Geosci. Model Dev. Discuss., 8, C4387–C4422, 2016
www.geosci-model-dev-discuss.net/8/C4387/2016/

Geoscientific Model Development Discussions
Open Access

Interactive comment on "Integrating peatlands into the coupled Canadian Land Surface Scheme (CLASS) v3.6 and the Canadian Terrestrial Ecosystem Model (CTEM) v2.0" *by* Y. Wu et al.

Y. Wu et al.

qiaoyuanwu@gmail.com

Received and published: 21 April 2016

We thank an anonymous reviewer and Dr. Nigel Roulet for the time and the care that they have

taken in providing helpful comments on our manuscript. Their comments, followed by our

responses in bold, are listed below.

Review 1 (Anonymous)

This manuscript presents a description and initial evaluation of a new peatland model

built for

the Canadian Land Surface Scheme (CLASS) and the associated Canadian Terrestrial Ecosystem

Model (CTEM). The manuscript does an excellent job introducing the importance of accurately

modeling peatlands within earth system models, and includes a well-written, brief review of

existing peatland models before presenting the new model developments. The conceptual

description of the new peatland model is well written and clear, and the evaluation against eddy

covariance flux measurements, water table measurements, and soil temperatures is well

presented and generally well designed. The model itself appears to do a good job of

incorporating current understanding of peatland vegetation and soil processes, and should be a

useful tool for simulating peatlands going into the future. I think there is some room for

improvement in the technical descriptions of the model equations and the flux data used to

evaluate the model. I also think that the conclusion that separate parameterizations for bogs and

fens are unnecessary for this model is not adequately supported by the results, and may need to

be reexamined or supported with more evidence.

We thank the reviewer for his/her time and for these thorough and thoughtful comments.

0.1 Model equations

I think some of the equations may need another step of proofreading. Some of the notation is

unclear, and there may be some errors in the equations. Specifically, Equations 10-17 may need

another look.

Eq. 10 and 12: The integrals do not look correct. Integrating temperature over depth doesn't

make much sense, unless it's intended to be an average temperature with depth. In that case, the

integral should be divided by the depth that it's being integrated over. In addition, dwt appears in

these equations and is never defined. Is it meant to be zwt? Q10,a and Q10,o are not explicitly defined.

I assume that these are derived from the Q10 function in Eq. 11 using either Ts,a or Ts,o from Eq. 12,

but this should be explicitly stated. Finally, I think the equations for fT,o and Ts,o should be

integrating from 0 (the soil surface) instead of 1 (which would be starting at 1 m depth).

We thank the reviewer for pointing out these inconsistencies. We have changed dp and dwt

to zp and zwt, and inserted the following sentence: "The Q10 values of the anoxic and

the oxic

zones of the soil are indicated as Q10,a and Q10,o." We have modified the equations as follows:

- Changed equation 10 to

- Changed equation 12 to

As a general point, it's not easy to visualize the k values resulting from equations 13 and 14.

These values are central to the resulting heterotrophic respiration, and a crucial part of the

argument that different fen and bog parameterizations don't matter for the model's accuracy. I

would suggest adding a figure that shows how k varies with water table depth, for both fen and

bog parameterizations. I tried making my own figure (which I've attached), but I'm not sure it's

totally accurate. Having such a figure in the paper would really help readers interpret the general

behavior of the model with respect to water table, and would be really helpful for understanding

why bog and fen parameterizations do or do not cause differences in simulated fluxes. When I

plotted these equations, the anoxic decomposition rates had some very sharp transitions at water

table depth of 0.3 m, which didn't seem very realistic.

We have added a figure illustrating the variation of the k values with water table depth, as suggested. The sharp transition of anoxic decomposition rate at 0.3 m is based on Figure 1b in Frolking (2001), reproduced below. As noted in section 2.1 of our manuscript, this value is widely accepted as a representative estimate of the depth dividing the acrotelm and catotelm. In reality, of course, this depth will vary among peatlands. When our peatland model is implemented in climate mode, it is planned that spinup tests will be run to assess the spatial variability of this depth, and adjustments will be made equations 13 and 14 if necessary. We have added the above explanation to the paper.

