Modeling spatial-temporal dynamics of global wetlands: comprehensive evaluation of a new sub-grid TOPMODEL parameterization and uncertainties

By Z. Zhang, N.E. Zimmermann and B. Poulter

In the present manuscript by Zhang et al., the main goal is to compare the use of three different digital elevation products for wetland model representation using a dynamic global vegetation model (DGVM), LPJ-wsl, and with it, improve the methane emissions in model simulations at a global scale. Traditionally, a topography model (TOPMODEL) is used and for that purpose a compound topographic index (CTI) is needed to describe the topographic profile at a sub-grid scale, which will define the aerial fraction of a grid cell that can be inundated (e.g. wetland). The CTI's distributions are generated from the DEM's and are products readily available at specific spatial resolutions. In this manuscript, the authors instead of using a discrete distribution of CTIs directly in their model configuration suggest to make use of a sub-grid discrete cumulative distribution function (CDF) of the CTIs to calculate the fraction of the grid cell that can be inundated (wetland area fraction F_{wet}). The wetland area fraction is then used in the model and can be directly related to the water table depth. The authors claim that the use of the pre-calculated wetland area fraction, instead of the CTI in global land vegetation models, is less costly computationally and that the parameterization is improved by "providing physical meaning" requiring fewer calibration parameters.

General issues:

Previous works have also fitted the CTI products to functions that represent a grid cell CTI value, such as in Kleinen et al., 2012 and Ringeval et al., 2012. Although this approach sounds reasonable, I am not convinced that by providing the inundated fraction in the grid cell the computational cost is considerably reduced. This might be true for some models but not in all cases and not in all resolutions. Furthermore, if this is true, an extra preprocessing of after the CTI grid cell fitting to obtain the inundated fraction implies an extra step beforehand that certainly adds more errors in the model input. The authors give a step towards this by reducing the uncertainties in the calculation of the maximum soil saturated fraction obtained from the CDFs, by introducing a parameterization to calibrate the maximum wetland fraction ($F^{\mathrm{wet}}{}_{\mathrm{max}}$) with "original" values (F_{max}) obtained from the CDF when the mean CTI is zero.
An interesting contribution from this manuscript is the comparison of the three DEM's (HYDRO1k, GMTED and HydroSHEDS) for wetland simulations in DGVMs, and arise the need of hydrological corrections before its use.

My major concern regarding this manuscript is that I find it still too descriptive for the model setup and I believe is still out of the scope of Biogeosciences. Despite the authors made an effort by adding few sentences regarding the analysis of modeled methane fluxes to test the wetland representation from the model, the authors rarely refer to the CH_4 fluxes application throughout the manuscript. The focus of the manuscript is still to simply compare the three DEM

products in their model setup and improve the F_{max} parameter in TOPMODEL, but they do not make any strong reference to the evaluation of methane fluxes or discuss further other papers that make this analysis. A clear example of this, are in the specific aims of the manuscript listed at the end of section 1, which are only focused on model improvement based on the analysis of using three different DEM's. Also in Discussion and Conclusions there is nothing regarding methane emissions. Therefore, I still find difficult to agree that this manuscript should be published in Biogeosciences in the current state, and I believe is still suitable for GMD.

Despite this, I made some comments that the authors may find useful to improve the current version of the manuscript. Some of the statements made by the authors are ambiguous and it needs several language corrections, this makes it sometimes hard to understand what the authors really mean. The wording is particularly hard to follow in the Discussions section, although I made some specific comments, I suggest that the authors revise carefully their sentences and re-arrange the wording for a clearer reading.

Still if these errors are corrected, and comments here included answered, I encourage the authors to make more emphasis in the CH_4 fluxes, e.g. include a specific aim in section 1 and discuss further other works that had published CH_4 fluxes using similar approaches (e.g. Kleinen et al., 2012). Also compare to more representative studies for the regions of interest with other methodologies (see my comments below for this). Therefore, I cannot support at this point the publication of this manuscript in its current form in Biogeosciences.

Major comments:

- The full name of an acronym should be always stated when is first mentioned in the paper. I could not find the full name of LPJ-wsl or LPJ-DGVM, please write it in full either in the Abstract or in the Introduction when is first mentioned (P17957, L23?). There are also other acronyms that should be written its name in full, please check this throughout the manuscript.

L14 – In the sentence: "… which has been proven to at least partly cause biases due to limited spatial resolution…", I don't think 1 km is a limited spatial resolution for such data sets, please elaborate here what the authors really mean with these sentence.
L26 – mention some examples of physical processes the authors refer to in this line (e.g.)

P17967-L26; P17968, L1-2. Although the correlation between the model simulated frozen-days and the in Fig. 3 agrees well, the authors speculate that the low correlation in East Siberia could be due to the nature of the data, while in the satellite observations it is included the ice condition in the vegetation canopy, snow layer and frozen water in the upper soil layer, in the model it is only considered the frozen state of the top soil, but if this is true, why in the southern regions of Siberia the correlation seems to agree better? I would expect that this behavior remain at least in most part of northern latitudes.

