model conditions.

4) Do the differences depend on the vegetation type?

Reply: We have rerun this test for two more sites. The first site is the very dry Sudan site from our MS and the second is a wet tropical site from the Amazon region (-1.40 S 56.25 W). The original site tested is in the N.E. USA (43.3 N 92.8 W). Figure 1 in the appended MS shows the new results. Yes, the magnitude *of change does depend on the vegetation type and the general climate of the gridcell. Both the NE USA site and the Amazon site have the highest productivity with a single tile of 50% sand and 50% clay while the Sudan site does much better with two tiles as one tile is then 100% sand (with the other being 100% clay) and that sand tile allows more plant available soil water similar to the Australian test site highlighted in MS Fig 9. Both the N.E. USA and Amazon sites generally have sensitivity up to about 8 tiles while the Sudan site appears sensitive to around 12 tiles. This does seem sensible since our results generally showed a much stronger proportional response to the tiling of soil textures from arid regions. We have added this expanded test along with discussion into the revised MS for Section 3.1. (Please see the appended MS for the changes).*

5) GPP abbreviations need to be explained at line12 on page7.

Reply:NPP, NEP, and HR were defined at the bottom of page 6.

6) Table 1. Are there any differences in individual components of evaporation (i.e. soil evaporation, transpiration) or runoff (surface runoff and base runoff)?

Reply: This question was also raised by another reviewer, and is a very valid point. We have now added plots showing the impacts upon the hydrologic cycle to the revised MS (Figs A7 - A9) along with discussion in Section 3.3.2. They are also discussed in more detail in the reply to Reviewer 4 (Eleanor Blyth) below.

7) Figure 9. Is theta_l in tile E larger than in other tiles? If so, can you explain why? Is it related to the representation of runoff processes?

Reply: No, tile E does not have a larger theta_l. We have replotted Figure 9 with the soil water content (theta_l) instead of plant available soil water (Fig 2 below). Tile E has the lowest soil water content while, as can be seen in the original MS Fig 9, it has the highest plant available water.

8) How much does the runtime increase in the global simulation? The variable number of tiles may effectively represent surface heterogeneity while saving computational resources.

Reply: This is an apt point. The offline CLASS-CTEM model timing scales roughly linearly with the number of tiles, but not additively (see dashed line in MS Fig 1; slope around 0.5). As a result, we are interested, as you say, to represent surface heterogeneity in a computationally efficient manner. This is also part of our motivation for why we first tested model sensitivity to the number of tiles. We could easily increase the number of tiles to a very large number but we would not gain any further information for the model and we would greatly increase the computation time. We added some text around this point as it was not emphasized in the original MS version.

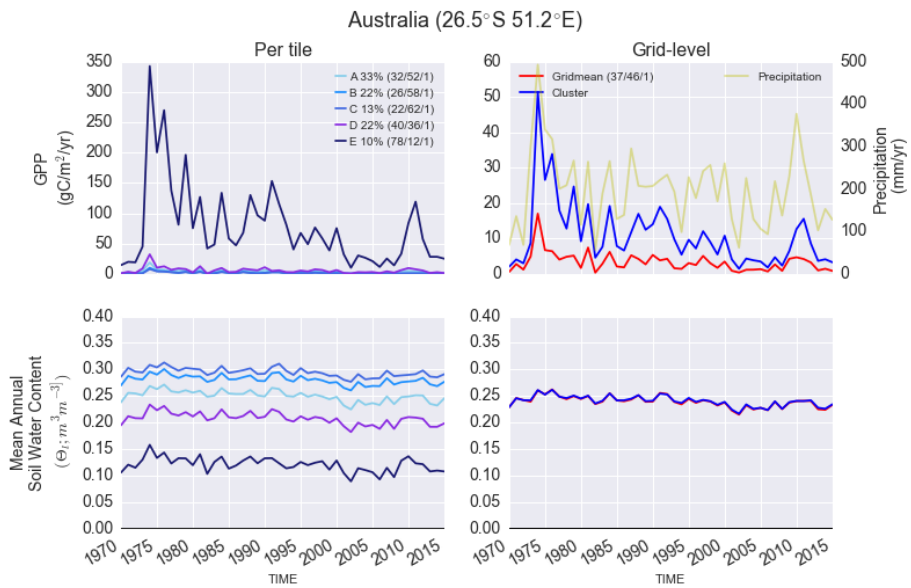

Figure 2: Figure 9 replotted with soil water content instead of plant available soil water.

Reviewer 3 : Anonymous

General comments:

The authors examine a clustering algorithm to represent sub-grid heterogeneity of soil texture for a coupled land-surface-ecosystem model (CLASS-CTEM). A newer clustering algorithm is applied to the CLASS-CTEM model soil texture inputs to define the sub-grid tiles in the model. The algorithm has not been previously applied to landsurface/hydrology/ecosystem models for sub-grid tiling and has the potential advantage of determining the number of clusters from the data; no a priori number of clusters needs to be defined. The CLASS-CTEM model shows sensitivity to tiling of soil texture in a synthetic experiment and two grid point simulations. An equilibrium is reached at around 7-8 tiles in the synthetic experiment. They then run global offline simulations.

Overall, clustering by soil texture using the algorithmic parameters discussed in the paper results in roughly 3-4 tiles per grid cell across the globe, with some cells containing up to 7-8+ clusters. The model does not show much sensitivity to ecosystem parameters (e.g. gross or net productivity) across the global uncoupled simulations, except in some arid/semi-arid regions along the margins of deserts. The paper is very easy to read and logically organized. I think the application of the new clustering algorithm is useful to the community and may have broader applicability; the experiments give insight into this particular clustering method-model combination. However, I have some major concerns. If the editor believes the authors can address them in the allotted time, then it should be acceptable for publication in GMD after revisions, otherwise I would suggest rejection with resubmission.

Reply: Thank you, we are glad you found our paper easy to read and of potential use to the community

Major Comments:

1) Lines 24-25 on page 2 (Introduction) mentions water balance partitioning (runoff vs ET) and then lines 17-18 on page 3 mention the offline run conclusions should hold for the coupled runs because grid-mean fluxes are passed back to the atmospheric model. However, there is no mention or presentation of any latent or sensible heat fluxes anywhere in the paper. It is very likely that the grid-mean values may be different in the regions of large net ecosystem productivity changes, which would then affect coupled simulations. I am not expecting fully coupled runs here, but presentation and discussion of flux partitioning changes (or lack thereof) since the primary function of the land-surface model is flux partitioning.

Reply: This is a good suggestion, thank you for it. We initially chose to focus on the carbon fluxes as that is our primary area of interest but we are happy to expand on the flux partitioning behaviour of the clustered vs. gridmean runs.

We now include the differences in sensible (new MS Fig A10) and latent (Fig 3 below) heat fluxes by season as well as the water use efficiency (WUE)(defined as GPP / evapotranspiration; new Fig A11). As in the MS figures, the dots indicate grid cells that are statistically significant (independent two sample t-test p-level <0.01). We have added plots to the revised MS and also updated Table 1. To incorporate the WUE and heat fluxes we have changed Section 3.3.2 extensively as can be seen in the appended PDF showing the changes in the manuscript. For the point regarding the impact of the tiling on coupled simulations, because it is difficult to tell apriori what impact the tiling might have, we have removed that statement (previously lines 17-18 on pg 3).

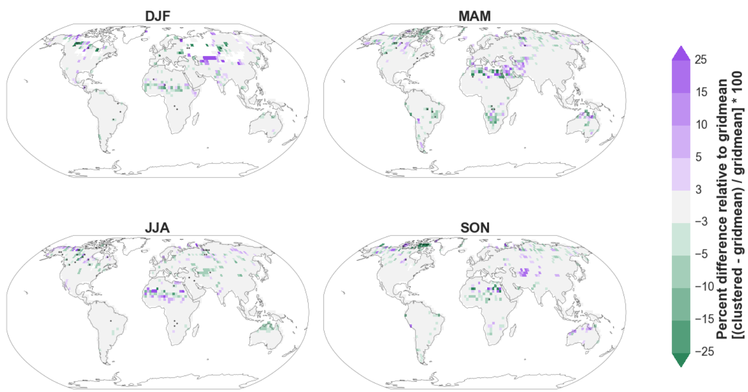

Figure 3: Percent difference relative to gridmean for latent heat fluxes. Positive values indicate the Cluster values are larger while negative values indicate the Gridmean values are larger.

2) Section 3.1: This section needs to be discussed in more detail. What was the vegetation distribution for this synthetic test? Does the vegetation distribution influence the sensitivity? If it does, perhaps another test would be useful. Does this then impact the equilibrium number of clusters?

Reply: This question was also raised by Reviewer 2 and is discussed above in detail in reply to their comment 4. To address the vegetation question directly, the vegetation at each of the now three test cells was taken to be the typical land cover for that grid cell used for our global runs. We were most interested in using this test to: first, ensure the model was sensitive to soil textural changes, and second, to determinine if the model threshold to soil texture changes has some limiting value. As a result choosing three locations with greatly different climate and vegetation should give a reasonable answer to those two primary questions. Those answers being, yes, the model is sensitive and it is likely around 7-8 tiles for most regions with arid regions closer to 12 tiles. So, in answer to the questions posed, yes, the equilibrium number of clusters is somewhat dependent upon location (climate and vegetation). In an ideal situation, one would only run the model with the bare minimum number of tiles to capture the important sub-grid heterogeneity as each additional tile carries a computational cost (dashed line in MS Fig 1).

3) The density-based approach of this clustering algorithm could be problematic. If only 57% of the high-resolution data are being included in clusters on average, you are effectively letting the algorithm determine that 43% of the data may not matter. It is particularly worrisome to see many grid points with clusters that contain only 10-20% of the high-resolution data. Just because the grid-mean texture values and these cluster values agree is not a good enough reason to say the clustering algorithm is successful. The "outliers" or lower-density regions may be important to ecology and/or hydrology. Seyfried et al. (2009) discuss hydrologic hotspots, portions of a catchment that are small but have disproportionate impacts on runoff. The same is seen here in Figure 9 for gross primary production (GPP). If the clustering is not considering areas of a grid cell that may be small individually (yet add up to a very significant portion of the grid-cell total in many cases), it is possible that some sub-grid variability is being missed (or possibly over-emphasized). For example, what percent of the large grid-cell in Figure 9 is classified on the high-resolution grid? How much of the high-resolution data is similar to (or higher) in sand content than cluster E? If the total area of the large grid cell has more(less) than 10% of the high-resolution points similar to cluster E (from expert evaluation), the grid-cell mean impacts of tiling may be under(over)-estimated.

Reply: The test of comparing grid-mean soil textures to the mean of the cluster values is a simple approach to ensure that the clustering algorithm could not heavily bias the soil texture for the grid cell as a whole. We agree that clustering *techniques are not ideal, nor are they infallable. However identifying "outliers" that may be important to ecology and/or hydrology requires expert assessment on a gridcell by gridcell level (It is also unclear how these hotspots would be identified based only on soil textural information found in global-scale datasets). This amount of intervention is not feasible considering the ca. 2000 grid cells globally. The use of a clustering analysis such as OPTICS allows an automated means to identify possibly important areas of similar soil textures, but the analysis can only rely upon the information given (in this case, soil textures). We know of no other algorithmic method to determine which high resolution gridcells are important to represent in the model while aiming to keep the computational cost of the model runs manageable, i.e. running as few tiles as can adequately represent the influence of soil heteogeneity.*

To clarify, the clustering algorithm does indeed consider all grid cells from the high-resolution dataset that fall within a CLASS-CTEM grid cell and if the high-resolution grid cells add up to a significant portion of the CLASS-CTEM grid cell total, they will be represented as a cluster. Many high-resolution grid cells with similar soil properties present regions of higher density to the clustering algorithm and would then be represented in the tiling.

Regarding the grid cell for the Australia example (MS Fig 9), that cell had 52% of the high resolution soil texture cells included in the five clusters. Below, we have plotted up that example site similar to MS Figures 2 and 3 (Figs 5 - 7). It is evident that the site has a highly heterogenous soil texture for sand and clay. Given the high heterogeneity, the determination of five clusters for this site, with their compositions evident in the figures below, appears reasonable (compare Fig 5 to Fig 7 below). Tile E (78% sand, 12% clay) attempts to capture the most sandy high-resolution gridcells.

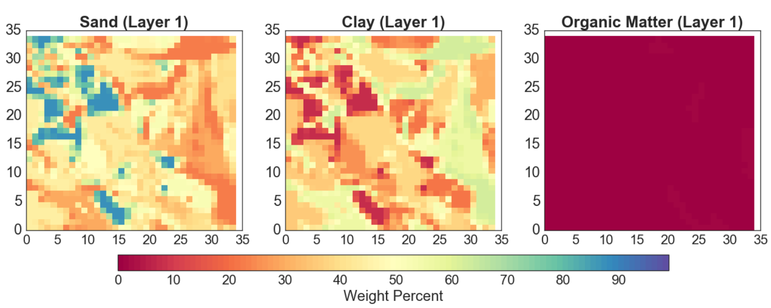

Figure 4: Australia example site from MS Fig 9. Composition of the high resolution soil textures from GSDE for the larger CLASS-CTEM grid cell (similar to MS Figs 2 and 3 top panel).

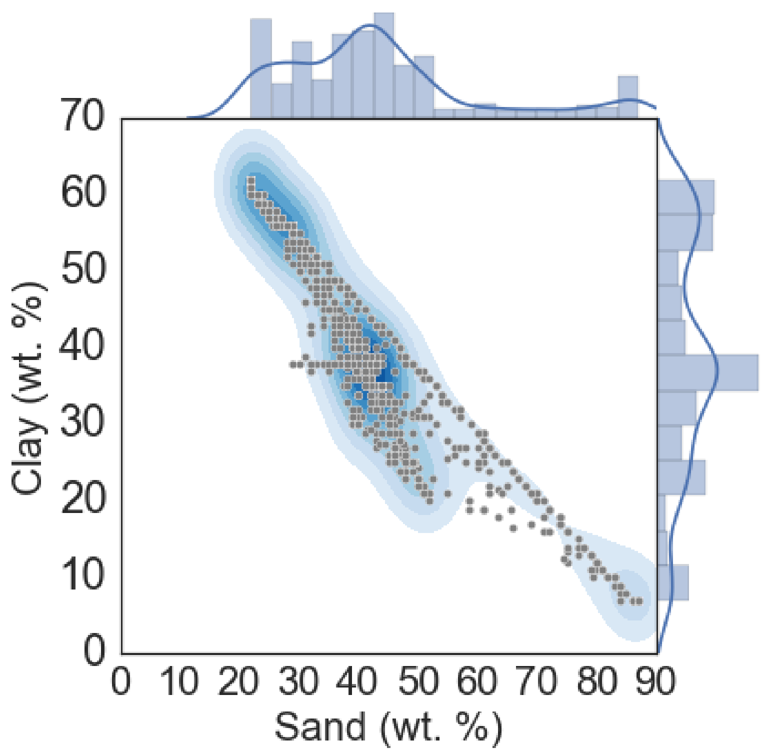

Figure 5: Australia example site from MS Fig 9. Joint distribution using kernel density estimation for the soil and clay content (similar to MS Figs 2 and 3 middle panel).

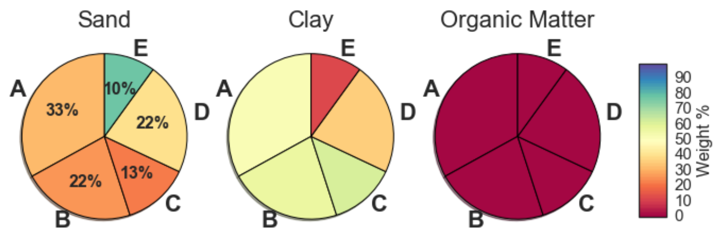

Figure 6: Australia example site from MS Fig 9. Pie charts showing the tile percent cover and composition (similar to MS Figs 2 and 3 bottom panel).

4) Why is vegetation considered constant across clusters (lines 1-2, page 6)? The GLC2000 dataset could have been mapped to the clusters to have variable PFTs with each tile. Soil texture and vegetation likely co-vary in some or many locations; it would be good to capture that relationship if present.

Reply: We agree this would be an interesting area for further study. While it is outside the scope of our present work, it would be possible to map the PFTs to the soil grid and thus associate PFTs with the cluster their underlying soil was mapped to. This is not overly straight forward since the remotely-sensed vegetation categories are not the same as the model PFTs but it would be very interesting to see if any improvement or greater influence could be detected. For this paper, since it was the first attempt, we tried to keep it simple to allow an easier understanding of how sensitive the model was to clustering soil textures. Our method could also possibly be extended to include other geophysical fields like soil depth, soil colour, slope, etc. Each new geophysical field clustered could allow, perhaps, a better representation of subgrid heterogeneity. We have also added some recommendations on future work to the conclusions section.

Minor Comments:

1) Page 5, line 27: Why 10%? Also, this is a 10% change relative to the grid-mean value?

Reply: Yes, this 10% change was relative to the grid mean value. The 10% limit was found via a sensitivity test and was chosen to allow some deviation from the grid mean but not to an extent where the tiling was producing a grid mean composition that was radically different. We looked at how many CLASS-CTEM grid cells were greater than this threshold in MS Figure A1.

Reviewer 4 : Eleanor Blyth

I enjoyed reading this paper - it is extremely well written and logical in its presentation. I think modellers will find themselves informed by it. However, I came away somewhat frustrated not to hear more about the other aspects of the model that might have been affected by this tiling approach. For instance, I would like to see plots of the soil moisture, the evaporation and the drainage (or groundwater recharge) from the models. These land-surface models such as CLASS are increasingly not only used for ecosystem dynamics and soil tiling might make more difference to other applications such as hydrology (droughts, groundwater recharge). This might light up the paper a little especially since the impact on GPP was so small. For instance, you might find a bigger story to tell with drainage since it is non-linearly related to soil moisture. I think a few extra plots should be enough.