As a final issue related to these equations: on page 10100, line 22: The model described here applies a factor of 0.025 to anaerobic decomposition, citing Frolking et al (2010). In Frolking et al (2010), the decomposition rate of anoxic carbon is in fact 0.001 (see Table 2 in that paper, where the parameter is described as "decomposition rate reduction factor at 'full persistent' anoxia.") The value of 0.025 used here actually appears in Frolking et al (2001), Table 1. Note

that in that context, 0.025 is the value used for bogs, and a value of 0.1 is used for fens. In this manuscript, the bog value of 0.025 is used for both peatland types, and as well as being referenced to the incorrect paper. This seems like an important omission given the later argument that fens and bogs are not significantly different in this model.

We have added the reference to Frolking et al. (2001). The value of the decomposition rate reduction factor is a matter of some uncertainty, varying in the Frolking papers as the reviewer has pointed out between 0.001, 0.025 and 0.1. In the absence of empirical support for the choice of which value to use, or for using different values for bogs and fens, we calibrated it based on a set of experimental simulations at the Mer Bleue bog (MB-Bog) and the Alberta treed fen (AB-fen). We found that using a constant value of 0.025 produced the best soil respiration results at these two sites. We have inserted the above explanation in the paper and also added a test in section 4.5 demonstrating the relative insensitivity of the model to this parameter (see the plot below).

0.2 Flux measurements used for evaluation

I think the origin of the eddy covariance fluxes used for evaluation should be described in more detail. In the manuscript, site parameters are listed in Table 4, but there is very little information

about the origin of the fluxes. Were they downloaded from the Fluxnet database, or individually

contributed by site PIs? Were they the result of standardized Fluxnet processing, or individual

site processing procedures? What FLUXNET-defined level of data were used? What kind of

filtering, gap filling, and quality control were done? Given that the observed fluxes in Figures 7-

9 are quite noisy, and the clearly unrealistic QE for one site-year in Fig. 5, it's important to know

whether these fluxes were screened for common sources of unrealistic values in eddy covariance

(low turbulence, wind directions identified as unrepresentative, equipment problems). Some of

the large outliers in Fig. 9 could be related to suboptimal atmospheric conditions, and it's

important to know whether these measurements were screened for these types of known issues

before being compared with the model.

ER and GPP are not strictly measured using eddy covariance, but are derived from NEP

measurements using a range of partitioning techniques. A commonly used method is to fit

nighttime NEP to a nonlinear function of temperature, and/or daytime NEP to a nonlinear

function of PAR (see Stoy et al, 2006, Desai et al 2008, and Lasslop et al 2009). If this

partitioning method was used, it makes comparisons with models problematic because the

modeled values are being compared to another [dataconstrained] model rather than to actual

observations. It's really important to describe the partitioning method so readers can

appropriately evaluate the results.

When calculating average fluxes, eddy covariance measurements are typically gap-filled,

because varying atmospheric conditions, equipment issues, and quality control invariably

produce gaps in data. What kind of gap-filling was applied to these eddy covariance

measurements before they were compared with the model results? Gap filling is usually

conducted in tandem with ER and GPP partitioning, and can introduce the same nonlinear

models to the dataset (see Moffat et al 2007 for a comprehensive review). Was any gap filling

applied to latent and sensible heat fluxes? If not, daily sums could be biased because data

availability is generally lower at night than during the day.

There is no discussion of the inherent uncertainty in eddy covariance measurements, which is

highly relevant when they are being used to evaluate a model. See Richardson et al (2006) for a

starting point.

We added the following paragraphs to section 3.1:

Data were obtained from the FLUXNET database (http://fluxnet.ornl.gov/). For each site

and for each downloaded variable, the highest available data level was used. The

meteorological drivers for the model were obtained from level 4 (gap-filled and quality-

controlled) data, except for the wind speed, which was obtained from level 3 and surface

pressure from level 2 data. Carbon fluxes were obtained from level 4 daily average data

when available. The observed GPP and NEP in the FLUXNET database were derived from

the observed NEP and the assumed relation between NEP, temperature and

photosynthetically active radiation (PAR). The remaining fluxes were averaged from half

hourly level 2 and level 3 data.

