

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

### **Anonymous Referee #2**

Received and published: 21 December 2018

Review of ‘A predictive algorithm for wetlands in deep time paleoclimate models’, by David J. Wilton et al.

#### General comments

This paper attempts to produce a statistical model (“predictive algorithm”) to estimate the global distribution of paleo-wetlands, and the associated methane emissions. The method (as far as I understand it) is to look at present-day wetlands, and find relationships between various climatic and vegetation variables that can be used to predict wetland extent. It turns out to be necessary to not just look at driving environmental variables at the same location, but to look at nearest neighbours (in parameter space – not spatially). A nearest neighbour model (NN) using just one NN is found to well represent present-day wetlands. However, the authors find that this approach applied

C1

to Eocene climate model data doesn’t produce a very realistic wetland distribution – based on proxy evidence from coal deposits. Using multiple nearest neighbours improves the model performance – three appears optimal.

I am fairly convinced this is a sensible and useful approach, but I must admit to being slightly baffled about the exact methods employed – I found the paper rather unclear in quite a few places. I would encourage the authors to revise the description of the methods to make it clearer. Some clarification on how this approach should be employed by the wider modelling community would also be appreciated – can the method be embedded within ESMs to calculate wetland emissions online? Or is it envisaged as an offline only tool? I wasn’t clear.

If the authors can clarify the methods and address the specific points below, then I believe this will be a very useful method for modellers to use.

#### Specific comments

L43 ESMs must either prescribe CH<sub>4</sub> concentrations as boundary conditions, or “incorporate dynamic methane fluxes from natural sources. . .”. If the latter, they must not only simulate the sources but also the sinks of the CH<sub>4</sub> (i.e the whole budget) in order to reasonably represent concentrations.

L55 ‘. . .no direct observations of wetland extent’ – it should be stated that there are however proxies, that you later utilise (i.e. coal deposits).

L60 ‘. . .mean monthly temperature drops below 0 °C at some point in the year. . .’ I found this slightly confusing. Do you mean if there is one (or more) month in the year below 0 °C, then that grid-cell is classified as producing methane? Clarify.

L71 So you are using DGVMs to simulate vegetation distributions, rather than using present-day observational datasets. It may be worth saying that the DGVMs have (presumably) been evaluated elsewhere.

L68 Is it worth briefly defining wetland? Perhaps earlier. E.g. the RAMSAR definition.

C2

Is it obvious how such definitions translate into a climate model-specific definition? (Water depth, etc.). What is the basis of the modern day reference data set of FW? Can you say it is 'known' or 'observed' FW?

L80 Typo: intercomparison

L87/90 Capitalise Nearest Neighbours or not? (Is it a well enough known method to be considered a proper noun? I don't know, but at least be consistent.)

L105 So SWAMPS is based on microwave satellite observations – what is the observational data that GLWD is based on?

L127 I didn't fully understand the scaling – are the mean/standard deviation global values?

L130 As previous comment – is the global mean 0?

L132 A modern-day test data set. . .

L134 conducted on -> driven by?

L136 Use the same terminology as l132 to avoid confusion, i.e.: "The paleoclimatic assessment of our model was performed using an early Eocene. . ." -> An early Eocene test data set was made using. . . ?

L145 It would be useful to provide a summary table of the test/reference data sets to clarify exactly how you are going to evaluate your approach; I didn't find the current explanation completely clear.

L163 The number of what? Months or grid cells?

L214 Is Rh an absolute or scaled (0-1) value? If absolute, what are the units? Similarly for GPP in the next equation (I guess it must be absolute value to make sense.)

L218 Is TMP soil, surface, or surface air temperature?

L226 Presumably me  $\geq 0$ ? Is there a test for mp  $\geq$  mo? What are the units of me?

C3

L232 'downscaled' – I think the definition of downscaling is to infer something at high resolution from something at low resolution. You seem to be using the word in the opposite sense. I don't think we (scientists) normally use 'upscaled', so I am unsure what to call this (degrading?), but I don't think it is downscaling (also l234).

L318 I think the term 'maxKNN' appears here for the first time and isn't defined. Is it just KNN with  $K > 1$ ? (As suggested by l316.)

L364 In a similar vein to the last comment – why not just 3NN rather than max3NN?

L383 '...both [FW and CH4 emissions] have their highest values in summer months. . .' This is not so clear in Figure 8 for SDGVM. It is clear in Figure 9.

L409 '...their respective impacts of soil water balance. . .'. Clarify. Is this just a typo of -> on?

L409 I got a bit confused here about EVT. It seems EVT is from the vegetation models; but EVT must also be calculated in the underlying climate model – I guess with a much more simplified vegetation scheme. Is there a large discrepancy between the EVT in the vegetation and climate models? Isn't this a bit of a problem? This decoupling of the simulated water budget between the climate model and the vegetation model should be clearly explained earlier in the methods section, and the implications discussed here.

L423 Global monthly mean FW for the Eocene. . .

L572 Figure 1 caption – Annual monthly maximum. . .

L586 Figure 4 caption and y-axes – clarify these are CH4 emissions – what are the units? (Tg CH4/month?)

L628 Incorrect punctuation for list.

---

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

C4