Reply: Thank you Dr. Blyth for your comments. Yes, we agree that it would be good to expand the variables examined in the paper. To be perfectly honest, we (primarily the lead author) worried that, since the effect overall was modest, it was better to keep the paper short and succinct to avoid boring a reader. Thankfully, each reviewer seems to have found the study worthwhile and thus we are happy to expand upon the original manuscipt to look at these other variables. We originally did look at soil moisture (MS Fig 7) but have now included runoff (revised MS Fig A9), evapotranspiration (revised MS Fig A7), transpiration (revised MS Fig A8), as well as sensible and latent heat fluxes and water use efficiency (revised MS Figs A10, A11).

While the global effect of clustering on runoff was relatively small (MS Table 1), we were curious if the changes would become more substantial if we looked at them grouped by large river basins as revised MS Fig A9 shows that it can be relatively large on a grid cell level. The mean annual runoff from the model was compared to river discharges (Arora, 2001) (Fig 7). This was a very simple comparision where the runoff was not routed thus we assumed that all runoff estimated by CLASS-CTEM in a river's drainage basin made it directly to the river mouth (ignoring evaporation in streams, loss to and contributions from groundwater, etc.). The improvement from clustering is relatively small on the whole with the standard error essentially unchanged (indeed it takes to the fifth decimal point to determine a difference; Gridmean 0.037099 vs. Cluster 0.037044, using SciPy stats linregress function). Because this is such a rudimentary comparison, we don't intend to include this in the revised MS, but we include it here for the reviewer's interest. We now include discussion and figures for sensible and latent heat fluxes, WUE, evapotranspiration, transpiration, and runoff in the revised MS. (Marked up MS provided below)

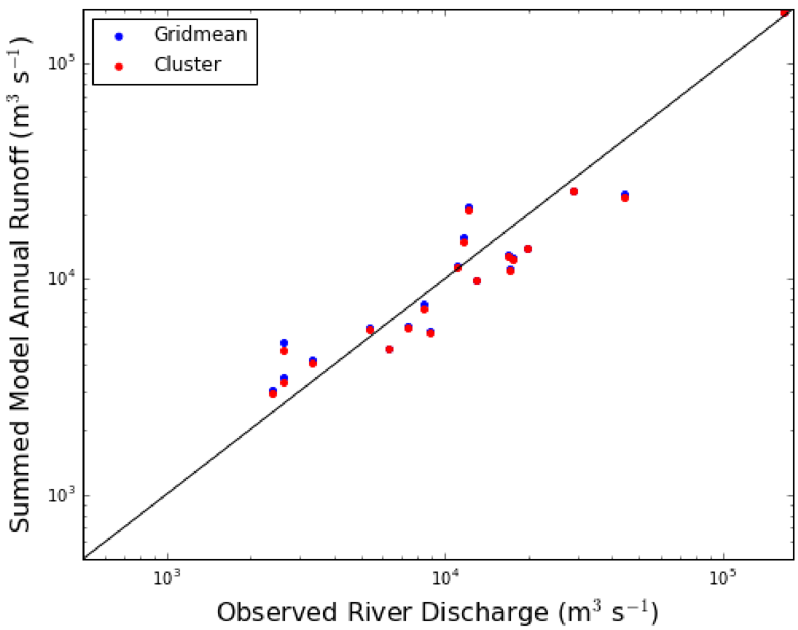

Figure 7: Simplistic river discharge comparison against model runoff. Each point represents a major river basin.

Bibliography

Arora, V. K.: Streamflow simulations for continental-scale river basins in a global atmospheric general circulation model, Adv. Water Resour., 24(7), 775–791, doi:10.1016/S0309-1708(00)00078-6, 2001.

Tiling soil textures for terrestrial ecosystem modelling via clustering analysis: a case study with CLASS-CTEM (version 2.1)

Joe R. Melton[1], Reinel Sospedra-Alfonso[2], and Kelly E. McCusker[2,3]

[1]Climate Research Division, Environment and Climate Change Canada, Victoria, B.C., Canada
[2]Canadian Centre for Climate Modelling and Analysis, Climate Research Division, Environment and Climate Change Canada, Victoria, B.C., Canada
[3]School of Earth and Ocean Sciences, University of Victoria, Victoria, B.C., Canada. Now at University of Washington, Department of Atmospheric Sciences

Correspondence to: Joe Melton (joe.melton@canada.ca)

Abstract. We investigate the application of clustering algorithms to represent sub-grid scale variability in soil texture for use in a global-scale terrestrial ecosystem model. Our model, the coupled Canadian Land Surface Scheme - Canadian Terrestrial Ecosystem Model (CLASS-CTEM), is typically implemented at a coarse spatial resolution (ca. 2.8° $\times$ 2.8°) due to its use as the land surface component of the Canadian Earth System Model (CanESM). CLASS-CTEM can, however, be run with tiling of the land surface as a means to represent sub-grid heterogeneity. We first determined that the model was sensitive to tiling of the soil textures via an idealized test case before attempting to cluster soil textures globally. To cluster a high-resolution soil texture dataset onto our coarse model grid, we use two linked algorithms (OPTICS (Ankerst et al., 1999; Daszykowski et al., 2002) and Sander et al. (2003)) to provide tiles of representative soil textures for use as CLASS-CTEM inputs. The clustering process results in, on average, about three tiles per CLASS-CTEM grid cell with most cells having four or less tiles. Results from CLASS-CTEM simulations conducted with the tiled inputs (Cluster) versus those using a simple grid-mean soil texture (Gridmean) show CLASS-CTEM, at least on a global scale, is relatively insensitive to the tiled soil textures, however differences can be large in arid or peatland regions. The Cluster simulation has generally lower soil moisture and lower overall vegetation productivity than the Gridmean simulation except in arid regions where plant productivity increases. In these dry regions, the influence of the tiling is stronger due to the general state of vegetation moisture stress which allows a single tile, whose soil texture retains more plant available water, to yield much higher productivity. Although the use of clustering analysis appears promising as a means to represent sub-grid heterogeneity, soil textures appear to be reasonably represented for global scale simulations using a simple grid-mean value.

The works published in this journal are distributed under the Creative Commons Attribution 3.0 License. This licence does not affect the Crown copyright work, which is re-usable under the Open Government Licence (OGL). The Creative Commons Attribution 3.0 License and the OGL are interoperable and do not conflict with, reduce or limit each other.

©Crown copyright 2016

1 Introduction

Representation of sub-grid variability is a challenging problem in large-scale modelling applications such as Earth System Models (ESMs). ESMs are commonly run at relatively coarse spatial resolutions due to the computational costs associated with these complex models. The terrestrial component of an ESM is also commonly tied to the grid cell size or truncation level of the atmosphere, making it difficult to resolve smaller scale processes. Heterogeneity in precipitation, vegetation, soils (Boone and Wetzel, 1999), topography, and snow cover (Nitta et al., 2014) on spatial scales much smaller than common model grid cells can cause surface fluxes to vary non-linearly across a grid cell. To address this issue, several modelling groups have adopted either a tiling approach, in which each grid cell is divided into a mosaic of tiles with a different tile given for each landscape feature (Avissar and Pielke, 1989; Koster and Suarez, 1992; Essery et al., 2003), or a statistical approach whereby the sub-grid heterogeneity is represented by a probability density function (e.g. Famiglietti and Wood (1994); Pielke et al. (1991); Boone and Wetzel (1999)).

The division of a grid cell into tiles has been attempted for characteristics such as hydrological parameters (Wood et al., 1992; Arora et al., 2001), vegetation present (Molod and Salmun, 2002; Li and Arora, 2012; Melton and Arora, 2014; Ke et al., 2013), land cover change (Landry et al., 2016), precipitation (Arora et al., 2001), elevation (Ke et al., 2013), land surface properties (Avissar and Pielke, 1989), and maximum infiltration (Decharme and Douville, 2005). Many of these reports that tiled the land surface used relatively easily observed, and hence classified, characteristics of the landscape, i.e. vegetation presence/absence, elevation band, vegetation type, etc. To our knowledge the tiling of soil texture has never been reported. We hypothesize that the use of tiled soil textures, rather than taking simple grid-mean values, will result in more realistic model simulations due to the non-linear influence of soil texture on soil hydrological and thermal characteristics. Soil moisture is one of the most important determinants for partitioning of surface fluxes of moisture and heat from net radiation (Shao and Henderson-Sellers, 1996) and precipitation into evapotranspiration and total runoff (Dirmeyer et al., 1999), as well as having a strong influence on vegetation productivity, and the terrestrial carbon cycle, which is of primary interest here. Soil texture influences on plant productivity and community structure should be especially strong in regions with low water availability as has been observed in semi-arid and arid regions (Archer et al., 2002; Hook and Burke, 2000; English et al., 2005). To test our hypothesis, we use clustering algorithms on a recently released high-resolution soil textural dataset. Clustering analysis searches for patterns in datasets based upon their natural structure or grouping. Some examples of clustering analysis in the Earth system sciences includes remote determination of inundated areas (Prigent et al., 2001), land use management zones (Li et al., 2007), ecoregion delineation (Kumar et al., 2011), and fire regimes (Archibald et al., 2013). Given high-resolution soil textural information, a clustering analysis can determine regions of similar soil textures (e.g. river valleys, mountainous slopes) that are smaller than the size of ESM grid cells. The soil textures of these distinct regions can then be used as a tile to allow representation of this sub-grid heterogeneity in the model without requiring a smaller model grid. Newman et al. (2014) used k-means clustering analysis to determine tiles based, primarily, on the vegetation types present and thus were able to provide the k term (number of clusters) a priori. In clustering soil texture it is desirable to allow the number of soil clusters to vary per grid cell and not be specified a priori. Our study thus presents two new approaches: the use of a clustering algorithm to determine tiles that does not require a priori information on the number of tiles per gridcell and using tiled soil texture to represent sub-grid heteorogeneity.

In the following sections we, i) introduce CLASS-CTEM and the clustering algorithms (Section 2), ii) evaluate the soil textural tiles found by the clustering algorithms and the resulting CLASS-CTEM outputs against simulations that use a simple grid cell mean soil texture and against an observation-based dataset of gross primary productivity (Section 3), and iii) discuss these results and give conclusions for the utility of this approach (Section 4).

2 Methods

2.1 CLASS-CTEM

All simulations were run with the Canadian Land Surface Scheme (CLASS v. 3.6; Verseghy (2012)) coupled to the Canadian Terrestrial Ecosystem Model (CTEM v.2; Melton and Arora (2016)). Together CLASS-CTEM forms the land surface component of the Canadian Earth System Model (CanESM), but the simulations presented here were performed off-line to permit easier interpretation. ~~Since CLASS-CTEM uses grid-mean quantities for the coupling of the land surface to the atmosphere -even if the simulation used tiles for the land surface (see Fig. 1 in Melton and Arora (2014)) - the offline simulations general conclusions should be applicable for the coupled model as well.~~

CLASS operates at a half-hour timestep and performs the land surface water and energy balance calculations. In simulating the energy balance of the land surface and its interactions with the atmosphere, CLASS uses vegetation attributes such as leaf area index (LAI), canopy mass, rooting depth, and vegetation height. The temperature, and liquid and frozen water contents of three soil layers, of 0.1, 0.25, and 3.75 meters thickness, are determined prognostically. The CLASS parameterization for mineral soils follows that of Cosby et al. (1984) and Clapp and Hornberger (1978) (see Appendix). Organic soils, defined as those cells having an organic matter weight percent greater than 30, are modelled as peat following Letts et al. (2000). The daily mean soil temperature, soil moisture, and net radiation from CLASS are passed to CTEM at the end of each day. CTEM then calculates the vegetation and carbon dynamics. While most of CTEM operates on a daily timestep, the carbon assimilation from photosynthesis and canopy conductance occur on the CLASS timestep. CTEM calculates the carbon uptake and respiratory costs of nine plant functional types (PFTs) which map directly to four PFTs that CLASS uses. The CLASS PFTs (with corresponding CTEM PFTs in parentheses) are needleleaf tree (needleleaf evergreen and needleleaf deciduous), broadleaf tree (broadleaf evergreen, broadleaf cold deciduous, and broadleaf drought/dry deciduous), crop (photosynthetic pathway C3 and C4), and grass (C3 and C4). CTEM tracks carbon flow through the leaves, stem and roots of the living plants and the litter and soil C for the detrital pools. For global simulations, CLASS-CTEM is typically run at a grid cell resolution of approximately 2.8° by 2.8° corresponding to a grid cell size of ca. 98 000 km^2 at the equator and ca. 49 000 km^2 at 45°

latitude. CLASS-CTEM has been validated against observation-based datasets from site-level to global (e.g. Peng et al. (2014); Melton et al. (2015); Melton and Arora (2016)).

2.1.1 CLASS-CTEM simulation details

The simulations were forced with version 7 of the Climate Research Unit-National Centers for Environmental ~~Protection~~ Prediction (CRU-NCEP) meteorological dataset covering 1901 - 2015 (Viovy, 2016). The meteorological inputs are disaggregated from 6 hourly to half-hourly as laid out in Melton and Arora (2016). To spin up the model, the climate years 1901 – 1925 were repeatedly cycled over until the model reached equilibrium (which is defined by the net biome production simulated to be less than 0.1% of net primary productivity). During the spinup, the land cover and population densities (used by the fire disturbance scheme) were set to that of year 1850 with a global atmospheric CO_2 concentration of 284.87 ppm. After the spinup, the transient simulation ran from 1851 to 2015 with atmospheric CO_2 concentrations from Meinshausen et al. (2011). The land cover is derived from the Global Land Cover 2000 (GLC2000) data set for year 2000 (Bartholomé and Belward, 2005). The GLC2000 data is then mapped to the corresponding CTEM PFTs and we use the HYDE v. 3.1 data set (Hurtt et al., 2011) to change crop area with time. The distribution of the C3 and C4 photosynthetic types for the crops and grasses is based upon Still et al. (2003). To run from 1851 - 2015, the climate was cycled over twice from 1901 – 1925 for the years 1851 – 1900, then allowed to run freely from 1901 - 2015. All simulations had land use change impacts (prescribed changes in crop cover from 1851 - 2015) as well as fire disturbance.

2.2 High-resolution Soil Texture Dataset

The Global Soil Dataset for use in Earth System Models (GSDE) (Shangguan et al., 2014) is available at 5 arc minute resolution from http://globalchange.bnu.edu.cn/research/soilw (Accessed July 23rd, 2015). GSDE has eight soil layers of depths: 4.5, 9.1, 16.6, 28.9, 49.3, 82.9, 138.3, and 229.6 cm. CLASS-CTEM's requirements for soil textural information include weight percent sand, clay, and organic matter (OM) as well as soil depth (Verseghy, 2012). We retain CLASS-CTEM's typical soil configuration of three soil layers with layer bottom depths of 10 cm, 35 cm, and 410 cm. The soil silt weight percent is found taking the remainder of 100% - sand - clay.

In each GSDE 5 arc minute grid cell, the soil textural values for depths of 4.5 and 9.1 cm were averaged for the clustering of model soil layer 1. Model layer 2 spanning 10 – 35 cm is assumed to be representable by the mean of GSDE layers 16.6 and 28.9 cm and the bottom model layer spanning 35 – 410 cm by the mean of GSDE layers 49.3, 82.9, 138.3, and 229.6 cm.

GSDE does not contain information about soil depth thus the model default soil depth for each grid cell was used (Zobler, 1986). CLASS-CTEM assumes that any part of the ground column below the soil depth is bedrock and simulates water flow only in the soil part of the ground column, while the temperature dynamics are simulated over both the soil and bedrock sections.

2.3 Clustering Analysis

Clustering analysis is primarily a tool for database mining in the information sciences but it has had applications in the earth sciences, predominantly for spatial pattern analysis of remote sensing databases (e.g. Prigent et al. (2001); Archibald et al. (2013)). For the purpose of representing the spatial heterogeneity of soil textures, a clustering analysis algorithm ideally would independently identify the number of clusters without requiring per-grid cell information, beyond the high resolution soil textural information. After a literature survey, we chose the Ordering Points to Identify the Clustering Structure (OPTICS) algorithm (Ankerst et al., 1999; Daszykowski et al., 2002). OPTICS is a density-based clustering algorithm where clusters are determined to be areas of higher density than the rest of the dataset. Data points in more sparse regions are considered to be noise. Another common clustering algorithm, k-means was not used as it requires the number of clusters as an input parameter and while there are techniques to diagnostically estimate the number of clusters, they are often ambiguous and their results can differ greatly depending on technique chosen (Chiang and Mirkin, 2010).

OPTICS does not directly produce a clustering of the data but rather a hierarchical representation of the data that shows its density-based structure. A second step, using the algorithm of Sander et al. (2003), then produces the clusters. The OPTICS algorithm searches a neighbourhood of a predefined radius (ϵ) for clusters that contain a minimum number of points ($minPts$). We set ϵ to infinity and $minPts$ to 5% of the number of data points in the gridcell (sensitivity to the $minPts$ parameter is discussed further in Section 3.3.1). The parameters for the algorithm of Sander et al. (2003) were taken directly from their paper.