In the model evaluation, it must be borne in mind that eddy covariance measurements of

turbulent fluxes of energy, water and carbon are subject to inherent uncertainties and

errors related to atmospheric conditions such a low turbulence and wind direction, or to

equipment malfunction. For this reason we selected a relatively large number of test sites

with multi-year datasets, and focused on long-term averages for the validation. We also

included in the evaluation variables such as water table depth, soil temperature and snow

depth, which are not dependent on turbulent flux measurements.

0.3 Conclusions regarding differences between bogs and fens

Figure 13 shows differences in r2 and RMSE after changing model parameters related to the

separate fen and bog parameterizations. From the K-SWAP test, it is clear that these separate

parameterizations do not significantly change the model's fidelity to observations (measured

using those two error metrics) over the time scales being investigated here. Based on this, the

authors conclude that "it is not necessary to distinguish between fens and bogs." I think this

conclusion is not supported by this analysis for a few reasons.

1. These results do say something about the specific parameterization being used here, but that is

not enough to draw general conclusions about modeling fen and bog ecosystems. Based on the figure I attached, it's clear that, with this parameterization, decomposition rates below the water

table are so low as to be essentially negligible. Any differences between fen and bog

decomposition rates below the water table would therefore have very little influence on total

fluxes. However, the previous manuscript that is the source of a key parameter fanoxic in fact had a

very large difference between bogs and fens originally (which is not reproduced in this

manuscript). It's possible that a different parameter set could yield very large differences

between simulations of bogs and fens. So, it's correct to say that this model with these

parameters does not predict much difference between bogs and fens. But a different

parameterization that produces equally good (or better) results compared to observations might

be much more sensitive. I think it's premature to conclude that the difference between fens and

bogs can be ignored entirely.

As noted above, we have added to the sensitivity analysis a test of the effect of varying fanoxic.

We have also included cumulative plots of the carbon fluxes for this and the other two

sensitivity tests (reproduced below), to demonstrate that for our purposes at least, and to a

first order approximation, we can neglect the differences between fens and bogs for

our

particular application in climate models.

2. These simulations incorporated differences between bogs and fens that a global model would

not have access to. Specifically, the plant functional types used to drive the model runs are

different between fens and bogs, and could drive large differences in global model simulations

depending on what vegetation is assumed to dominate different peatlands. Real fens and bogs

have very different dominant plant communities, hydrology, and soil properties that can drive

differing ecological behaviors (for example, see Sulman et al 2010). On the other hand, some

studies have concluded that peatland type is not the primary driver (e.g. Humphreys et al 2006).

A review of literature related to ecological differences between fens and bogs and how they

might affect or not affect model simulations would really add to the discussion. The section of

the discussion addressing this issue (Section 4.5) does not contain any citations to literature

addressing observed contrasts or similarities between fen and bog ecology and biogeochemistry, and this argument is begging for some more context.

We agree that real fens and bogs have very different plant communities, hydrology etc. In global climate model applications, these vegetation distributions will be either assigned on the basis of global land cover datasets, or derived from spin-up runs with a dynamic vegetation model. In either case, the location of peat soils will be specified at the model initialization stage, and the various peatlands will develop either bog or fen characteristics depending on the climatic forcings. We have added the following sentences in the introductory section outlining the differences between bogs and fens:

Peatlands can be classified as either fens or bogs. Bogs are dependent upon precipitation for water and nutrients while fens receive additional contributions from ground and surface waters (Rydin and Jeglum, 2006). The different sources of nutrients between bogs and fens leads to differences in their physical state including hydrology, soil and water chemistry, vegetation, and nutrient availability. These differences can lead to differences in the fluxes of carbon from these fens vs. bogs, e.g. fen methane emissions are more sensitive to vegetation type but less sensitive to temperature than bogs (Turetsky et al. 2014). Fens

generally produce the most methane with water tables at or above the peat surface, while bogs produce the most methane with the water table below the peat surface (Turetsky et al. 2014).

Rydin, H. and Jeglum, J.: The Biology of Peatlands, Oxford Univ. Press, Oxford, United Kingdom., 2006.