- It is misleading the explanation of F_{max} and $F^{wet}{}_{max}$. To what I understood from the manuscript, F_{max} is taken fro the satellite observations and used to calibrate $F^{wet}{}_{max}$ which is then used to obtain the wetland area fraction F_{wet}. However, the authors repeat in the manuscript that what they propose is a "calibration of F_{max}", shouldn't be $F^{wet}{}_{max}$? Please correct me if I am wrong or otherwise, be more explicit and careful in the description of the method and correct where necessary in the manuscript.

- The newly available DEM product from the Centre for Ecology and Hydrology (an improvement from HYDRO1k from 30" res to 15" res) https://data.gov.uk/dataset/high-resolution-global-topographic-index-values1, should be at least mentioned and discuss how this new product can improve the representation of wetlands at global scale and how this can be combined with the F_{max} (or $F^{wet}{}_{max}$?) calibration proposed in this manuscript.

Specific comments:

P17954,
L2 – spatio-temporal
L16 – Define here what DEM stands for

P17957,
L10 – Add citation year for Ward and Robinson (2000)
L12 – is really 1 km limited?
L26 – e.g. physical processes

P17958,
L16-17 – remove parenthesis in Hodson et al., 2011 AND Wania et al., 2013
L17 – "and is a function of two scaling …"
L17 – the authors does not define f_{ecosys} and $r_{CH4:C}$ in the text, nor say how they are obtained
L24 - delete "contributed as"

P17959,
L18 – move parenthesis before "Cosby" to before "1984" (Cosby et al., (1984))

P17960,
L20 – add in parenthesis after the name the acronym CTI

P17961,
L14 – delete "furthermore"
L18 – to my understanding a gamma function can be also exponential, and this in the end is a similar treatment than the gamma function, thus not reducing the computational cost.

P19762,
L3 – "… topographic information generated by fitting the …"

L4 – add a comma after CTI
L4 - here the authors should be more specific on "observed maximum wetland fraction" stating that this information was obtained
L15- write the meaning here of SWAMPS-GLWD

P17963,
L4 – write the meaning of HWSD
L4 – reference for the HWSD soil texture database?
L8 – replace "more" by "mainly"
L10 – latitudes
L19 – write the spatial resolution of the DEMs after they are mentioned in the following lines

P17964,
L14-20 – Here it is a misleading whether the authors generated ONE single CTI maps based on the three DEM products or if there were THREE CTI maps been one per DEM product. This becomes confusing along the manuscript, particularly arriving at Figure 7. See my comment below for it.
L20-25 – Here it is not really clear in the paragraph if GMTED was also used to generate the global CTI map despite was not hidrologically corrected as the other two DEM products? What do the authors mean with "retaining GMTED DEM without hydrologically correction" ?
L25 – change "hydrologically" by "hydrological"

P17965,
L4 – "generating a global catchment map"
L9 – "The description of the Dem products used in this study are summarized in Table 2"
L13- here the word spin up is separated, while in L18 is a single one (spinup), the correct should be separated

P17966,
L27 – Poulter et al., 2015

P17967,
L24 – "in those regions"

P17969,
L4 – correct here and throughout the manuscript that CH4 is with subscript (i.e. CH_4)
L5-10 – As stated in the caption of Figure 6, the authors should mention here the DEM product used is Hydro-SHEDS for TOPMODEL. However, this is confusing since earlier in the manuscript the authors mention that they generate a mean CTI map of the three DEM products to actually "calibrate" TOPMODEL, so why here it is only comparing Hydro-SHEDS?
L5-10 – I would try to avoid using the expression "calibrated TOPMODEL" and "non-calibrated TOPMODEL" for the correction on the maximum fraction of wetland extent. This is what it was actually corrected ($F_{\max}$) but TOPMODEL itself not only provides the maximum fraction.

L14-19 – I am not convinced with the comparison of results from the West Siberian Lowland to the CARVE observations in Alaska. Although both are boreal wetland regions, there are published works that match better the region of interest in question. I would rather use for example previous observations at least in the Siberian region with other techniques like Eddy covariance like the works of Parmentier et al., 2011 (J. of Geophys. Res.) or Wille et al., 2008 (Global Change Biology).
L22-25 –Figure 7 is not really well explained here nor in the Figure caption. What do the authors mean with the prefix BASIN and GRID? This part needs more detailed information in the simulations description before it is presented in the results. If they are the aggregation schemes they briefly mention in the introduction, then the authors need to refer to them by their name there. Furthermore, the authors mention "both datasets" but they should be specific to what they mean (e.g. the results from the simulations with BASIN and GRID aggregation schemes?). I honestly, don't see much the sense of this figure plus it is hard from it to visually look at the "uncertainties" of the parameterization.
L27 – replace "differing" by "different"

P17970,
L5 – replace "sensitivity" by "sensible"
L10-12 – I thought GMTED was not hydrologically corrected?
L11- Add the degrees symbol to 60 N
L16-17 – replace "estimation" by "estimates"
L18 – replace "paddy" by "paddies"
L21 – replace "digitalized" by "digitized"
L22 – move the word "directly" after " … when comparing …" at the end of line 20
L25 – I guess it should say " ... due to permanent wetlands that are hard to detect by GIEMS."
L27 – please elaborate here more about the satellite inundation data sets, what the authors really mean with "non-specific measurement of inundation"?
L28 – This paragraph is also misleading, do the authors meant to say that the definition of wetland in this work is in agreement to the definition used by the National Wetlands Working Group? Please also reference this in the reference section as: National Wetlands Working Group, 1988. *Wetlands of Canada,* Ecological Land Classification Series, No. 24. Canada Committee on Ecological Land Classification. Sustainable Development Branch, Environment Canada and Polyscience Publications Inc. Montreal, Quebec, Canada.