2.3.1 Application of OPTICS and the Sander et al. (2003) clustering algorithm

The boundaries of each CLASS-CTEM grid cell (1958 total land cells) were used to determine which high-resolution GSDE grid cells would fit within each model cell. Around 1100 GSDE cells fit within a CLASS-CTEM grid cell. From these GSDE cells, all points that were not land (lakes, rivers, etc.) were masked out. If the CLASS-CTEM grid cell did not contain more than 100 GSDE cells (which is about 340 km^2 at the equator), the CanESM soil textural information was used for that grid cell. This occurred for four CLASS-CTEM grid cells and is a result of the land mask used by CLASS-CTEM, which is the same as in the CanESM where the exact placement of the land cells is determined somewhat by the needs of the ocean model. The remaining 1954 CLASS-CTEM grid cells were then individually clustered using the OPTICS and Sander et al. (2003) algorithms. The clustering algorithms choose which GSDE grid cells are considered part of the clusters determined for each CLASS-CTEM grid cell. GSDE grid cells that, in soil texture space, are far from regions of higher density are considered noise and excluded from clusters (see Section 2.3 above), thus the percent of GSDE cells clustered varies between CLASS-CTEM grid cells. We checked the weighted mean of the clusters against the simple mean of the GSDE grid cells for each CLASS-CTEM grid cell and if the difference between them was greater than 10% for sand, clay, or OM, we assigned that cell the simple gridmean soil textures. This 10% limit was exceeded for 53 CLASS-CTEM grid cells, or $<3\%$ of the total. The vast majority of the CLASS-CTEM grid cells above this 10% limit were cells where the clustering algorithm had found only one cluster (Fig. A1). The clustering algorithms were applied to the GSDE grid cells for the first model layer (0 – 10 cm depth). For simplicity, the clustering found in the first layer was then applied to the layers below, i.e. we did not cluster the lower layers separately rather we apply the clustering assignment of each GSDE grid cell from layer one to each of the lower layers. As our study is mostly focused on the determining the impact of sub-grid soil texture on the model outputs, this simple approach is likely sufficient. Each cluster was assigned the same soil depth. Other model inputs like meteorological forcing and prescribed vegetation cover was the same for each cluster, i.e. each tile within a CLASS-CTEM grid cell had the same PFT fractional coverages on each tile (for e.g. if the CLASS-CTEM grid cell had 30% needleleaf evergreen tree, 50% C3 grass and 20% bare ground coverage, each tile would have that same PFT distribution applied to it.).

3 Results and Discussion

3.1 Model sensitivity to tiling

We performed a simple test to ascertain model sensitivity to soil texture, the number of soil tiles, and if this sensitivity has a saturating number of tiles using three example sites: N.E.USA (temperate; 43.3 °N 92.8 °W), Amazon (tropical; -1.40 °S 56.25 °W) and Sudan (arid ; 12.6 °N 28.1 °E). For this test, we first ran a simulation ~~of a~~ at each test grid cell with a soil texture of 50% sand and 50% clay. We then ran different simulations with an increasing number of tiles at each site but with the same proportion of sand and clay percentages for the grid cell weighted mean. ~~Further~~ These further simulations were i) two tiles each covering half the grid cell (one with 100% sand and the other 100% clay), ii) three tiles each covering a third (one with 100% sand, one with 50% sand and 50% clay, and the third 100% clay), iii) four tiles each covering a fourth (one with 100% sand, one with 75% sand and 25% clay, one with 25% sand and 75% clay, and the fourth 100% clay), iv) five tiles each covering a fifth (one with 100% sand, one with 75% sand and 25% clay, one with 50% sand and 50% clay, one with 25% sand and 75% clay, and the fourth 100% clay), etc. up until 20 tiles. All tiles were assigned the same vegetation, soil depth and an OM content of 0%.

Some example carbon cycle outputs are plotted in Fig. 1~~and~~ . As we increase from one tile to two, the model outputs show drops of slightly less than 10% to almost 20% ~~as we increase from one tile to two~~for the N.E. USA and 30 to 50% at the Amazon site, but show an increase in some variables of up to six-fold for the Sudan site. The change in the carbon outputs from the one-tile simulations then decreases and stabilizes, indicating that the model is not sensitive to ~~a greater numbers of tilesthan~~ an increasing number of tiles. The threshold number of tiles at which the carbon outputs stabilize is around seven or eight ~~.~~ for the N.E. USA and Amazon while the Sudan site is around 12. The Sudan site has low productivity (NPP ca. 300 $gCm^{-2}yr^{-1}$) due to arid conditions (annual precipitation ca. 400 mm yr^{-1}), and it demonstrates a strong response of the carbon cycle to soil texture. Additionally, since model runtime increases proportionally to the number of tiles (see dashed line in Sudan plot of Fig. 1), to manage computational cost only the minimum number of tiles that allows an adequate representation of sub-grid heterogeneity should be run.

These results demonstrate the model should indeed be sensitive to tiling of the soil texture and that 'too many' tiles is not necessary, ~~however this~~ and would indeed come with a computational cost penalty. However this sensitivity test is relatively unrealistic in its choice of soil texture for the clusters so the next section looks ~~at two example grid cells~~again at the Sudan test site as well as one in Brazil.

3.2 Site-level simulations

3.2.1 Evaluation of soil textural clusters

~~Looking first at~~ We first examine example grid cells from Sudan and Brazil (Figs. 2 and 3). These sites were chosen because they are from relatively arid regions and therefore soil moisture variations should play a role in the vegetation dynamics. Figures 2 and 3 show the high resolution GSDE grid cell textures for the top CLASS-CTEM soil layer. The clustering algorithm found three clusters for both example grid cells. The weight percent of clay, sand and OM for each cluster can be seen in Figs. 2 and 3 and compared to the original GSDE grid cells. The joint distribution using kernel density estimation for the sand and clay soil contents is also shown. The clustering is able to effectively capture the distinct soil textural regions apparent in both grid cells. Another example cell with a more heterogeneous GSDE soil texture is shown in Fig. A2.

3.2.2 Influence on model outputs

The CLASS-CTEM simulated net primary productivity (NPP), heterotrophic respiration (HR), and net ecosystem productivity (NEP) for the Sudan and Brazil example sites are shown in Figs. 4 and 5, respectively. Model outputs such as NPP, HR and NEP are important components of the terrestrial C cycle but they are also useful indicators of changes in soil hydrology and thermal regimes since their calculation is influenced by the soil environment as a whole. To investigate the influence of the clustering algorithm, the per tile results are shown alongside the model results taken at the grid-level (as a weighted mean) for the clustering simulation ('Cluster') and the model result if a simple mean of the GSDE soil texture for the CLASS-CTEM grid cell was used ('Gridmean').

The Sudan grid cell shows relatively large differences between the three tiles determined by the clustering algorithm. The NPP of tile C (with 36% sand, 31% clay and 2% OM) is generally very low which draws down the grid-level NPP for the Cluster simulation, however not greatly as this tile only occupies 8% of the grid cell. The other tiles (A: 91% sand, 4% clay and 1% OM covering 62% of the grid cell and B:67% sand, 15% clay and 1% OM covering 30% of the grid cell) can also differ greatly especially for HR and NPP. The NPP and NEP is generally higher for the Gridmean simulation while the HR is higher for the Cluster simulation. The different sensitivity of CLASS-CTEM's simulated NPP and HR to each tile's soil texture is at least partially due to the model formulation of these processes. In CLASS-CTEM, GPP, a component of NPP, depends upon a soil moisture stress term that uses the volumetric water content to determine the degree of soil saturation (Equations A5 - A7 in Melton and Arora (2016)) whereas the HR calculation depends on soil matric potential (Melton et al. (2015) and Equations A33 - A36 in Melton and Arora (2016)). Soil matric potential is calculated as,

$$\Psi = \Psi_{sat} \left(\frac{\theta_l}{\theta_p + \theta_i} \right)^{-b} \tag{1}$$

where b is the Clapp and Hornberger b term (Cosby et al., 1984), Ψ_{sat} is the soil moisture suction at saturation, and θ_i, θ_l, θ_p is the volumetric ice, liquid, and pore (air) content of the soil layer, respectively. The Cluster grid-level NPP is also slightly more variable than the Gridmean simulation and appears to respond strongly to precipitation changes in this arid grid cell.

The NPP, HR and NEP values at the Brazil test site (Fig. 5) are all higher than the Sudan test site due, in part, to the higher precipitation in the region. The Brazil site's Gridmean simulation generally has similar NPP and NEP to the Cluster simulation but higher HR. This appears to reflect the relatively similar behaviour between the three tiles determined for this location. The largest difference between tiles is for HR which is lower for tile C (43% sand, 36% clay and 3% OM) compared to the sandier tiles, A (91% sand, 4% clay and 1% OM) and B (67% sand, 15% clay and 1% OM).

The differences between the Cluster and Gridmean simulation for these two grid cells indicates that 1) the model is sensitive to soil textural differences, especially for more arid sites, and 2) the influence of clustering soil textures is not uniform and will depend on the conditions unique to each grid cell. We next look at the influence of the clustering on global simulations.

3.3 Global simulations

3.3.1 Evaluation of soil textural clusters

On a global scale, the clustering algorithm found, on average, slightly more than three clusters per CLASS-CTEM grid cell (3.1 ± 1.5 ; Fig. A3) with few cells having more than five clusters. The global distribution of the number of clusters and the percent of GSDE grid cells that formed the clusters is shown in Fig. 6. The number of clusters found by OPTICS shows a lower number of clusters in parts of the United States, Europe and China, with higher numbers generally found for South America and part of Africa. There appears to be some dependence between the number of clusters and the original source soil map that was incorporated into GSDE (c.f. Fig. 1 in Shangguan et al. (2014) for the distribution of source maps incorporated into GSDE). The regions of two original source maps, the General Soil Map of the United States (GSM) and the Soil Database of China for Land Surface Modelling, appear to correlate well with areas of, primarily, single tiles, as determined by the clustering algorithms. The soil textural information from these regions is of higher quality (pers. comm. W. Shangguan, 2016) with more observations contributing to a higher spatial heterogeneity in the original maps incorporated into GSDE. This higher spatial heterogeneity could have lead the clustering algorithms to find no distinct clustering by effectively increasing noise and obscuring the regions of higher density of soil textural points that indicate a cluster. The GSM map also covers Alaska but given the sparse population and remoteness of the region, the soil textural information could be of poorer quality, and hence lower spatial heterogeneity, for that state. Western Europe could also have higher quality soil data but it is only a sub section of the European Soils Database. To understand if the selection of the $minPts$ parameter caused the predominance of single tiles in these regions, we reduced $minPts$ from 5% to 1% of the number of data points in the gridcell. This did reduce the number of grid cells with only single tiles in China, the US, and Europe but it also greatly increased the number of tiles everywhere else (Fig. A4). The mean number of clusters found increased to 11.2 ± 5.2 with some grid cells having up to 20 cells. Since the model is not sensitive to more than about 7 tiles (see Section 3.1), the original $minPts$ value used appears more appropriate for the majority of the land surface.

The percent of GSDE grid cells that were included in clusters is, on average, 57.0 ± 20.1, as not all GSDE soil textural values are necessarily determined to fall within a cluster (as discussed in Section 2.3.1). The clustering does not, however, appreciably shift the simple grid-mean texture of the CLASS-CTEM grid cells (Figs A5 and A6), i.e. the raw gridmean is similar to the weighted mean of the clusters. The spatial distribution of the percent of GSDE cells clustered is shown in Fig. 6. Areas of northern Eurasia, southeastern Australia and the Prairie region of Canada appear to have lower percentages of GSDE grid cells clustered while areas like northern Africa and the high-latitudes of Canada have higher percentages clustered although the pattern on the whole is relatively heterogeneous.

3.3.2 Influence on model outputs

Global totals of CLASS-CTEM outputs for tiled (Cluster) and grid-mean (Gridmean) simulations for 1996–2015, along with observation-based estimates, are presented in Table 1. The general impact of the clustering integrated over the globe is small ~~. The physical processes such as evaporation~~ for most variables. Evapotranspiration (ET), transpiration, and runoff show small differences of around 1% or less. There are some seasonal and regional differences for ET between the Cluster and Gridmean simulations (Fig A7) but they are generally not statistically significant (independent two sample t-test p-level <0.01). The transpiration component of ET is relatively unchanged globally between the Cluster and Gridmean simulations (Table 1) while changes at the gridcell level indicate a partitioning shift between evaporation and transpiration with some arid regions showing more transpiration for the Cluster simulations (Fig. A8 and Fig. A7). Runoff also has some seasonal differences with more grid cells significantly different between Cluster and Gridmean simulations (Fig A9) and while the Cluster simulation has generally higher runoff, the signal is quite mixed. Globally, latent heat fluxes are less influenced by the tiling than sensible heat fluxes with a 0.4% difference compared to 3.7%, respectively. Seasonal maps of latent heat fluxes show little difference between the two simulations (not shown) while there is a general increase in sensible heat fluxes of the Cluster simulation over the Gridmean for all seasons in arid regions (Fig. A10).

Some variables for the carbon cycle also show similar relatively small changes with the largest changes occurring for NEP with a 4% difference between simulations and net biome productivity (NBP) with a 5% difference. The largest difference is observed for water use efficiency (WUE; defined as GPP / ET) with a percent absolute difference of about 33%. The higher mean annual global WUE of the Cluster simulation is also closer to an observation-based estimate (Xue et al., 2015) than the Gridmean simulation. The change in WUE between the two simulations will be discussed in greater detail below.

Regional differences can be much larger than the generally modest global differences found between the two simulations. The annual mean simulated soil moisture per soil layer shows some regions to differ by more than 20% between the Cluster and Gridmean simulations (Fig. 7) with more grid cells showing statistically significant differences between the simulations with increasing soil depth. The Cluster simulation has generally drier soils than the Gridmean simulation, with larger differences visible for arid regions, such as northern Australia, the Middle East, and Mongolia (which have low soil moisture so small changes in absolute amounts will appear as a larger percent change than the same absolute change in a more humid region), while the northern latitudes are wetter for some of the Canadian high north and western Siberia as well as areas of Indonesia and other parts of Southeast Asia. These patterns are not uniform and can also differ by soil model layer as is the case in the Saharan region where the second layer is generally wetter for the Cluster simulation than the Gridmean, but drier in the third layer. The drier first soil layer in the Cluster simulations leads to an increase in sensible heat fluxes over the Gridmean simulations as seen in Fig. A10 and Table 1.

The principle regions of an increase in soil moisture for the Cluster simulation over the Gridmean are in large peatland complexes such as the West Siberian Lowlands, the Hudson's Bay Lowlands, the Mackenzie River delta, and parts of Indonesia. These peatland regions are strongly influenced by the tiling due to the soil OM threshold above which the peat soils parameterization of Letts et al. (2000) is applied (soil OM >=30%; see Section 2.1). The higher porosity, greater hydraulic conductivity variation with depth, and differing thermal properties all cause greater changes in soil moisture when a grid cell or tile is treated as peat soil as opposed to a mineral soil. The differences in soil moisture appear to be relatively stable throughout the year with relatively little seasonal variation (not shown).

Changes in soil moisture will influence vegetation through changes in water supply and water stress. The mean annual gross primary productivity (GPP) as simulated by CLASS-CTEM is plotted in Fig. 8. An observation-based estimate (Beer et al., 2010) is provided for reference against the Gridmean simulation GPP. The relative percent difference between the Gridmean and Cluster simulations can be large in arid regions while relatively small elsewhere (again with the understanding that small absolute changes appear relatively larger for areas of low GPP than for the same absolute change in a region of higher productivity). The areas of significant difference are similar to the regions that saw the significant changes in total soil moisture (Fig. 7) including central Australia, Saharan Africa, and other arid regions. In these arid regions the Cluster simulation produces higher GPP values than the Gridmean simulation. However in these regions the soil moisture in the total soil column was less in the Cluster simulation than the Gridmean simulation (with the exception of Saharan Africa where the second soil layer increased in soil moisture). In non-arid regions, the general effect of the clustering was to slightly lower GPP (resulting in a slightly smaller global GPP ; Table 1). Regions like the peatland complexes that showed significant changes in soil moisture (Fig. 7) are already moist and rarely experience water stress thus these changes in soil moisture have little impact upon productivity. We investigated the temporal dynamics of GPP in the regions that differed significantly between the Cluster and Gridmean simulations using the dataset of Jung et al. (2011), and found a small improvement in root mean square deviation of the cluster simulation over the gridmean, but it was smaller than the uncertainty of the Jung et al. (2011) dataset which is relatively large in these regions due to sparse observations (e.g. Fig S2 in Beer et al. (2010)) and low productivity.

To understand how lower soil moisture could lead to higher GPP, we selected a grid cell in Australia that saw a large increase in GPP with lower soil moisture (Fig. 9). This grid cell has five tiles determined by the clustering algorithms. Of these tiles, the fifth (tile E; 78% sand, 12% clay, and 1% OM) has much higher productivity than the others, while only occupying 10% of the grid cell. The higher GPP of tile E can be understood by looking at its plant available soil water, θ_a ($\mathrm{m}^3/\mathrm{m}^3$), which we approximate using the soil's field capacity, θ_{fc}, and wilting point, θ_w,

$$\theta_a = \theta_{fc} - \theta_w \tag{2}$$

The GPP formulation of CLASS-CTEM is sensitive to θ_a (Melton and Arora, 2016) and thus enforces stomatal closing to limit water loss during periods of low moisture availability, resulting in lower productivity. From Fig. 9, we can see tile E, on an annual basis, generally has some plant available water while the other tiles, and Cluster weighted mean (Fig. 9 bottom right) and the Gridmean simulation, are commonly strongly water limited resulting in higher GPP for tile E than other tiles and the Gridmean simulation.