Turetsky, M. R., Kotowska, A., Bubier, J., Dise, N. B., Crill, P., Hornibrook, E. R. C., Minkkinen, K., Moore, T. R., Myers-Smith, I. H., Nykänen, H., Olefeldt, D., Rinne, J., Saarnio, S., Shurpali, N., Tuittila, E.-S., Waddington, J. M., White, J. R., Wickland, K. P. and Wilmking, M.: A synthesis of methane emissions from 71 northern, temperate, and subtropical wetlands, Glob. Chang. Biol., 20(7), 2183–2197, 2014.

3. The evaluation shown in Figure 13 is not really adequate to establish that there is no important difference between results using fen and bog parameterizations. The only data presented are RMSE and r2, which only allow evaluation of the results with regard to a quite noisy observation-based dataset. It's quite likely that the parameter change introduces a small but significant persistent bias in heterotrophic respiration. This might not show up over short (several

year) time

scales, but could lead to large differences in peat carbon pools after decades or centuries of

integration (which are the time scales of greatest interest for peatlands). Because eddy covariance

data includes inherent uncertainty due to turbulence and micrometeorological variations, even an

important difference in model predictions could be obscured by this minimum noise level in the

analyses used here. It would be much more illuminating to see a comparison of modeled ER, or

cumulative NEP, with the different parameter sets in order to evaluate how sensitive the model is

to these differences. Even a comparison of time series between model and observations might

reveal some persistent biases at seasonal or annual time scales that are too small to show up in

the total RMSE and r2 numbers.

We agree that plots of cumulative NEP provide instructive additional information, and we

have added them to the sensitivity analysis as shown in the plots presented above.

0.4 Additional specific comments

- 10091, Line 10: A net C uptake of 3.3 GtC/year compared to the 5.0 GtC/year net C uptake

seems awfully high. Are these estimates directly comparable?

The net C uptake value of 3.3 GtC/yr was calculated using a maximum weighting method,

which has 46 GtC of biomass of bryophytes, in contrast to 4.0 GtC in the average scenario.

This result is thus not directly comparable, and we have removed it.

- 10094, line 22-28: The discussion of peatland and non-peatland fractions and PFT fractional

cover seems out of place, since the rest of the manuscript only discusses single-point simulations.

Was this sub-grid-scale heterogeneity actually included in the simulations? If so, what basis was

used to determine peatland fractions, and fractional PFT cover?

Sub-grid scale heterogeneity was not included in the simulations reported in this paper.

The section in question describes how the peatland model will be implemented in future

coupled climate model runs. We have revised the wording to make this clearer.

- 10096, line 15: It would be helpful to have the units for wm here.

We have added the units, which are kg water per kg dry mass

- 10098, Equation 7: θm does not seem to be defined anywhere. Is this the same as wm?

We have added the definition: θm represents m3 water per m3 soil.

- 10098, line 20: I think the 4.6 factor should have units of μmol m$-$2s$-$1 per W m$-$2

We thank the reviewer for pointing this out, and have made the correction.

- 10099, line 7-11: Were fractional PFT coverages included in these simulations? How were they

parameterized based on the limited land-cover data from sites? Peatlands typically have open,

patchy vegetation. Did the model incorporate this heterogeneity? Were multiple overlapping

PFTs used for each site, or just one?

The model has three possible peatland PFTs parameterized. As noted in section 3.2, the

FLUXNET database was used to assign values to the areal coverage of the PFTs for each

site. We have added a line in Table 4 showing the total fractional vegetation coverage.

- 10101, line 1: What is the model time step?

The time step is 15-30 min for the calculations associated with CLASS and daily for those

associated with CTEM. We have clarified this in section 2 of the manuscript.

- Equations 13 and 14: What is the justification for the 0.3 m cutoff? It seems fairly arbitrary.

The text says these equations are from Frolking et al (2001), but the table says the parameters are

from the McGill Wetland Model.

See our response in the "Model equations" section above. We have added the Frolking

et

al. (2001) reference to the table title.

- 10101, line 13: There is no equation for Chum. Is it just a constant rate, or a fraction of

decomposition?