P17971,
L7 – SON is not a season but the acronym of a list of months that accumulated corresponds to a season (autumn), please rephrase correctly (replace the word seasons by months).
L9 – what do the authors mean here with "masked estimates"? ambiguous
L10 – pluralize latitude
L11 – an area cannot be higher, only larger
L12 – rephrase, seasons are not unfrozen, you can instead say "… from longer periods of unfrozen and relatively water saturated soils in the model data"
L16 – replace "seasons" by "months" (or "SON seasons" by "autumn")
L22 – replace "underestimated" by "underestimate"

L24 – replace “estimates” by “data sets”
L24 – replace “base” by “based”

P17972,
L4 – here the authors refer to the “grid” experiments as “tile-based”, please keep consistency with your nomenclature here and throughout the manuscript
L10 – “the” Pearson’s correlation coefficient
L13- Define what is a “Transcom region”?, it was only mentioned before in the figure caption of Fig. 2 and also in caption of Fig. 8
L17-18 – This sentence is a confirmation of previous works, like Kleinen et al., 2012. Taking this into account I would rather make more emphasis throughout the manuscript that the aim of the correction in the maximum wetland extent is to actually improve the representation of wetlands by the models using TOPMODEL at a regional scale. This has to be highlighted even in the abstract section.

P17973
L7-8 – wording of sentence a bit strange, I suggest: “ … TOPMODEL with calibrated parameters as described in this study, allows the dynamical simulation of wetlands, in particular their geographic location and extent.”
L9-13 – this sentence is particularly hard to follow, please re-arrange the wording to make it clearer
L21 – strange wording, do the authors mean: “… in absolute values, which is necessary for global wetland modeling.”?. I would modify this sentence since is confusing in the way is written now.
L23 – change to: “…because the physical processes are described in a robust way”
L25 – “allows the retrieval of the maximum water saturated fraction ($F_{\max}$) of a model grid cell, which is defined by …”

P17974,
L2 – Replace “This” by “The”
L14 – pluralize “application”
L15 – pluralize “parameterization”
L16 – “fine scale”
L16 – “which complicates the comparison to inventories”
L17-22 – the wording of this paragraph is wrong, and hard to follow, please correct it.

P17975,
L14 – “ … size and location that make hard to reconcile a single definition for wetlands”
L15 – pluralize “parameterization”
L19 – pluralize “area”
L25 – elaborate in “limitation therein”
L18 – and complete paragraph should be moved to the introduction since this is a better start for the background knowledge and motivation of this study. This paragraph will certainly improve the flow of the method if it is moved forward in the manuscript.

L26 – move “during the last decade” to the beginning of the sentence

P17976,
L2 – “from regional to global scales”
L2 – The reference Lin et al., must be separated as: Lin et al., 2010; Lin et al., 2013; the first one corresponds to Kairong Lin and the second to a different author (Shengpan Lin)
L6 – “benefit”
L7 – “creating a more realistic representation …”
L9 – “This is supporting the ideas of …”
L16 – “closed depressions”
L24 – “As a result”

P17977,
L23 – “describe”
L25 – “need”

P17978,
L27-28 – “Remotely sensed global inundation is prone to underestimate small wetland areas,…”

P17979,
L3 – “This raises the need for benchmark dataset useful to generate accurate products with lower uncertainties”
L14 – “and captured well the spatio-temporal …”

References
- P17980, L24 – Update the reference by Bohn et al., 2015a (not in discussion anymore)
Missing reference USGS, 2000 (cited in P17964, L5-6)

Figures
Besides specific comments on figures’ captions mentioned before, here are some more comments.
Figure 1 – replace the symbol lambda with the horizontal line on top by lambda with subscript m as in the text. Also in the label of the x-axis lambda should have the subscript l corresponding to the local CTI value. Change this also in the legend of the figure
Figure 2 – the figure caption must be considerably improved, by making reference to the panels and their meaning, also by editing the text (italics, subscripts, etc.)
Figure 4 – add year “Tanocai2009” in both title of subplot and caption
Figure 5 – Include in the caption the area of study (e.g. Amazon River Basin or Lowland Amazon Basin)
Figure 6 – Change the units of CH4 emissions with the area unit before the time unit (e.g. g CH_4 m^{-2} yr^{-1})
Figure 8 – replace “variation” by “variability”