The large changes in the global mean WUE seen in Table 1 can also be observed regionally and seasonally (Fig. A11). Outside of arid regions the Gridmean simulation has a slightly higher WUE, although it is generally not statistically significant. Arid regions show greatly increased WUE principally due to the higher GPP discussed previously (Fig. 8), while the increase in evapotranspiration is muted by a compensating shift in transpiration vs. evaporation as discussed above –Thus, while the effect of tiling on WUE appears larger than other variables, some of the impact is simply due to how WUE is formulated (ratio of two variables and commonly given as a global mean value, not a global sum) and its sensitivity to small changes in its components.

The large influence of the clustering in arid regions demonstrates the impact of soil texture when water limitations are important. In these arid regions, the amount of water in the soil column is low and thus soil textural changes that allow greater θ_a are important, while regions with plentiful moisture are much less influenced by soil texture since water stress is less frequent and the soils generally contain sufficient water for photosynthesis. CLASS-CTEM's pedotransfer functions could also be limiting the influence of the tiling of soil textures. The range in θ_a for the Cosby et al. (1984) pedotransfer functions, as implemented in CLASS-CTEM, for soils ranging from the most disparate USDA texture classes ('sand' to 'silt' to 'clay') only covers a θ_a range of 0.08. Using another pedotransfer function may cause CLASS-CTEM to have a greater sensitivity to soil textural changes. For example, the range in θ_a using the Saxton and Rawls (2006) pedotransfer functions is over double the range of CLASS-CTEM's implementation of Cosby et al. (1984) (θ_a high to low range of 0.2). Additionally, the Cosby et al. (1984) pedotransfer functions (Fig. A12 and equations A1 - A8), while non-linear, are relatively linear in the regions of most soil textures (Fig. A5). Additionally, the GPP moisture-stress response of CLASS-CTEM could be quite different from another model thus the effects could be somewhat ~~dependant~~ dependent upon the model used.

4 Conclusions and Future work

Soil texture influences soil hydrology and temperature and is commonly assigned simple mean values across large grid cells. The sub-grid heterogeneity of soil texture can be represented by tiling of the land surface. To test the sensitivity of our model, CLASS-CTEM, to soil texture, we ran simulations of ~~an artificial grid cell~~ three artificial test grid cells with increasing numbers of tiles but the same grid mean soil texture. CLASS-CTEM's carbon cycle outputs were sensitive to the tiling with some outputs changing ~~over 15and~~ greatly, but displaying a saturating effect ~~around~~ dependent upon the climate of the grid cell that ranged between 7 ~~or 8~~ - 12 tiles. We then used two linked clustering algorithms (OPTICS (Ankerst et al., 1999; Daszykowski et al., 2002) and Sander et al. (2003)) to cluster high resolution soil textures over the relatively coarse CLASS-CTEM model grid (ca. 2.8° by 2.8°). After determining the impact of this tiling at two locations, we ran global simulations using tiled soil textures against those with a simple grid mean soil texture. The difference between the two simulations on a global scale were generally relatively small (<5%) but could be large regionally (>20%). The areas that felt the largest impact due to the soil texture tiling were in arid or peatland regions. Peatland regions were more sensitive to the tiling due to the model parameterization of peatland soils, that is subject to a minimum organic matter limit, and that could be exceeded for single tiles while the simple grid mean remained below the limit. Arid regions saw the largest impact upon GPP due to those regions' general state of moisture stress on the vegetation whereas the peatland regions generally have abundant soil moisture. Tiles that retained higher levels of plant available water in arid regions would greatly increase GPP causing the grid-level GPP to rise above that simulated when using a simple grid mean soil texture.

In water-limited regions, the inverse-texture hypothesis as put forth by ~~Noy-Meir (1973); Sala et al. (1997)~~ Noy-Meir (1973) and Sala et al. (1997) predicts that coarse-textured soils will support higher above-ground plant net primary productivity than fine-textured soils. This hypothesis has been supported by observations across precipitation gradients (Lane et al., 1998) and we also find this in our simulations for semi-arid and arid regions (Figs. 4, 5, and 9). The role of soil texture is even stronger on plant community composition based on both observations (Lane et al., 1998; Dodd et al., 2002; Dodd and Lauenroth, 1997; Fernandez-Illescas et al., 2001) and modelling studies (Bucini and Hanan, 2007). While our model does have a parameterization for competition between plants for ground coverage (Melton and Arora, 2016), we do not presently have shrub PFTs. As the major interactions in these regions is between grasses and shrubs, our competition parameterization is unlikely to appropriately capture the dynamics of plant cover due to soil texture as has been reported by observational studies.

We suggest the following as some possible future directions for this work. First, mapping of the PFTs so the PFTs observed for each point in the GSDE grid are assigned to the same tile as their underlying soil textures. Presently we give all tiles in the grid cell the same composition of PFTs whereas the underlying soil conditions could lead to notable differences in what PFTs exist in a location. Second, the Zobler (1986) soil depth dataset is markedly shallow compared to a more recent dataset (Pelletier et al., 2016) . Introducing the newer soil dataset into the clustering could allow greater distinction of clusters within a tile, for e.g. river valleys with deep soil columns with surrounding shallow soil uplands. Lastly, the use of different pedotransfer functions could yield more model sensitivity to the clustering. An examination of pedotransfer functions would need to look carefully at the impact of the function for test sites with well understood soil conditions.

While the performance of the tiled grid cells in the arid regions is encouraging, the overall impact of tiling on the terrestrial C cycle is relatively small and thus the use of a simple grid mean soil texture is likely sufficient for most applications. For ~~large scale~~ large-scale applications with a special interest in arid regions, selectively tiling those regions could be useful for capturing the impact of soil heterogeneity on plant productivity.

5 Code availability

The python code used for the OPTICS and Sander et al. (2003) algorithms as well as the CLASS-CTEM fortran code is available. Please email the first author for access to the git repository.

Appendix A: CLASS-CTEM soil pedotransfer functions

Figure A12 demonstrates the non-linear relationships between soil texture and the hydrologic soil state variables. The saturated hydraulic conductivity, K_{sat} (m/s; Fig. A12) is found from the weight percentage sand content, X_{sand} as (Cosby et al., 1984; Verseghy, 2012) :

$$K_{sat} = 7.0556 \times 10^{-6} \exp(0.0352 X_{sand} - 2.035) \tag{A1}$$

while the pore volume, θ_p ($\mathrm{m}^3/\mathrm{m}^3$; Fig. A12), is also calculated using X_{sand} (Cosby et al., 1984; Verseghy, 2012) ,

$$\theta_p = (-0.126 X_{sand} + 48.9)/100.0 \tag{A2}$$

The soil moisture suction at saturation, Ψ_{sat} (m; Fig. A12), uses X_{sand},

$$\Psi_{sat} = 0.01 \exp(-0.0302 X_{sand} + 4.33) \tag{A3}$$

The hydraulic parameter b (unitless; also called the Clapp and Hornberger B term) is calculated, via the weight percentage clay content, X_{clay} (Cosby et al., 1984; Verseghy, 2012) as,

$$b = 0.159 X_{clay} + 2.91 \tag{A4}$$

The hydraulic conductivity of the soil, K (m/s), is then related to the soil's volumetric liquid water content, θ_l ($\mathrm{m}^3/\mathrm{m}^3$) via the Clapp and Hornberger (1978) relationship:

$$K = K_{sat}(\theta_l/\theta_p)^{(2b+3)} \tag{A5}$$

In CLASS-CTEM, the field capacity of soil moisture, θ_{fc} ($\mathrm{m}^3/\mathrm{m}^3$; Fig. A12), is found by setting K in equation A5 to 0.1 mm/d (1.157 $\times$ 10^{-9} mm/s), and then solving for the liquid water content,

$$\theta_{fc} = \theta_p(1.157 \times 10^{-9}/K_{sat})^{1/(2b+3)} \tag{A6}$$

The field capacity of the lowest permeable layer, $\theta_{fc,b}$ ($\mathrm{m}^3/\mathrm{m}^3$), accounts for the permeable depth of the whole overlying soil column, z_b (m), and is found via Soulis et al. (2011)

$$\theta_{fc,b} = \theta_p/(b-1)(\Psi_{sat} b/z_b)^{1/b}[(3b+2)^{(b-1)/b} - (2b+2)^{(b-1)/b}] \tag{A7}$$

At the wilting point, the soil moisture suction, Ψ_{wilt} is set to 150 m. The volumetric water content at the wilting point, θ_w ($\mathrm{m}^3/\mathrm{m}^3$) is then calculated as,

$$\theta_w = \theta_p \left(\Psi_{wilt}/\Psi_{sat}\right)^{1/b} \tag{A8}$$

The thermal regime of the soil is also influenced by soil texture. The volumetric heat capacity of soils in CLASS-CTEM, C_g ($\mathrm{J/m^3/K}$) is derived from the volume fraction (V) and volumetric heat capacity of clay and silt, C_{fine}, sand, C_{sand}, and organic matter (OM), C_{OM}, components of the soil matrix as a weighted average:

$$C_g = \Sigma(C_{sand}V_{sand} + C_{fine}V_{fine} + C_{OM}V_{OM})/(1-\theta_p) \tag{A9}$$

In a similar manner, the soil thermal conductivity, τ_g ($\mathrm{W/m/K}$), is calculated via a weighted average of the components' thermal conductivities:

$$\tau_g = \Sigma(\tau_{sand}V_{sand} + \tau_{fine}V_{fine} + \tau_{OM}V_{OM})/(1-\theta_p) \tag{A10}$$

Organic soils, defined as those cells having an organic matter weight percent greater than 30, are assigned values of K_{sat}, θ_p, θ_{fc}, Ψ_{sat}, b, K, C_g, and τ_g based on peat texture following Letts et al. (2000). The model's first soil layer is assumed to be fibric peat, the second as hemic peat and the bottom soil layer as sapric peat.

Author contributions. J. R. M. initiated the study, performed the clustering, ran the model simulations, performed the analysis and wrote the first draft of the manuscript. K. E. M. suggested clustering analysis, provided coding help for the python scripts as well as interpretation and statistical assistance. R. S.-A. helped choose clustering algorithms, compared results to observations, and helped determine CLASSCTEM's sensitivity to the number of clusters. All authors contributed to the final manuscript.

The authors declare that they have no conflict of interest.

Acknowledgements. R.S.-A. was supported by a Natural Sciences and Engineering Research Council of Canada (NSERC) Visiting Fellowship. K.E.M. was supported by the CanSISE Network, which is funded by the Natural Science and Engineering Research Council of Canada (NSERC) under the Climate Change and Atmospheric Research (CCAR) programme. We thank Michal Daszykowski for sharing his coding of the OPTICS algorithm and Brian H. Clowers for sharing his porting of this code into Python. We also thank Amy X. Zhang for sharing her Python code implementing the algorithm of Sander et al. (2003). The high-resolution soil textural database was kindly shared by Wei Shangguan. We thank Vivek Arora and Christian Seiler for providing comments to an earlier version of this manuscript. We also thank our four reviewers (H. Zheng, E. Blyth, and two anonymous) for their comments which improved our manuscript.

References

Ankerst, M., Breunig, M. M., Kriegel, H.-P., and Sander, J.: OPTICS: Ordering Points to Identify the Clustering Structure, in: Proceedings of the 1999 ACM SIGMOD International Conference on Management of Data, SIGMOD '99, pp. 49–60, ACM, New York, NY, USA, doi:10.1145/304182.304187, 1999.

Archer, N. A. L., Quinton, J. N., and Hess, T. M.: Below-ground relationships of soil texture, roots and hydraulic conductivity in two-phase mosaic vegetation in South-east Spain, J. Arid Environ., 52, 535–553, doi:10.1006/jare.2002.1011, 2002.

Archibald, S., Lehmann, C. E. R., Gómez-Dans, J. L., and Bradstock, R. A.: Defining pyromes and global syndromes of fire regimes, Proc. Natl. Acad. Sci. U. S. A., 110, 6442–6447, doi:10.1073/pnas.1211466110, 2013.

Arora, V. K., Chiew, F. H. S., and Grayson, R. B.: Effect of sub-grid-scale variability of soil moisture and precipitation intensity on surface runoff and streamflow, J. Geophys. Res., 106, 17 073–17 091, doi:10.1029/2001JD900037, 2001.

Avissar, R. and Pielke, R. A.: A Parameterization of Heterogeneous Land Surfaces for Atmospheric Numerical Models and Its Impact on Regional Meteorology, Mon. Weather Rev., 117, 2113–2136, doi:10.1175/1520-0493(1989)117<2113:APOHLS>2.0.CO;2, 1989.

Bartholomé, E. and Belward, A. S.: GLC2000: a new approach to global land cover mapping from Earth observation data, Int. J. Remote Sens., 26, 1959–1977, doi:10.1080/01431160412331291297, 2005.

Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., Rödenbeck, C., Arain, M. A., Baldocchi, D., Bonan, G. B., Bondeau, A., Cescatti, A., Lasslop, G., Lindroth, A., Lomas, M., Luyssaert, S., Margolis, H., Oleson, K. W., Roupsard, O., Veenendaal, E., Viovy, N., Williams, C., Woodward, F. I., and Papale, D.: Terrestrial gross carbon dioxide uptake: global distribution and covariation with climate, Science, 329, 834–838, doi:10.1126/science.1184984, 2010.

Boone, A. and Wetzel, P. J.: A simple scheme for modeling sub-grid soil texture variability for use in an atmospheric climate model, J. Meteorol. Soc. Japan, 77, 317–333, 1999.

Bucini, G. and Hanan, N. P.: A continental-scale analysis of tree cover in African savannas, Glob. Ecol. Biogeogr., 16, 593–605, 2007.

Chiang, M. M.-T. and Mirkin, B.: Intelligent Choice of the Number of Clusters in K-Means Clustering: An Experimental Study with Different Cluster Spreads, J. Classification, 27, 3–40, doi:10.1007/s00357-010-9049-5, 2010.

Clapp, R. B. and Hornberger, G. M.: Empirical equations for some soil hydraulic properties, Water Resour. Res., 14, 601–604, doi:10.1029/WR014i004p00601, 1978.

Cosby, B. J., Hornberger, G. M., Clapp, R. B., and Ginn, T. R.: A Statistical Exploration of the Relationships of Soil Moisture Characteristics to the Physical Properties of Soils, Water Resour. Res., 20, 682–690, doi:10.1029/WR020i006p00682, 1984.

Daszykowski, M., Walczak, B., and Massart, D. L.: Looking for natural patterns in analytical data. 2. Tracing local density with OPTICS, J. Chem. Inf. Comput. Sci., 42, 500–507, doi:10.1021/ci010384s, 2002.

Decharme, B. and Douville, H.: Introduction of a sub-grid hydrology in the ISBA land surface model, Clim. Dyn., 26, 65–78, doi:10.1007/s00382-005-0059-7, 2005.

Dirmeyer, P. A., Dolman, A. J., and Sato, N.: The Pilot Phase of the Global Soil Wetness Project, Bull. Am. Meteorol. Soc., 80, 851–878, doi:10.1175/1520-0477(1999)080<0851:TPPOTG>2.0.CO;2, 1999.

Dodd, M. B. and Lauenroth, W. K.: The influence of soil texture on the soil water dynamics and vegetation structure of a shortgrass steppe ecosystem, Plant Ecol., 133, 13–28, doi:10.1023/A:1009759421640, 1997.

Dodd, M. B., Lauenroth, W. K., Burke, I. C., and Chapman, P. L.: Associations between vegetation patterns and soil texture in the shortgrass steppe, Plant Ecol., 158, 127–137, doi:10.1023/A:1015525303754, 2002.

English, N. B., Weltzin, J. F., Fravolini, A., Thomas, L., and Williams, D. G.: The influence of soil texture and vegetation on soil moisture under rainout shelters in a semi-desert grassland, J. Arid Environ., 63, 324–343, doi:10.1016/j.jaridenv.2005.03.013, 2005.

Essery, R. L. H., Best, M. J., Betts, R. A., Cox, P. M., and Taylor, C. M.: Explicit Representation of Subgrid Heterogeneity in a GCM Land Surface Scheme, J. Hydrometeorol., 4, 530–543, doi:10.1175/1525-7541(2003)004<0530:EROSHI>2.0.CO;2, 2003.

Famiglietti, J. S. and Wood, E. F.: Multiscale modeling of spatially variable water and energy balance processes, Water Resour. Res., 30, 3061–3078, doi:10.1029/94WR01498, 1994.

Fekete, B. M., Vörösmarty, C. J., and Grabs, W.: High-resolution fields of global runoff combining observed river discharge and simulated water balances, Global Biogeochem. Cycles, 16, 2002.

Fernandez-Illescas, C. P., Porporato, A., Laio, F., and Rodriguez-Iturbe, I.: The ecohydrological role of soil texture in a water-limited ecosystem, Water Resour. Res., 37, 2863–2872, 2001.

Hook, P. B. and Burke, I. C.: Biogeochemistry in a Shortgrass Landscape: Control by Topography, Soil Texture, and Microclimate, Ecology, 81, 2686–2703, doi:10.2307/177334, 2000.

Hurtt, G. C., Chini, L. P., Frolking, S., Betts, R. A., Feddema, J., Fischer, G., Fisk, J. P., Hibbard, K., Houghton, R. A., Janetos, A., Jones, C. D., Kindermann, G., Kinoshita, T., Goldewijk, K. K., Riahi, K., Shevliakova, E., Smith, S., Stehfest, E., Thomson, A., Thornton, P., van Vuuren, D. P., and Wang, Y. P.: Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands, Clim. Change, 109, 117–161, doi:10.1007/s10584-011-0153-2, 2011.

Jasechko, S., Sharp, Z. D., Gibson, J. J., Birks, S. J., Yi, Y., and Fawcett, P. J.: Terrestrial water fluxes dominated by transpiration, Nature, 496, 347–350, doi:10.1038/nature11983, 2013.