The humification rate (Chum) is an assumed fraction of decomposition which varies by PFT.

The values of the coefficient are provided in Table 2 (variable "humicfac"). We have

added this explanation to the text.

- 10105, line 8-17: Were the parameter changes applied before or after spinup? If they were only

applied to the spun-up model, than any significant changes that would have accumulated over

100 years would be ignored. These could be important in an earth system modeling context.

The parameters were changed prior to each spinup.

- 10106, line 11-13: Hummock-hollow topography is very typical of bogs. Did this affect any

other study sites? I think it's worth discussing this issue in more depth, with respect to all of the

sites and how these topographical variations could affect the model. See Dimitrov et al (2010),

Baird et al (2009), Loisel and Yu (2013), etc for some good discussions of issues related to

microtopography.

We neglected the effects of hummock-hollow topography because the model has been developed for global simulations, and on global scales there is no good information on how to parameterize this. Wu et al. (2013) found that for the McGill Wetland Model, microtopography was not very important for upscaling the net carbon exchange. In the global peatland model by Wania et al. (2013), hummocks and hollows were included but were arbitrarily assigned a distribution of 50%-50%. We decided that considering the lack of information on fractional coverage of hummocks and hollows globally, their inclusion was not warranted in our model.

- Section 4.2, 4.3: All of these evaluations used daily averages, correct? It might help to state this explicitly at the top of the section.

These evaluations are based on daily averages or daily totals. We have clarified this in the manuscript.

10108, line 23: NEP in the model is calculated by subtracting ER from GPP. In eddy co-variance measurements, NEP is the measured quantity, while ER and GPP are derived from NEP

(and therefore may contain additional errors).

We have added the following to the beginning of section 4.3:

In eddy-covariance measurements, as noted in section 3.1 above, GPP and ER are obtained by partitioning the observed NEP on the basis of empirically derived relationships. In the case of modelled carbon fluxes, on the other hand, NEP is calculated by subtracting ER from GPP, therefore the error in the NEP simulations accumulates the errors in GPP and ER. Bearing in mind these caveats . . .

10109, line 2: "Model errors for the extreme values at these two sites" implies that the eddy covariance values are "truth". Eddy covariance is an inherently noisy measurement because it relies on atmospheric turbulence. Furthermore, large spikes could be due to inadequate screening for poor meteorological conditions. I wouldn't place too much confidence that these big outliers in eddy covariance fluxes are actually real ecological fluxes. This is where it's important to check what kind of screening was done on the flux measurements.

We have reworded the sentence as follows: "The discrepancy with the modelled values, contributing to the low r2 values for these two sites, might be due either to weaknesses in

the model or to inadequate screening of the eddy covariance measurements."

10110, line 5-7: If this site were included in the Figures 7-9, readers could see what was going on

much more easily.

This site was not included because since the high NEP is largely a product of the tree cover,

it clouds the interpretation of the peatland model performance, and is therefore not as

informative in this regard as the other sites.

10111: I think a bit more explanation of the Taylor diagrams would be helpful here. I don't think

they're really common enough to forego a sentence or two about how to read them.

We have added the following explanation:

Taylor diagramsÂăprovide a graphical summary of how closely modelled data match

observed data (Taylor, 2001). The radial spokes represent the level of correlation and the x

and y axes show the standard deviation. The standard deviation of the observations is

plotted on the x axis, and the RMSE of the modelled values is indicated by the concentric

contours around this point. Since we have eight pairs of modelled and observed points for

each diagram, we normalized the data by dividing each of the standard deviations and

the

RMSEs by the standard deviation of the observations associated with each point, so that all

the observation points fall at 1 on the x axis.

Section 4.5: I think these results would be stronger if readers could see a bit more than just r2 and

RMSE. Changes in modeled values between runs, or changes in mean bias, would be useful

additions to this section.

We have added plots of cumulative NEP to the analysis, as shown above.

Figure 1: There is no key for a lot of the notation in this figure. The soil layers are also a bit

confusing. The model seems to calculate peat depth prognostically, but this diagram implies that

there are fixed depths for fibric, hemic, and sapric layers. That doesn't appear to be the case in

the actual model equations.

