Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2018-213-AC1, 2018

Geoscientific Model Development Discussions
Open Access
EGU

Interactive comment on "A Predictive Algorithm For Wetlands In Deep Time Paleoclimate Models" *by* David J. Wilton et al.

David J. Wilton et al.

d.j.wilton@shef.ac.uk

Received and published: 14 December 2018

We thank the referee for their comments and respond to the points they raise below.

1. "The main confusion I have is on the validation of this approach. It is not convincing that using one reference dataset to train their algorithm, and then evaluate the simulated results with the same reference dataset. It would be necessary to compare with independent inundation products to justify their approach, or the authors need to provide the uncertainty in the estimated inundation using their approach given that there are large uncertainties in wetland extent among existing inundation products (Melton et al., 2013)." There is no training and evaluation in the sense that would normally be understood from a machine learning perspective. For the Eocene results, section 3.2,

Printer-friendly version

Discussion paper

we clearly have no wetland data with which to train and evaluate our predictions. We simply use the coal deposits as a proxy, comparing those to our wetland predictions to give us the best value of K for the maxKNN approach with this particular data set. We are happy to improve the text in this section to make this clearer.

Nor are we using a training set for the modern day test data, section 3.1. These results were included simply to show whether some form of nearest-neighour approach might, in principle, be useful (lines 236-238); we were exploring the potential of this approach. It was a test that if failed would have meant we would not have continued developing a nearest neighbour method; it would have been another unsuccessful attempt along with those briefly discussed in section 2.3. That the method passed this test merely indicated we could explore some form of nearest neighbour method in the context of the Eocene climate. If this is what the referee is referring to, then we will improve the text in section 3.1 to make this clearer.

2. “The logic of this approach is a bit confusing to me. If I understand it correctly, this nearest neighbor-based algorithm implicitly assumes the locations of wetlands should close to each other and inundation is correlated with eight variables the authors proposed. But according to the modern dataset, is there any analysis/evidence prove that this relationship exist” The nearest neighbour approach assumes that sites with similar values of wetland fraction should have some similarity in terms of their values of the 8 climate & vegetation variables we use; or to put it another way, if sites with similar FW show no similarity at all between their values of at least some of those 8 variables, then a nearest neighbour approach will simply not work. There is certainly no simple correlation between FW and those 8 variables in the modern day data, as we briefly explain in our “Initial unsuccessful models” section 2.3; a multiple-linear regression on those 8 variables did not produce a good predictive model of FW. This suggests that any relationship between FW and those 8 variables must be complex. We are happy to add to the text to further explain this, the best place probably being at the end of section 2.3.

3. “Fan (2011) suggest that water table depth is a key to simulate wetland distribution - at least it is an important variable to capture the distribution of peatlands in high latitudes as some of the peatlands don’t show inundated condition but still emit CH4.”

We use soil water content, defined as the amount of water in in the top 1m of soil. This is produced by both vegetation models whereas water table depth is not.

4. “I’m not sure that comparing the simulated wetland distribution with coal deposit can be helpful as the authors have already mentioned some of the limitations using coal deposit. Also, it’s hard to tell how good the fit is from reading Figure 7.” Clearly coal deposits are not an ideal proxy for wetland fraction, but they are all we have. Without them we would have had no way of deciding on a value for K in the maxKNN algorithm. Therefore, despite the limitations, they are useful to explore this approach.

5. “It would be great to address a bit more about the background why it’s important to develop a dynamic inundation algorithm for deep time paleoclimate simulation and what’s the current status of research on this topic.” As explained in the introduction, there is great interest in understanding how the extent of wetlands changed through geological time and what role that could have had on methane cycling. However, there is currently only one model-based approach for deep time paleoclimates (Beerling et al., 2011). The goal of this paper is to explore other methodologies and compare them to this original work, better understanding the potential of the new approaches and the robustness of the previous work. We will add additional text to the introduction to address this further.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-213, 2018.