Jiménez, C., Prigent, C., Mueller, B., Seneviratne, S. I., McCabe, M. F., Wood, E. F., Rossow, W. B., Balsamo, G., Betts, A. K., Dirmeyer, P. A., Fisher, J. B., Jung, M., Kanamitsu, M., Reichle, R. H., Reichstein, M., Rodell, M., Sheffield, J., Tu, K., and Wang, K.: Global intercomparison of 12 land surface heat flux estimates, J. Geophys. Res., 116, D02 102, doi:10.1029/2010JD014545, 2011.

Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., Arneth, A., Bernhofer, C., Bonal, D., Chen, J., and Others: Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations, Journal of Geophysical Research: Biogeosciences (2005–2012), 116, 2011.

Kauppi, P.: New, low estimate for carbon stock in global forest vegetation based on inventory data, Silva Fenn., 37, doi:10.14214/sf.484, 2003.

Ke, Y., Leung, L. R., Huang, M., and Li, H.: Enhancing the representation of subgrid land surface characteristics in land surface models, Geoscientific Model Development, 6, 1609–1622, doi:10.5194/gmd-6-1609-2013, 2013.

Koster, R. D. and Suarez, M. J.: A Comparative Analysis of Two Land Surface Heterogeneity Representations, J. Clim., 5, 1379–1390, doi:10.1175/1520-0442(1992)005<1379:ACAOTL>2.0.CO;2, 1992.

Kumar, J., Mills, R. T., Hoffman, F. M., and Hargrove, W. W.: Parallel k-Means Clustering for Quantitative Ecoregion Delineation Using Large Data Sets, Procedia Comput. Sci., 4, 1602–1611, doi:10.1016/j.procs.2011.04.173, 2011.

Landry, J.-S., Ramankutty, N., and Parrott, L.: Investigating the Effects of Subgrid Cell Dynamic Heterogeneity on the Large-Scale Modeling of Albedo in Boreal Forests, Earth Interact., 20, 1–23, doi:10.1175/EI-D-15-0022.1, 2016.

Lane, D. R., Coffin, D. P., and Lauenroth, W. K.: Effects of soil texture and precipitation on above-ground net primary productivity and vegetation structure across the Central Grassland region of the United States, J. Veg. Sci., 9, 239–250, 1998.

Le Quéré, C., Andrew, R. M., Canadell, J. G., Sitch, S., Korsbakken, J. I., Peters, G. P., Manning, A. C., Boden, T. A., Tans, P. P., Houghton, R. A., Keeling, R. F., Alin, S., Andrews, O. D., Anthoni, P., Barbero, L., Bopp, L., Chevallier, F., Chini, L. P., Ciais, P., Currie, K., Delire, C., Doney, S. C., Friedlingstein, P., Gkritzalis, T., Harris, I., Hauck, J., Haverd, V., Hoppema, M., Klein Goldewijk, K., Jain, A. K., Kato, E., Körtzinger, A., Landschützer, P., Lefèvre, N., Lenton, A., Lienert, S., Lombardozzi, D., Melton, J. R., Metzl, N., Millero, F., Monteiro, P. M. S., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., O'Brien, K., Olsen, A., Omar, A. M., Ono, T., Pierrot, D., Poulter, B., Rödenbeck, C., Salisbury, J., Schuster, U., Schwinger, J., Séférian, R., Skjelvan, I., Stocker, B. D., Sutton, A. J., Takahashi, T., Tian, H., Tilbrook, B., van der Laan-Luijkx, I. T., van der Werf, G. R., Viovy, N., Walker, A. P., Wiltshire, A. J., and Zaehle, S.: Global Carbon Budget 2016, doi:10.5194/essd-8-605-2016, 2016.

Letts, M. G., Roulet, N. T., Comer, N. T., Skarupa, M. R., and Verseghy, D. L.: Parametrization of peatland hydraulic properties for the Canadian land surface scheme, Atmosphere-Ocean, 38, 141–160, doi:10.1080/07055900.2000.9649643, 2000.

Li, R. and Arora, V. K.: Effect of mosaic representation of vegetation in land surface schemes on simulated energy and carbon balances, Biogeosciences, 9, 593–605, doi:10.5194/bg-9-593-2012, 2012.

Li, Y., Shi, Z., Li, F., and Li, H.-Y.: Delineation of site-specific management zones using fuzzy clustering analysis in a coastal saline land, Comput. Electron. Agric., 56, 174–186, doi:10.1016/j.compag.2007.01.013, 2007.

Meinshausen, M., Smith, S. J., Calvin, K., Daniel, J. S., Kainuma, M. L. T., Lamarque, J.-F., Matsumoto, K., Montzka, S. A., Raper, S. C. B., Riahi, K., Thomson, A., Velders, G. J. M., and van Vuuren, D. P.: The RCP greenhouse gas concentrations and their extensions from 1765 to 2300, Clim. Change, 109, 213–241, doi:10.1007/s10584-011-0156-z, 2011.

Melton, J. R. and Arora, V. K.: Sub-grid scale representation of vegetation in global land surface schemes: implications for estimation of the terrestrial carbon sink, Biogeosciences, 11, 1021–1036, doi:10.5194/bgd-10-16003-2013, 2014.

Melton, J. R. and Arora, V. K.: Competition between plant functional types in the Canadian Terrestrial Ecosystem Model (CTEM) v. 2.0, Geoscientific Model Development, 9, 323–361, doi:10.5194/gmd-9-323-2016, 2016.

Melton, J. R., Shrestha, R. K., and Arora, V. K.: The influence of soils on heterotrophic respiration exerts a strong control on net ecosystem productivity in seasonally dry Amazonian forests, Biogeosciences, 12, 1151–1168, doi:10.5194/bg-12-1151-2015, 2015.

Molod, A. and Salmun, H.: A global assessment of the mosaic approach to modeling land surface heterogeneity, J. Geophys. Res. D: Atmos., 107, ACL 9–1–ACL 9–18, doi:10.1029/2001JD000588, 2002.

Newman, A. J., Clark, M. P., Winstral, A., Marks, D., and Seyfried, M.: The Use of Similarity Concepts to Represent Subgrid Variability in Land Surface Models: Case Study in a Snowmelt-Dominated Watershed, J. Hydrometeorol., 15, 1717–1738, doi:10.1175/JHM-D-13-038.1, 2014.

Nitta, T., Yoshimura, K., Takata, K., O'ishi, R., Sueyoshi, T., Kanae, S., Oki, T., Abe-Ouchi, A., and Liston, G. E.: Representing Variability in Subgrid Snow Cover and Snow Depth in a Global Land Model: Offline Validation, J. Clim., 27, 3318–3330, doi:10.1175/JCLI-D-13-00310.1, 2014.

Noy-Meir, I.: Desert Ecosystems: Environment and Producers, Annu. Rev. Ecol. Syst., 4, 25–51, 1973.

Pelletier, J., Broxton, P., Hazenberg, P., Zeng, X., Troch, P., Niu, G., Williams, Z., Brunke, M., and Gochis, D.: Global 1-km Gridded Thickness of Soil, Regolith, and Sedimentary Deposit Layers, doi:10.3334/ORNLDAAC/1304, 2016.

Peng, Y., Arora, V. K., Kurz, W. A., Hember, R. A., Hawkins, B. J., Fyfe, J. C., and Werner, A. T.: Climate and atmospheric drivers of historical terrestrial carbon uptake in the province of British Columbia, Canada, Biogeosciences, 11, 635–649, doi:10.5194/bg-11-635-2014, 2014.

Pielke, R. A., Dalu, G. A., Snook, J. S., Lee, T. J., and Kittel, T. G. F.: Nonlinear Influence of Mesoscale Land Use on Weather and Climate, J. Clim., 4, 1053–1069, doi:10.1175/1520-0442(1991)004<1053:NIOMLU>2.0.CO;2, 1991.

Prigent, C., Matthews, E., Aires, F., and Rossow, W. B.: Remote sensing of global wetland dynamics with multiple satellite data sets, Geophys. Res. Lett., 28, 4631–4634, doi:10.1029/2001GL013263, 2001.

Randerson, J. T., Chen, Y., Werf, G. R., Rogers, B. M., and Morton, D. C.: Global burned area and biomass burning emissions from small fires, Journal of Geophysical Research: Biogeosciences (2005–2012), 117, doi:10.1029/2012JG002128, 2012.

Sala, O. E., Lauenroth, W. K., and Golluscio, R. A.: Plant functional types in temperate semi-arid regions, Plant functional types: their relevance to ecosystem properties and global change, pp. 217–233, 1997.

Sander, J., Qin, X., Lu, Z., Niu, N., and Kovarsky, A.: Automatic Extraction of Clusters from Hierarchical Clustering Representations, in: Advances in Knowledge Discovery and Data Mining, Lecture Notes in Computer Science, pp. 75–87, Springer Berlin Heidelberg, doi:10.1007/3-540-36175-8_8, 2003.

Saxton, K. E. and Rawls, W. J.: Soil Water Characteristic Estimates by Texture and Organic Matter for Hydrologic Solutions, Soil Sci. Soc. Am. J., 70, 1569–1578, doi:10.2136/sssaj2005.0117, 2006.

Schlesinger, W. H. and Jasechko, S.: Transpiration in the global water cycle, Agric. For. Meteorol., 189–190, 115–117, 2014.

Shangguan, W., Dai, Y., Duan, Q., Liu, B., and Yuan, H.: A global soil data set for earth system modeling, Journal of Advances in Modeling Earth Systems, 6, 249–263, doi:10.1002/2013MS000293, 2014.

Shao, Y. and Henderson-Sellers, A.: Validation of soil moisture simulation in landsurface parameterisation schemes with HAPEX data, Glob. Planet. Change, 13, 11–46, doi:10.1016/0921-8181(95)00038-0, 1996.

Soulis, E. D., Craig, J. R., Fortin, V., and Liu, G.: A simple expression for the bulk field capacity of a sloping soil horizon, Hydrol. Process., 25, 112–116, doi:10.1002/hyp.7827, 2011.

Still, C. J., Berry, J. A., Collatz, G. J., and DeFries, R. S.: Global distribution of C3 and C4 vegetation: Carbon cycle implications, Global Biogeochem. Cycles, 17, 6–1–6–14, doi:10.1029/2001GB001807, 2003.

Trenberth, K. E., Fasullo, J. T., and Kiehl, J.: Earth's Global Energy Budget, Bull. Am. Meteorol. Soc., 90, 311–323, doi:10.1175/2008BAMS2634.1, 2009.

Trenberth, K. E., Fasullo, J. T., and Mackaro, J.: Atmospheric Moisture Transports from Ocean to Land and Global Energy Flows in Reanalyses, J. Clim., 24, 4907–4924, doi:10.1175/2011JCLI4171.1, 2011.

Verseghy, D.: CLASS – The Canadian land surface scheme, Climate Research Division, Science and Technology Branch, Environment Canada, 2012.

Viovy, N.: CRU-NCEP Version 7, http://dods.extra.cea.fr/store/p529viov/cruncep/V6_1901_2014/, accessed Aug 1, 2016.

Wood, E., Lettenmaier, D., and Zartarian, V. G.: A land-surface hydrology parameterization with sub-grid variability for general circulation models, J. Geophys. Res., 97, 2717–2728, 1992.

Xue, B.-L., Guo, Q., Otto, A., Xiao, J., Tao, S., and Li, L.: Global patterns, trends, and drivers of water use efficiency from 2000 to 2013, Ecosphere, 6, art174, doi:10.1890/ES14-00416.1, 2015.

Zobler, L.: A world soil file for global climate modelling, title of the publication associated with this dataset: NASA Technical Memorandum 87802, 1986.

Table 1. Global annual values for CLASS-CTEM model outputs based on simulations using grid-mean soil textures (Gridmean) and tiled simulations derived from the clustering analysis (Cluster). Values are an average over the period 1996 – 2015.

CLASS-CTEM Output	Cluster	Gridmean	Percent absolute ~~difference~~ difference	Observation-base
~~Evaporation (~~Evapotranspiration (ET; 10^3 km^3 yr^{-1})	78.3	78.6	0.5	83.9±9.9 (Trenbertl
Transpiration (T; 10^3 km^3 yr^{-1})	21.1	21.2	0.3	62±8 (Jasechko et al., 2013) , 45±4.5 (
T/ET (%)	27.0	27.0	0.1	61±15 (Schlesinger and
Runoff (10^3 km^3 yr^{-1})	32.8	32.4	1.1	38.3 (Fekete et a
Latent heat fluxes (W m^{-2})	44.9	45.2	0.4	39 ± 2 (Jung et al., 2011) , 38.5 37 - 59 (Jiménez e
Sensible heat fluxes (W m^{-2})	25.5	24.6	3.7	41 ± 4 (Jung et al., 2011) , 27 18 - 57 (Jiménez e
Water Use Efficiency (g C kg^{-1} water)	1.47	1.10	32.8	1.70 (Xue et al
Gross primary productivity (GPP) (Pg C yr^{-1})	133.1	133.6	0.4	123 ± 8 (Beer et
Vegetation biomass (Pg C)	555.00	558.46	0.6	300 - 536 (Forest
Soil carbon mass (Pg C)	1132.1	1119.6	1.1	1922^c (Shangguan
Area burnt (10^4 km^2 yr^{-1}	484	505	4.2	464 (Randerson e
Net ecosystem productivity (NEP) (Pg C yr^{-1})	4.6	4.8	4.0	
Net biome productivity (NBP) (Pg C yr^{-1})	1.0	1.1	5.0	1.0 – 2.5^d (Le Quér

Percent absolute difference is calculated as abs{100 - [(clustered value / grid-mean value) * 100]}. [a] Value from eight reanalyses for 2002 – 2008, except ERA-40 which was for the 1990s. [b] As sum version of CLASS-CTEM does not simulate permafrost C pools. [d] Range of all estimates across 1990-2015 time period.

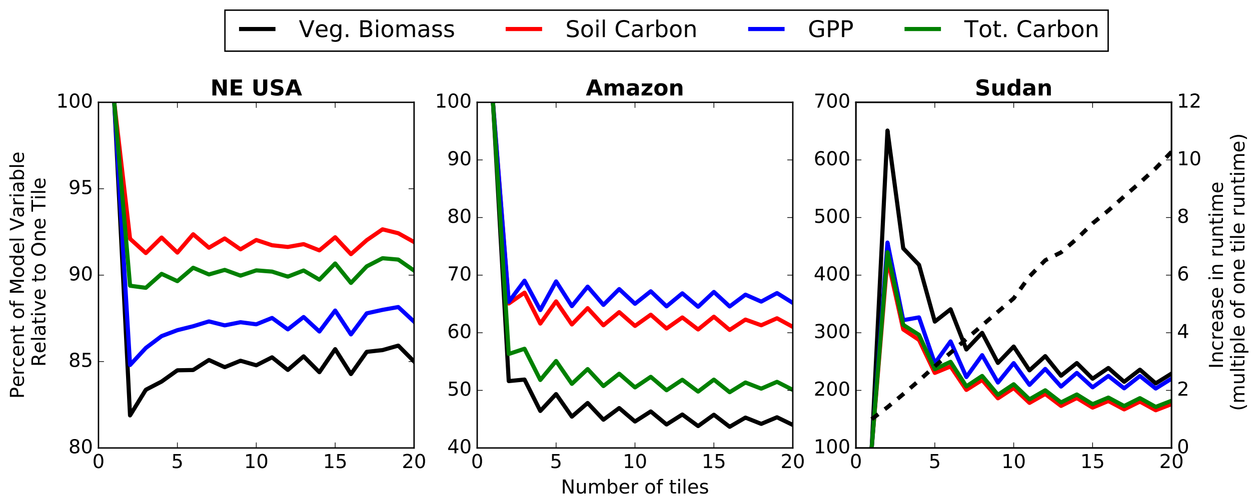

Figure 1. Sensitivity test of CLASS-CTEM to the number of tiles (clusters) for ~~a single~~ three test grid ~~cell~~cells. The texture of each tile as the number of tiles increases is described in Section 3.1. GPP is gross primary productivity. All simulations were run until a new equilibrium state was established. The increase in runtime of the model is displayed as a dashed line.