We have added the following sentences to the figure caption: "The symbols C, T and θ

represent carbon, temperature and soil water content respectively. The subscripts L,

S, R, H, and D represent leaf, stem, root, fresh litter and old litter respectively." As

explained in section 2.1 of the manuscript, the first nine peat layers have thicknesses

of 10 cm and the tenth, whose thickness can vary with time, contains the remainder of the peat depth.

Figure 5: What is going on with the "observed" fluxes in UK-Amo QE in 2006? Those do not

look like trustworthy measurements, and if they were included in the evaluation it casts doubt on

whether the resulting statistics are meaningful. It's probably worth checking with the PI if there

was some equipment problem in that year. Also, why are only 4 of the 8 sites shown?

The large scatter in the observed UK-Amo QE flux in 2006 is mentioned in the second

paragraph of section 4.2 as being probably due at least partly to instrumental errors. We

chose to show two representative fens and two representative bogs in order to allow the

figures to be legible and the presentation focused.

Figure 7-9: Why these six sites and not all 8?

For the same reason as above. For full comparisons, all the sites are included in figures 10

to 12 as well as in tables 5 to 7.

Review 2 (Dr. Nigel Roulet)

In this paper the authors combine a newer version of CLASS with a newly developed peatland

carbon model based on the structure CTEM. This paper is an interesting and useful addition to

the literature but I believe the authors need to do more to substantiate the models usefulness.

Their work represents part of a movement by some global modelling groups to incorporate

peatlands into the global models. The reason this is important is because peatlands represent the

highest carbon density ecosystems in the world so while they cover a relatively small fraction of

the land (4 to 6%) they contain up to 25% of the world's terrestrial biogenic carbon. Simulating

the sensitivity of their carbon stores to climate and land-use change is important for future

projections of global carbon cycling. Since the carbon dynamics of peatlands is so tightly

coupled to surface hydrology it is reasonable to ask if climate change will have a significant

impact on their carbon function in the future. The only way to address this question is through

modelling and that modelling requires a reasonable representation of peatland ecosystems at an

appropriate level for incorporation in global ecosystem models. The authors present the details

on the development of such a model and then provide an evaluation of their model output for six different peatlands: three bogs and three fens.

I think the authors have done a reasonable job but I do have some constructive criticisms of their

manuscripts as it current stands. My main concern is the authors provide little in explanation for

why their models produces the results it did. The evaluation of the model against measurements

is useful but it is very limited. They compare model out against observations but do not go into

detail of why the model is successful in some cases and not very successful in others. They

presentation of their results is very limited and not overly useful at time. Better visual

presentation of their results (see suggestions below) and the inclusion of some sensitivity

analyses would help demonstrate to the community the utility of the model they have developed.

Several things stand out as being quite unusual that I think the modelling community, and

certainly the peatland carbon modelling community will find unusual – for example the apparent

lack of the influence of initial conditions and the apparent lack of sensitivity to peatland wetness

on the carbon exchanges. These are quite at odds with the empirical observations from numerous

long-term measurement sites (some of these sites are included in this manuscript). The

simulation of wtd seems quite poor but it does not seem to really matter in the end? The authors

should do some more analysis to better assess where the uncertainties in the simulated results are

coming from (see suggestions below). All models are far from perfect but hopefully they are

useful. To determine the usefulness readers have to understand why the model does what it does

and why it does not do what it was expected to do.

We thank Dr. Roulet for his comments and his constructive criticism, and we endeavour to

address his concerns below.

Page 1 Ln 36 This depends on the peatland type. Mineral peatlands with Ca concentration much

above 2 mg/l there becomes less bryophytes and more sedges. Roughly this is the difference

between bogs and fens.

We have re-worded the sentence to read "bryophytes or sedges".

Page 2 Ln 7-10 Yes but Wu and Roulet showed bogs are quite resilient and fens are not.

Christensen et al and Ise et al. worked on poor fens. The conclusion from the literature you cite is that ombrogenic peatlands (bogs) may have sufficient resilience to maintain their sink function in

climate change but fens, which rely on additional external inputs of water, may not. This is the

crux of the problem in simulating the sensitivity of peatlands to climate change. We break forests

into different functional types. Similarly peatlands should not be seen as one type of ecosystem.

My guess is that that C store in peatlands is roughly 50 - 60% in bogs and 40 - 50%

fens.

As noted above in our response to the first reviewer, we have added some sentences in the

introduction outlining the differences between fens and bogs. We have also included the

following sentence: "Wu and Roulet (2014) showed that fens, which rely on external inputs

of water, may be particularly sensitive to changes in surface hydrology."

Page 2 Ln 17 I believe was should be were

We believe that it should be "was", referring back to "poor representation".

Page 2 Ln 28-34 Unlike the other three models discussed in the sentence HPM it us not process

based but it a phenomenological model. It is a one year time scale model and is not of the same

temporal scale as the process based models.

True. We have deleted the reference to HPM.

Page 6 Ln 23 It may be a more theoretically sound representation but "better" needs to be

justified by evidence.

We have changed "better" to "more detailed".

Page 9 Ln 4- 16 What did you do for the spin up of the peat profiles of each peatland? You

discuss the sensitivity of two of the initial parameters derived from the spin up at the end of the

results but this is an important issue to discuss up front. Later on you show that it might not

matter that the decomposition rates with depth are generalized and the presence or absence of

moss is not that important (not sure I understand this) but the hydraulic properties of the profile

are important to the simulation of the water table (wtd). Later on you show that you have

marginal success in simulated the wtds but it might not matter that much for GEP, ER, or NEP.

However, if this model will be extended to do methane this will be critical. At the very least the

authors should be explicit here on the decomposition coefficients or base respiration they assign to each layer to capture the drop in intrinsic OM quality. They also should show explicitly the hydraulic parameters for each layer so the readers can determine if the characteristic differences between fen peat and big peat are in the model parameter set or are not in the model parameter set. The authors do not address this issues until the end of the paper and it should be clear from the beginning.

Spin-up is discussed in section 3.2. The purpose of the spin-up was only to determine the C content in the vegetation pools; we have reworded the second paragraph to make this more clear. The decomposition rate coefficients and the moss depth were assigned for fens and bogs based on values in the literature, not on spin-up. The soil hydraulic parameters and decomposition parameters are listed in Tables 1 and 3.

Page 9 Ln 15-16 Does this assumes the relationship between C and density is the same across all peatlands. Is this true? The density - depth relationship and depth - age relationship can be quite different among individual peatlands and quite different between bogs and fens. The model

appears to have an ombrotrophic bog set up? Is this used for all the simulations?

Yes, we used one C-density relation for all the peatlands. It is true that density-depth-age relationships may differ substantially among peatlands, but this information is not available on a global scale, which is where our model will be applied.

Page 10 Ln 1 – 14 You use Taylor plots in your evaluation but do not mention this here?

We did not think it necessary, as Taylor diagrams include the r2 and RMSE.

Page 10 Ln 16 A general comment on the presentation of the times series in the result sections.

Time series are useful but they are difficult to sue to isolate if the uncertainty in the model result is random (ok assuming they cancel) versus systematic (which may or may not be OK). Scatter plots of simulated versus observed around a 1:1 line would reveal if there are some systematic errors - for wtd and LE there appears to be some systematic errors in the growing season. For NEE, GEP and respiration, it appears that for some peatlands the top and bottom 10% are systematically missed and in other sites both the GEP and ER are grossly under-estimated but because NEP is the difference in these two numbers the NEP does not look that bad. Using these plots does not necessarily negate the utility of the model but it helps the reader assess if the model is suitable or not for the task it is being developed for – the assessment of peatland carbon dynamics in changing climate conditions. For example St. Hilaire et al. (2010) showed that MWM truncated high GEPs and ERs but these represented less than 10% of the total exchanges so it did not matter for the overall annual exchange. Where this truncation may have implications for climate simulations is if there is a systematic shift into those conditions that are more favourable to greater GEP and/or ER. St, Hilaire et al explicitly show this so the reader is fully aware of the issues. In these models it is the accumulation of small systematic uncertainties that can give erroneous results over the long-term.