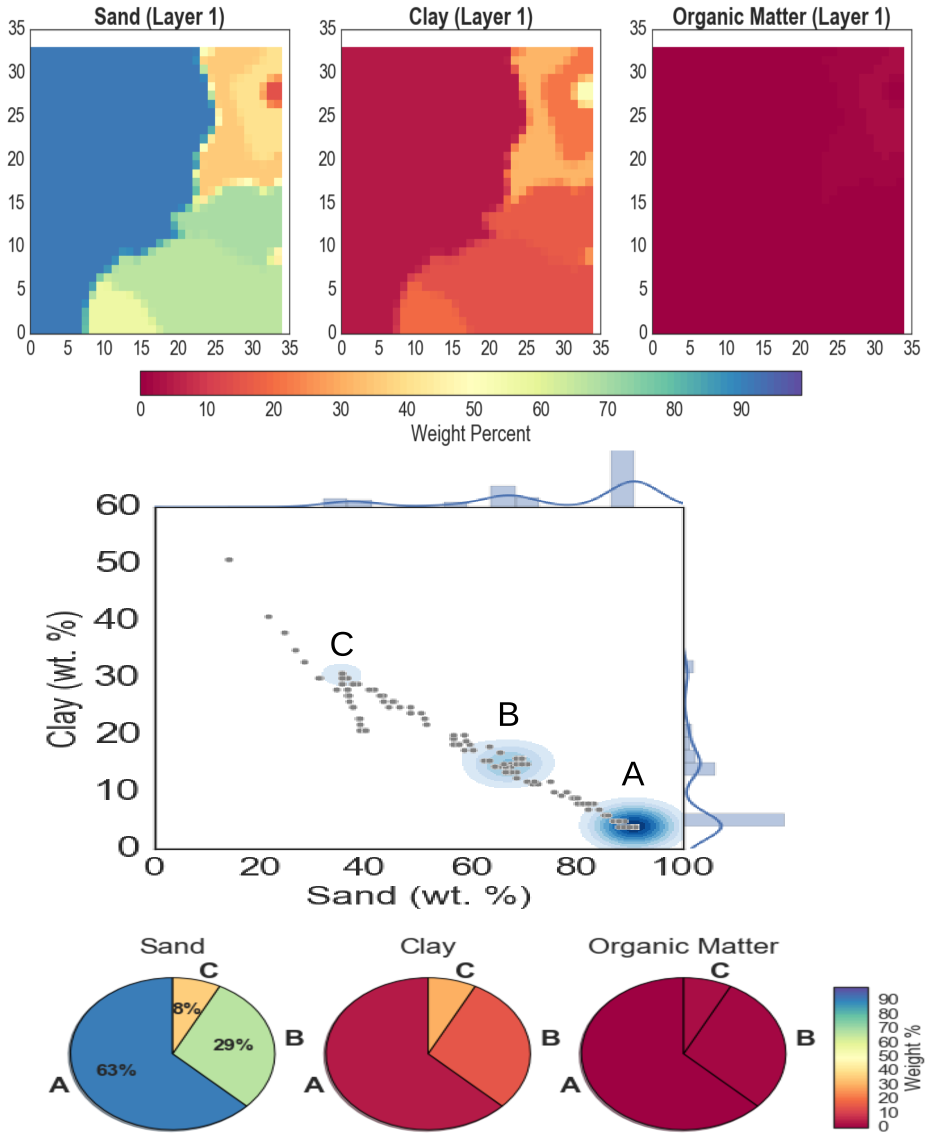

Figure 2. Example CLASS-CTEM grid cell located in Sudan (12.6 °N 28.1°E). The top panel shows the GSDE sand, clay, and organic carbon weight percents for GSDE cells within the CLASS-CTEM grid cell. Each GSDE grid cell is 5 arc min by 5 arc min. The numbers on the plot axes are the number of GSDE grid cells along that axis. The joint distribution using kernel density estimation for the soil sand and clay content is shown in the centre panel. The histograms on the axes and the blue colour scaling demonstrate qualitatively the number of GSDE grid cells sharing the similar soil textural space. The clustering algorithm found three clusters (labelled A, B, and C) with a fractional area per cluster and soil texture as shown in the pie charts. The pie charts can be visually referenced to the top panel which uses the same colour scheme, e.g. Cluster A covers 63% of the CLASS-CTEM grid cell with 91% sand, 4% clay and 1% OM. In the scatter plot the label is placed close to the cluster value to help illustrate the cluster relation in sand-clay space.

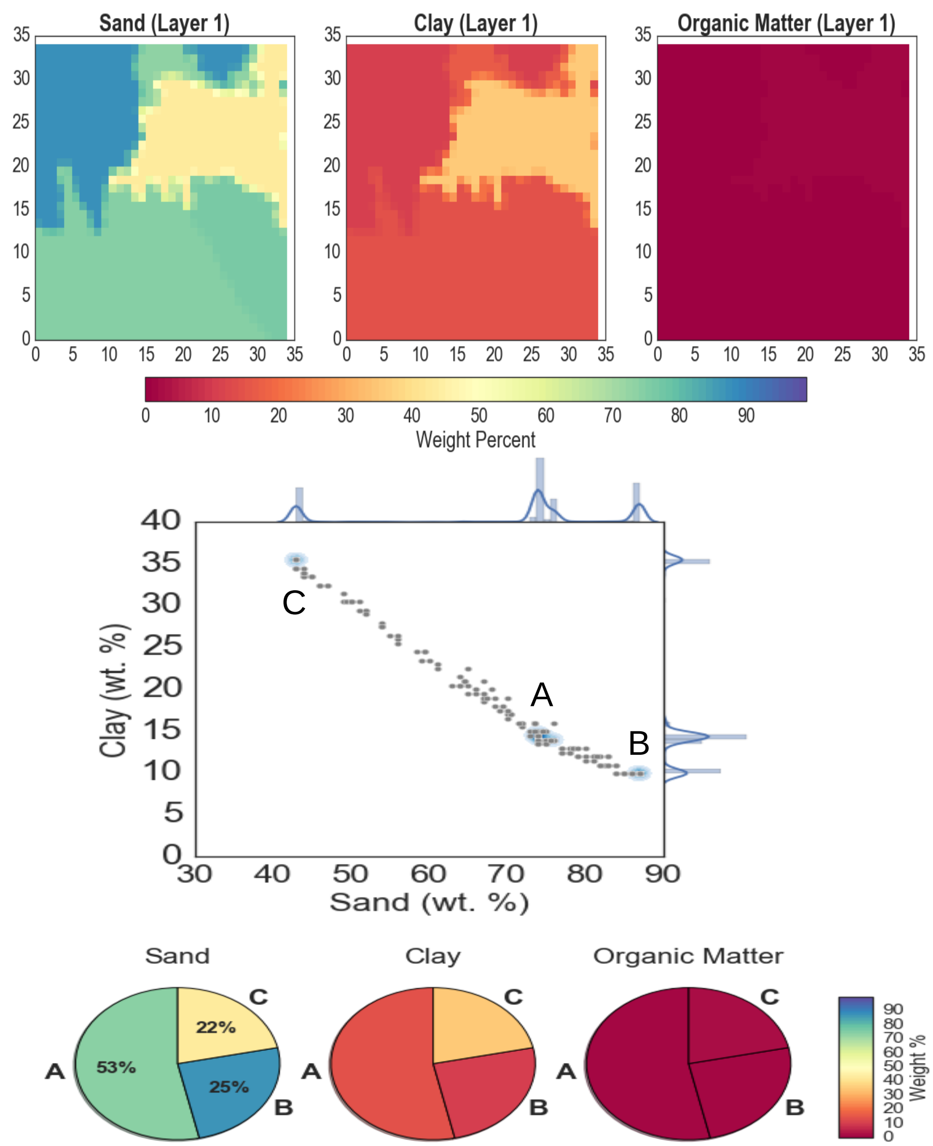

Figure 3. Similar to Fig. 2 for a CLASS-CTEM grid cell located in Brazil (9.8 °S 45.0°W).

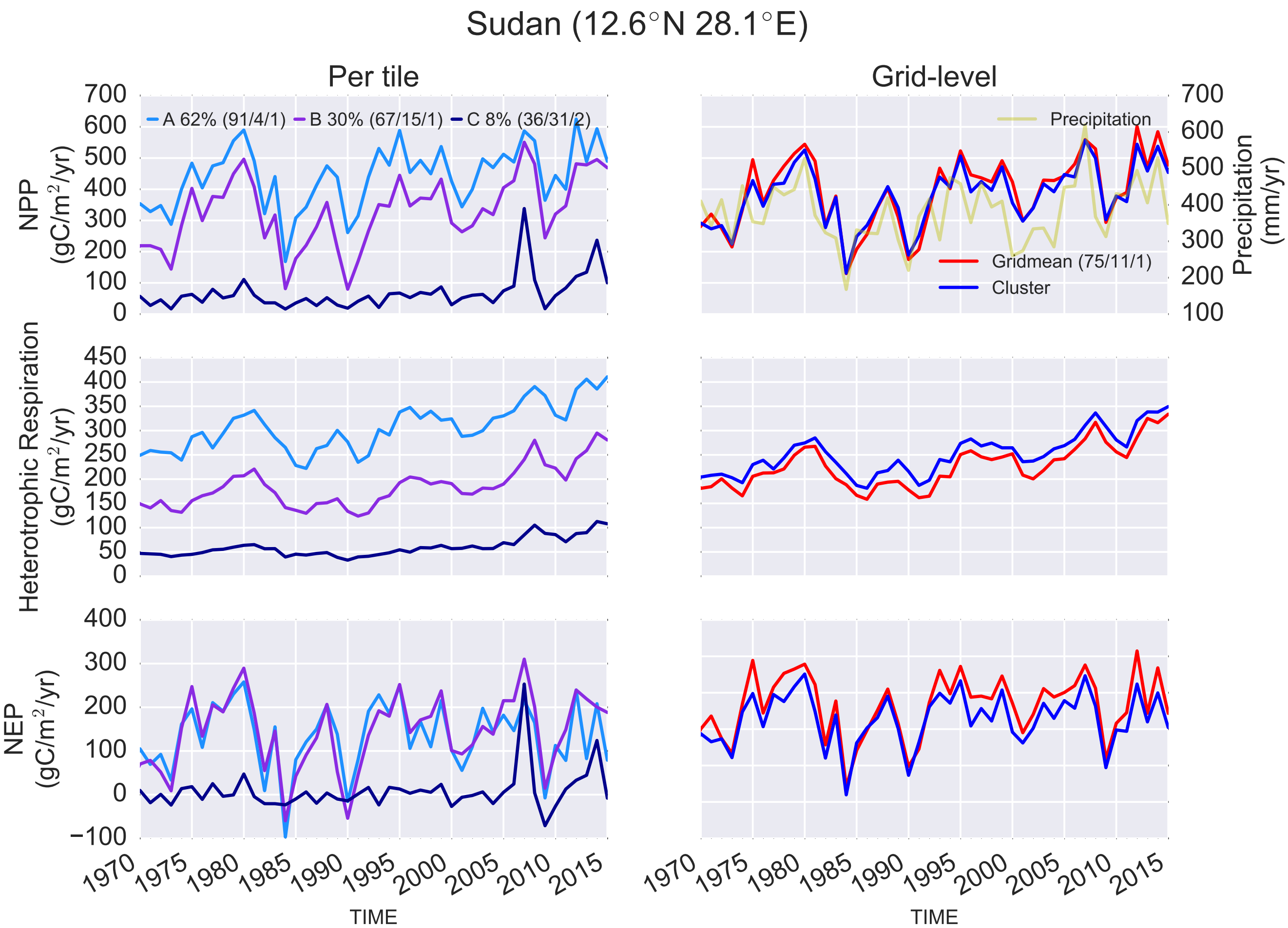

Figure 4. CLASS-CTEM simulated net primary productivity (NPP), heterotrophic respiration and net ecosystem productivity (NEP; NPP – heterotrophic respiration) for the same grid cell in Sudan as Fig. 2. The left column shows the model results per tile with soil textures listed as percent sand/clay/OM along with the tile percent grid cell coverage. The right panel is the model results at the grid-level for the Cluster simulation (weighted mean average of all tiles) and the Gridmean simulation (simple mean of GSDE soil textures). The annual precipitation for this grid cell from CRU-NCEP is included for reference in the upper right plot.

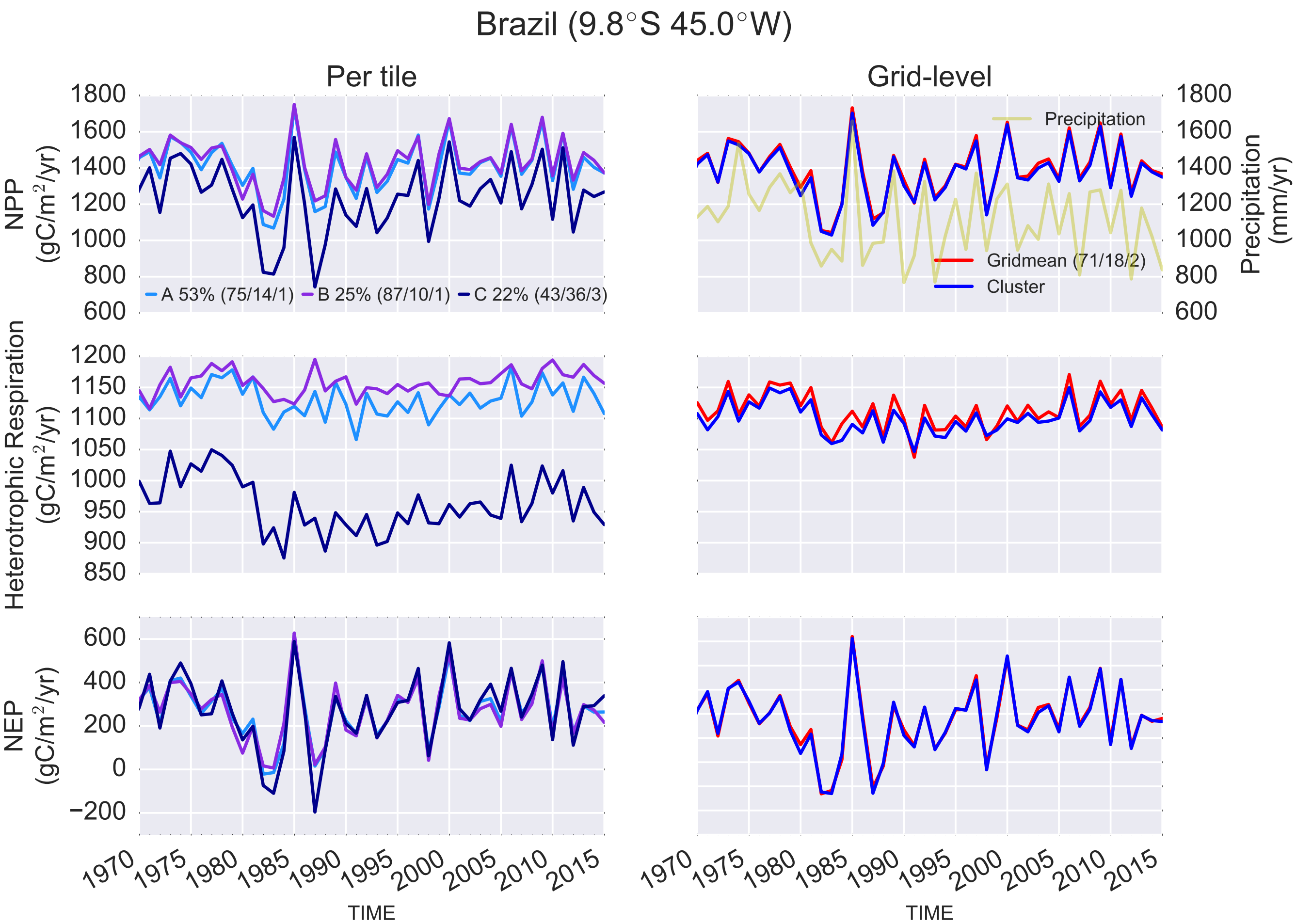

Figure 5. Same as Fig. 4 for a grid cell in Brazil (same cell as Fig. 3).

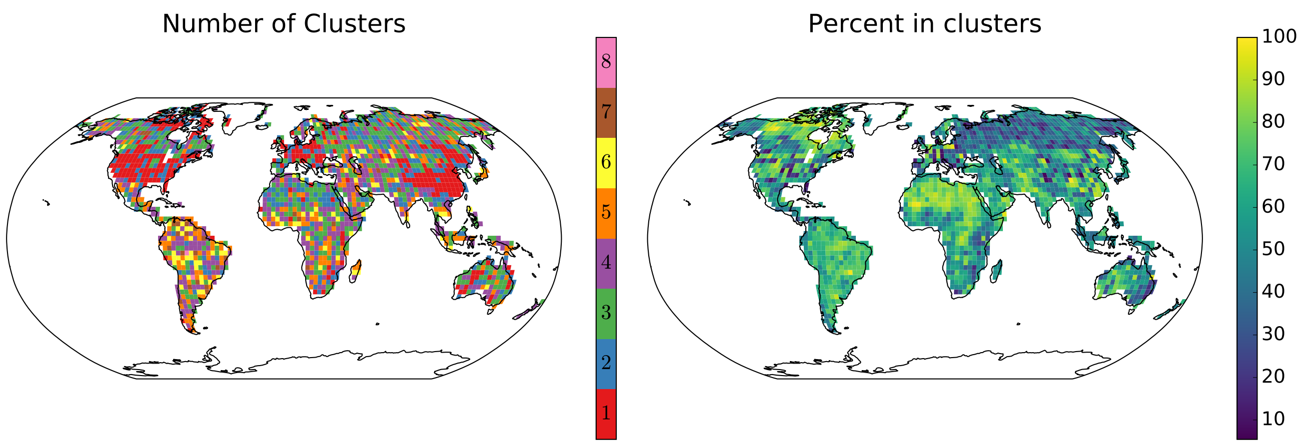

Figure 6. Global distributions of the number of clusters (tiles) found per CLASS-CTEM grid cell (left) and the percent of GSDE grid cells clustered per CLASS-CTEM grid cell (right).

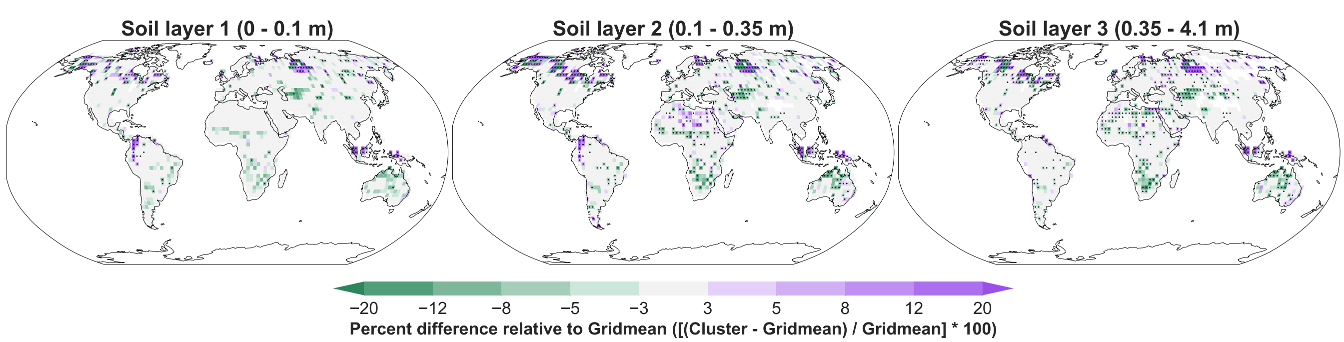

Figure 7. Percent difference in soil moisture per CLASS-CTEM soil layer between the Cluster and Gridmean simulations (mean of 1995–2015). Grid cells with soil moisture below 10^{-5} kg m^{-2} were masked out to prevent instances of divide by zero and overly large relative differences in regions of very little soil moisture. Positive values indicate the Cluster soil moisture is larger while negative values indicate the Gridmean soil moisture is larger. Dots indicate grid cells that are statistically significant (independent two sample t-test p-level <0.01)

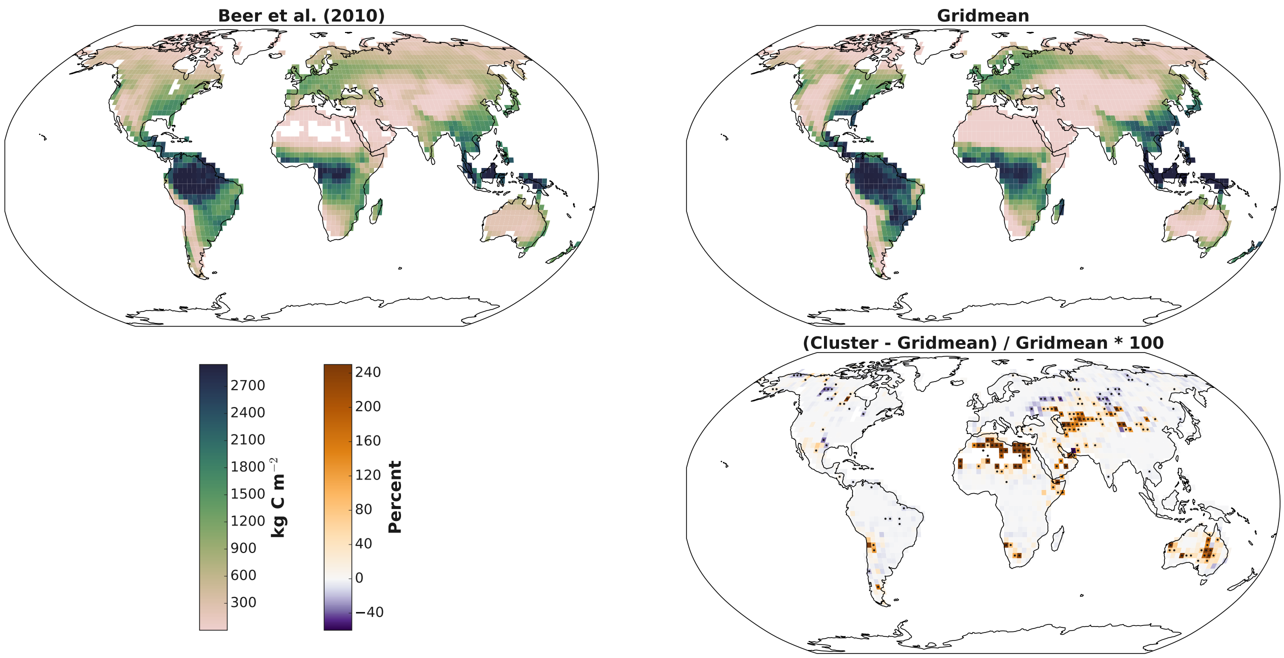

Figure 8. Mean annual gross primary productivity (1982 – 2008) from an observation-based dataset (Beer et al., 2010) (top left), the Gridmean simulation (top right), and percent relative difference between the Cluster and Gridmean simulations (bottom right). Note that many of the regions with the largest changes in GPP between the two approaches are also regions with low GPP hence the absolute change in GPP is generally small. Dots indicate grid cells that are statistically significant (independent two sample t-test p-level <0.01). For the areas of significant change in GPP between the Cluster and Gridmean simulations, comparison of Cluster and Gridmean simulations against observations was not significant after accounting for the observational uncertainty.

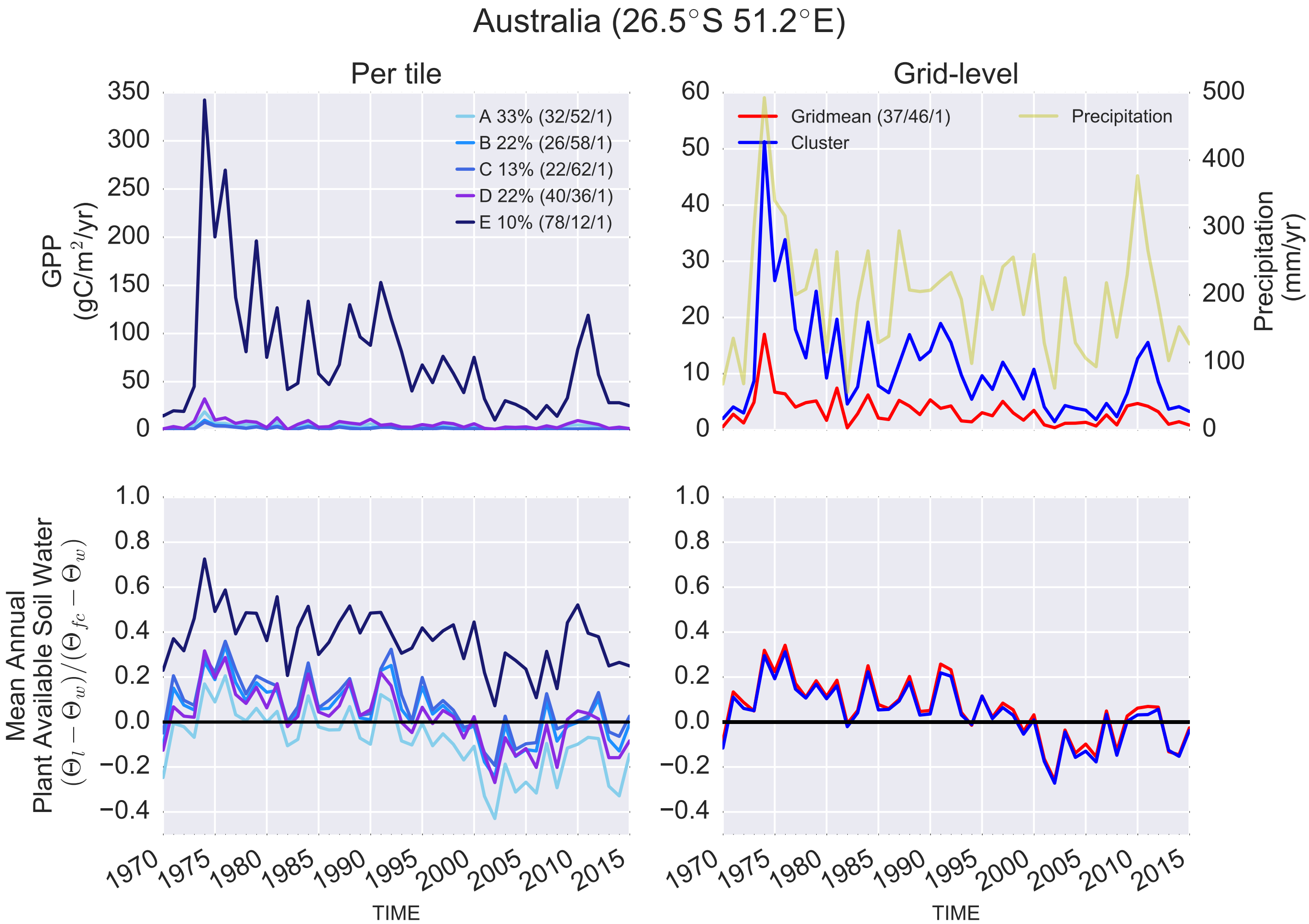

Figure 9. Australian grid cell that has higher GPP for the Cluster simulation than the Gridmean simulation, but lower soil moisture. A measure of the mean annual plant available soil water, scaled so that 1 is field capacity and 0 is wilting point, is calculated for the second model soil layer (0.1 - 0.35 m) and is described in Section 3.3.2. The annual precipitation for this grid cell from CRU-NCEP is included for reference in the upper right plot.

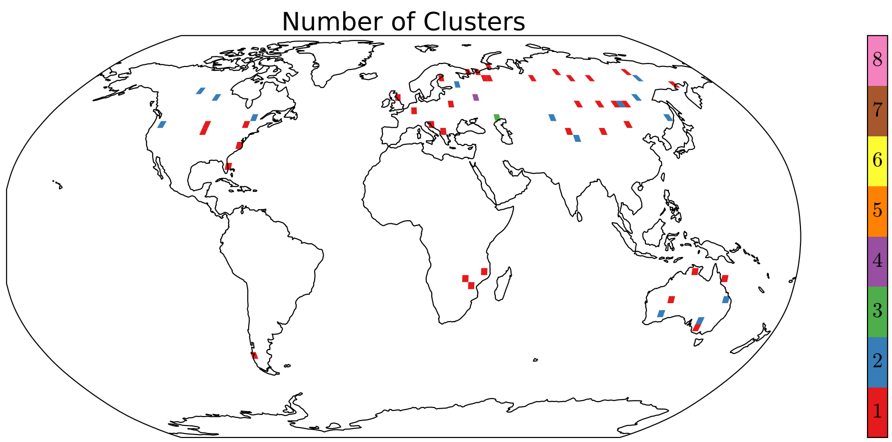

Figure A1. Map of the number of clusters for all CLASS-CTEM grid cells where the weighted mean of the clusters was more than 10% different than the simple mean of all GSDE grid cells within a CLASS-CTEM grid cell. These grid cells were then assigned the simple gridmean soil texture values for all simulations (see Section 2.3.1).

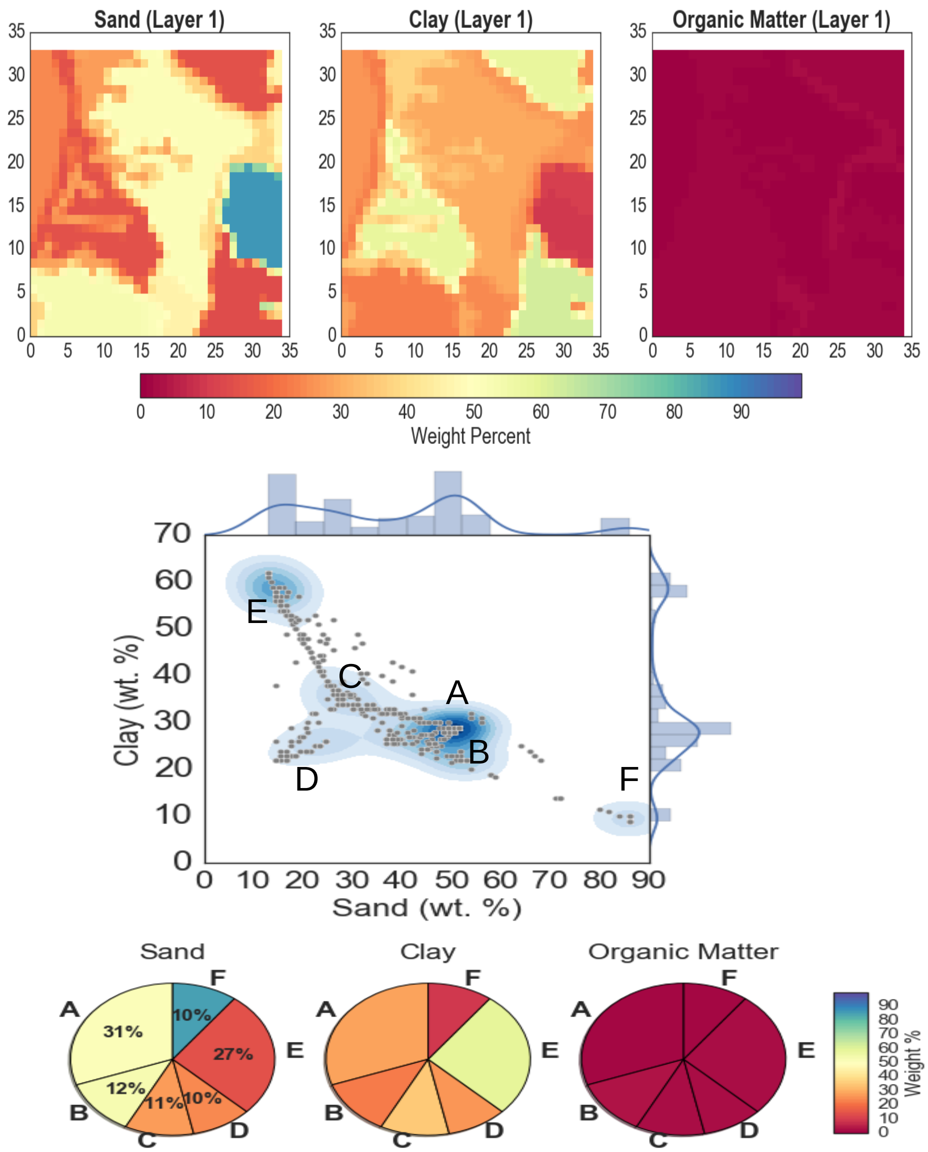

Figure A2. Similar to Fig. 2 for a grid cell in Paraguay. This figure demonstrates the clustering performance for a grid cell with a more heterogeneous soil texture. The clustering algorithm found six clusters for this grid cell. The clusters are labelled A through F in the bottom pie chart and middle scatter plot. In the scatter plot the label is placed close to the cluster value to help illustrate the cluster relation in sand-clay space.

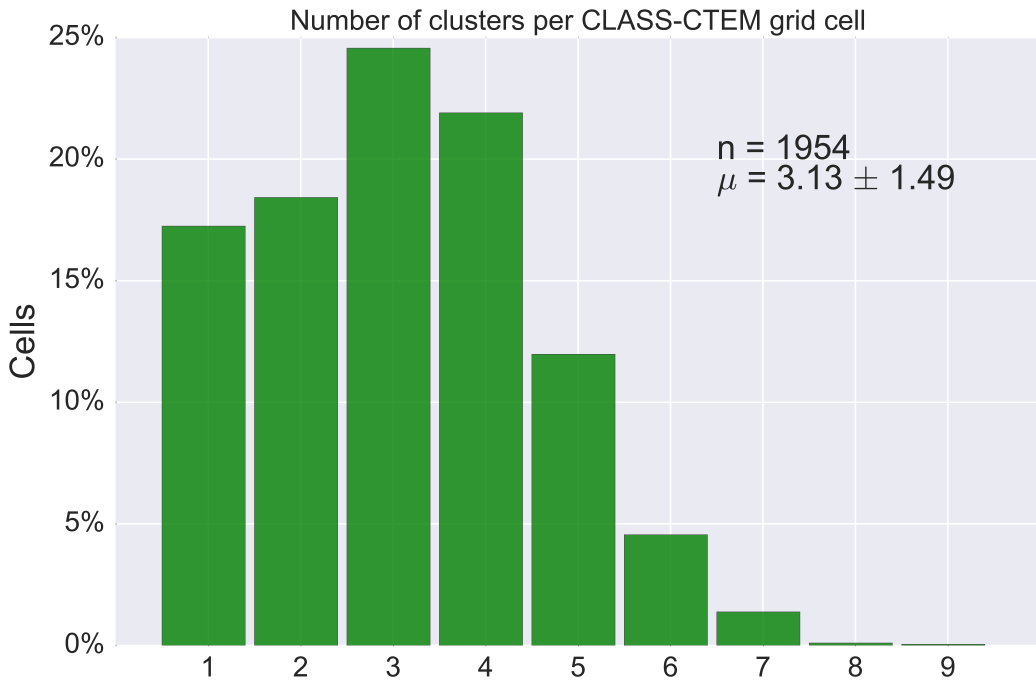

Figure A3. Number of clusters determined per CLASS-CTEM grid cell for the GSDE (Shangguan et al., 2014) dataset.

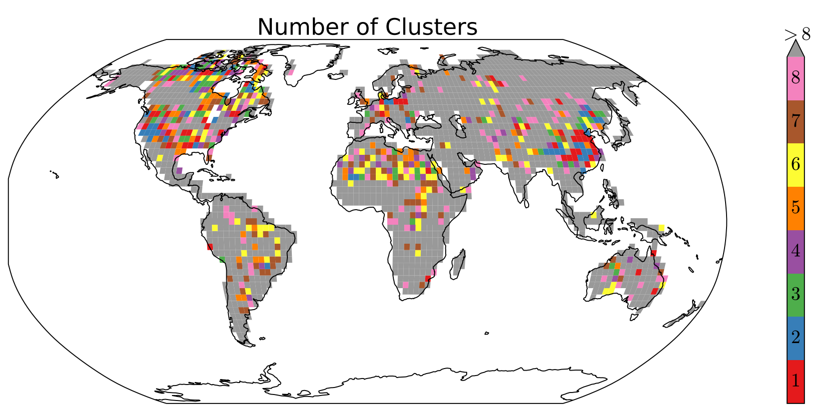

Figure A4. Global distributions of the number of clusters (tiles) found per CLASS-CTEM grid cell when $minPts$ (see Section 2.3) is set to 1% of the number of GSDE data points in the CLASS-CTEM gridcell. The gray regions have >8 tiles found by the clustering algorithms.

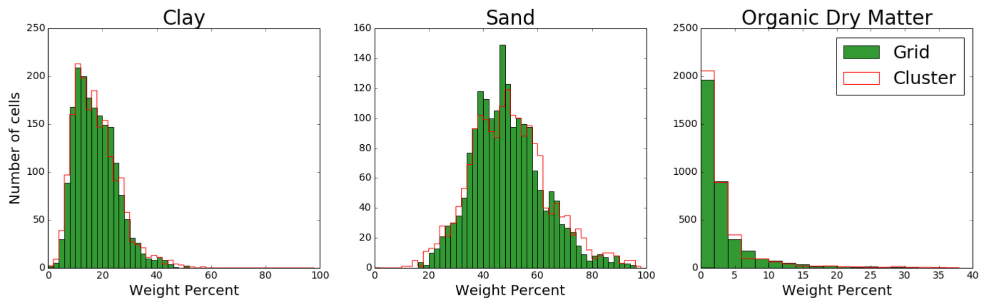

Figure A5. Histogram of the mean clay, sand, and organic matter content for CLASS-CTEM grid cells based on the simple mean value of all GSDE cells (green) or the weighted mean of the clusters within a CLASS-CTEM grid cell (red line).

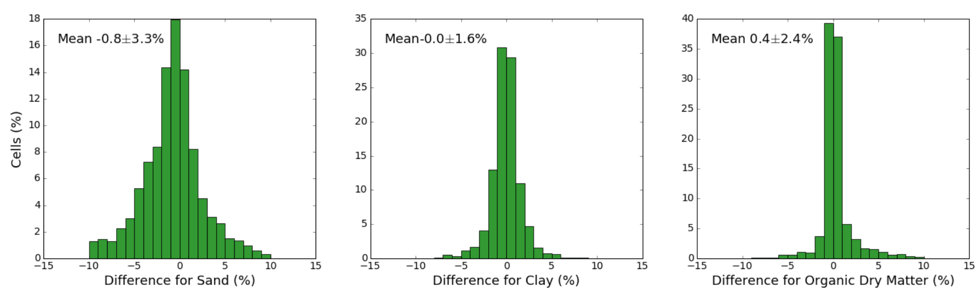

Figure A6. Histogram of the difference between the mean clay, sand, and organic matter content for CLASS-CTEM grid cells based on the simple mean value of all GSDE cells and the weighted mean of the clusters within a CLASS-CTEM grid cell.

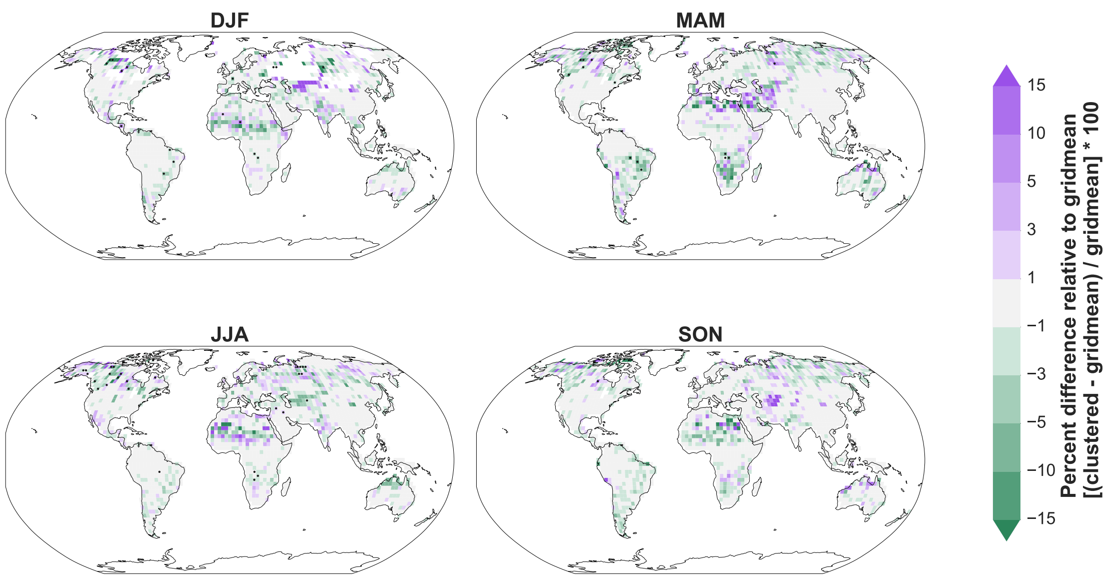

Figure A7. Percent difference in evapotranspiration (ET) between the Cluster and Gridmean simulations by season (mean of 1995–2015). Grid cells with monthly ET of $<10^{-5}$ mm water were masked out to prevent instances of divide by zero and overly large relative differences in regions of very small evapotranspiration. Positive values indicate the Cluster simulation ET is larger while negative values indicate the Gridmean simulation ET is larger. Dots indicate grid cells that are statistically significant (independent two sample t-test p-level <0.01)

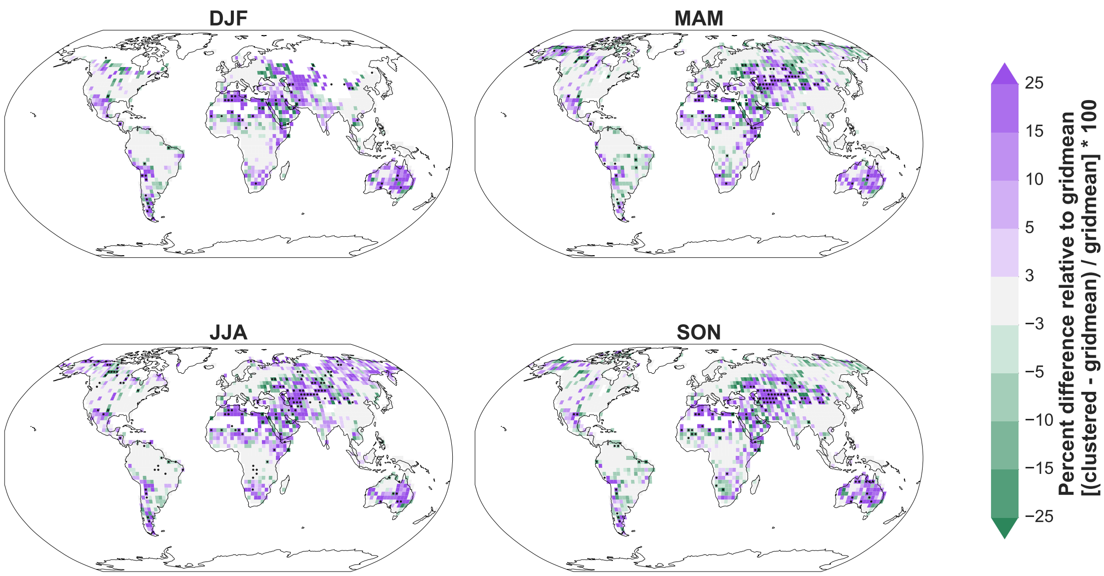

Figure A8. Percent difference in transpiration between the Cluster and Gridmean simulations by season (mean of 1995–2015) following Fig.A7.

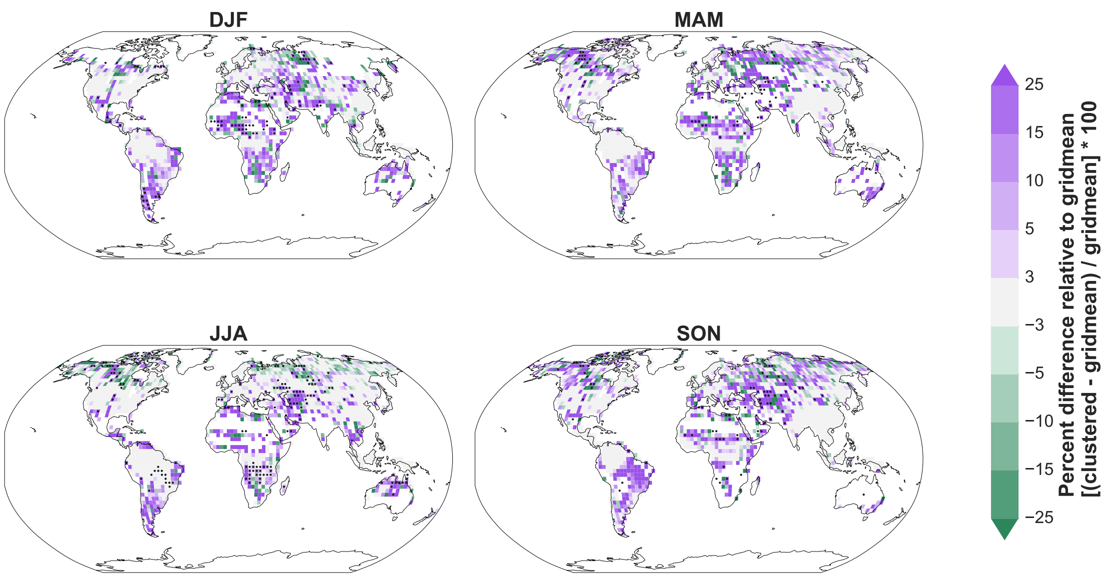

Figure A9. Percent difference in runoff between the Cluster and Gridmean simulations by season (mean of 1995–2015) following Fig.A7.

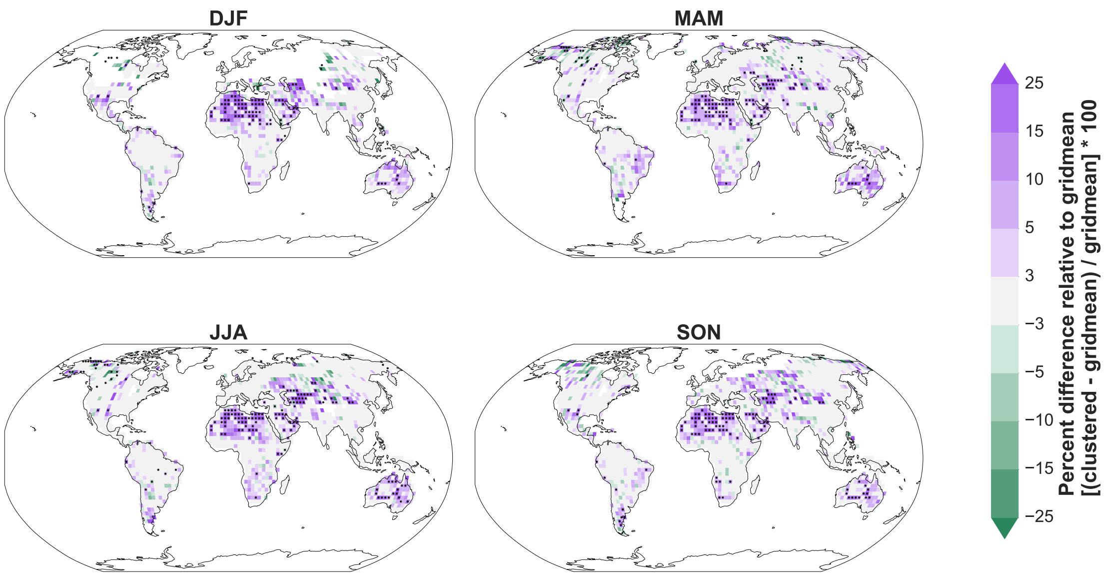

Figure A10. Percent difference in sensible heat fluxes between the Cluster and Gridmean simulations by season (mean of 1995–2015) following Fig.A7.

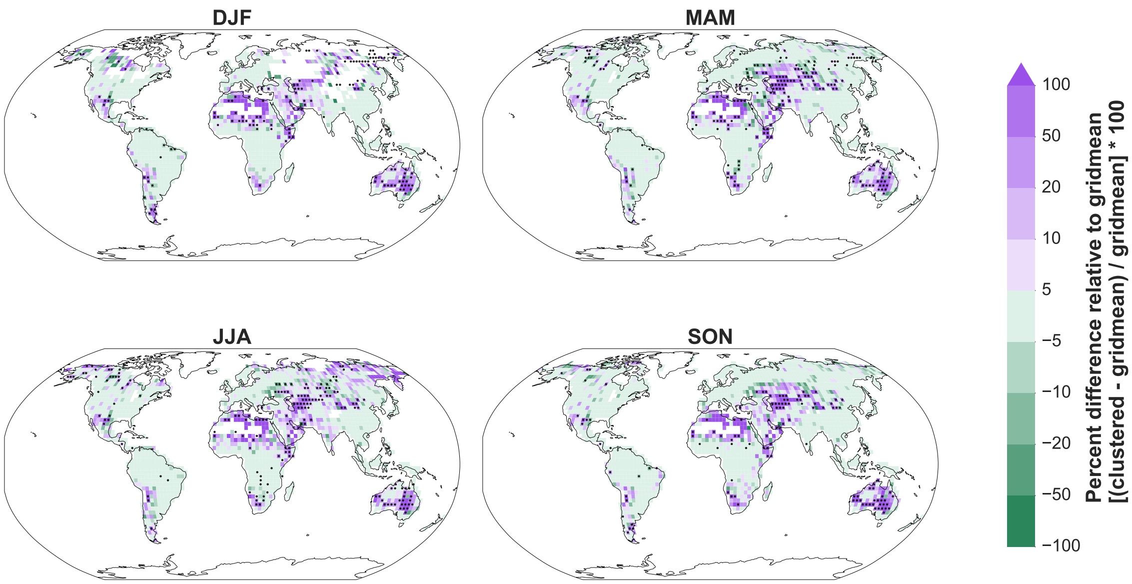

Figure A11. Percent difference in WUE between the Cluster and Gridmean simulations by season (mean of 1995–2015). Grid cells with monthly evapotranspiration of $<10^{-5}$ mm water were masked out to prevent instances of divide by zero and overly large relative differences in regions of very small evapotranspiration. Positive values indicate the Cluster WUE is larger while negative values indicate the Gridmean WUE is larger. Dots indicate grid cells that are statistically significant (independent two sample t-test p-level <0.01).

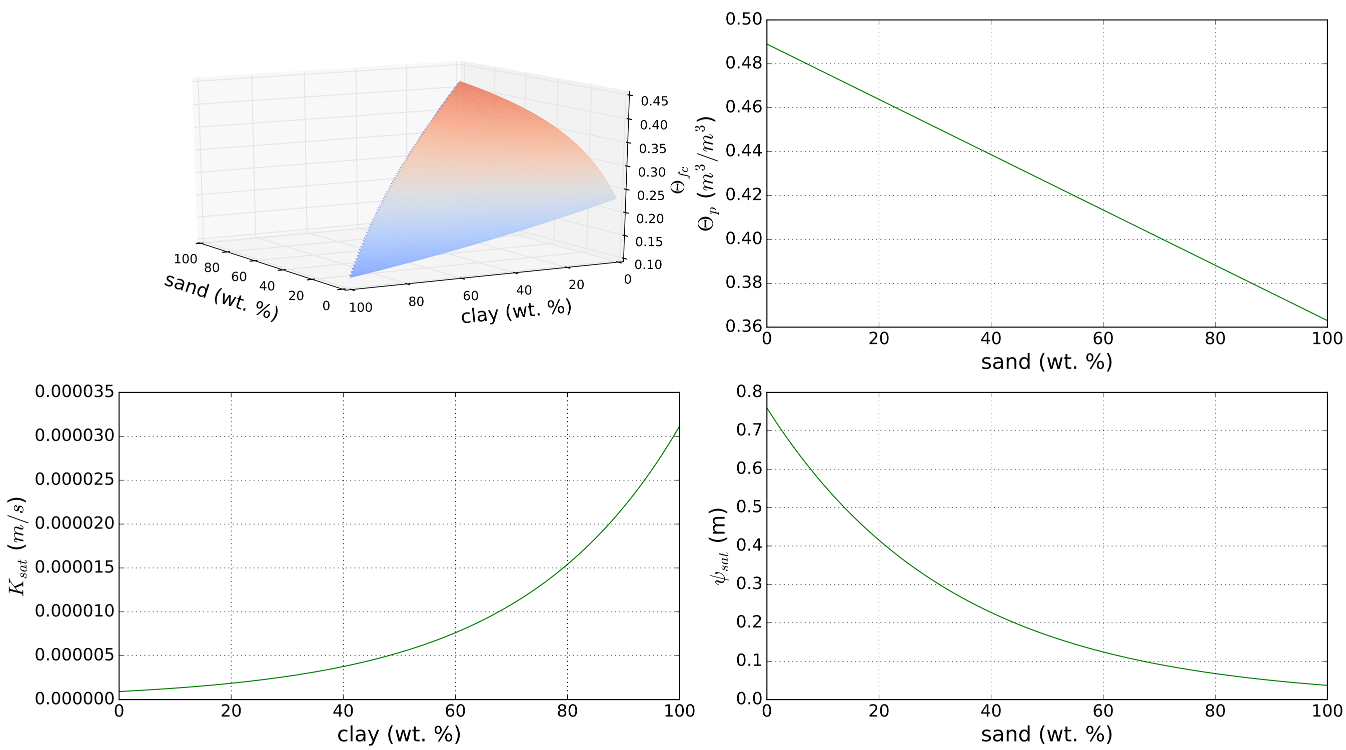

Figure A12. Relationships between soil texture and field capacity (θ_{fc}; upper left), pore volume (θ_p; upper right), saturated hydraulic conductivity (K_{sat}; lower left), and soil moisture suction at saturation (Ψ_{sat}; bottom right) following Cosby et al. (1984) as implemented in Verseghy (2012)