We have replaced Figures 7, 8 and 9 (GPP, ER and NEP) with scatter plots, and modified the discussion accordingly.

Page 10 Ln 18 On Figure 3 the value of 0.25 cm should be 0.25 m. This result is interesting. The authors are not aware but there are two papers (one in press and one in review at Ecohydrology)

that show that the wtd changes over time across the hummock - hollow at MB move in unison. This means if you know a wtd at a single point and you know where that point is with respects to the difference in hummock – hollow elevations then you can estimate of the wtds across MB. So the authors' explanation is plausible though I am surprised the model offset is the same height of the height of the hummocks above the hollows. This raises the question why the same problem did not arise for Fajemyran. It has micro-topography, maybe even greater, than MB?

This is an interesting point. We are not certain, but perhaps the water table was measured in a hollow at Fajemyran? Our thanks for pointing out the error in the Figure 3 caption, which we have corrected.

Page 10 Ln 30 – 35 The Wtds for the fens is poor in all three sites. They are generally between 0.1 to 0.3 m off. Part of this maybe the parameter set up is for bogs not fens. In terms of carbon function this has a much larger implication for fens than bogs - see Wu and Roulet 2014. Did you try any of the fen simulations by adding in extra water emulate additional water through groundwater seepage? We know fens receive some addition water from either surface inflows

and/or groundwater seepage. What you do not know is how much extra water. My gut feeling,

from visiting a few of these sites, is the extra water is quite small for AB, maybe about 5 to 10%

for Degero, and probably more for Lom. You compare in Fig. 5 measured and simulated Et and

it looks like Et is overestimated at most sites, hence I assume this is the reason for the problems

with the wtds? It's important to a better handle on where the problems are with the wtd estimates

because wtd is a critical variable to the NECB. It's even more important if you intend to

eventually use this model to get at methane.

These are good points. We have not attempted to introduce groundwater seepage because

on global climate model grid scales, unless a tiling approach is implemented this is

impossible to parametrize. We do have a student working on groundwater modelling

within CLASS, but this is only anticipated as being applicable at regional climate model

scales. When our peatland model is implemented with CLASS in the Canadian Regional

Climate Model we will definitely be interested in revisiting this groundwater question.

Page 11 Ln 4-6 As suggested in the general comment above scatter plots will reveal if there are

consistent biases in the simulated turbulent fluxes.

Page 11 Ln 29 Again scatters plots of GEP and ER around a 1:1 line and this will illustrate the

biases in the model relative to the size of the flux. It also illustrates the range of GEP and ER the

model does well and where it does not do well. This is important because NEP is the difference

of two much bigger numbers and it may do NEP reasonable well but for the wrong reasons. An

alternative way to illustrate the uncertainty is to analysis the residuals of the regressions between

observed and simulated and see if there are patterns. If the errors are random there should not be

any pattern to the residuals but if the residuals show a pattern this suggests structural issues with

the model.

As noted above, we have changed Figures 7-9 to scatterplots.

Page 11 Ln 33-36 It is interesting that the model does well on Kaa given that an appa mire - i.e.

it contains a lot of open water in the form of pools in the measurement footprint . Pools tend to

be large sources of CO_2 with no mechanism for the uptake of CO_2 (see work by Hamilton et al.

1994 and recent work by Pelletier et al (2014). This suggests that the model gets the 'right'

answer without accounting for the spatial variability. This is a little disconcerting.

We would argue that this does not necessarily imply that the model is getting the right

answer for the wrong reasons. It is possible that at Kaa the impact of the open water pools

is relatively limited.

Page 11 35-38 I also find this result reason for concern. In one case, RU, there is a huge mass of

old carbon that sustains a larger than simulated ER and in the other case, UK, there is a relatively

tiny mass of C that produces the same over-estimate. This does not really make sense to me

unless the respiration below 1 m depth is insignificant.

This is indeed the case. Generally respiration decreases with depth in soils, which is one of

the factors that allows the buildup of peat in these ecosystems. Changes to heterotrophic

respiration with depth in soils includes changes in oxygen transport (e.g. diminishing

vascular tissues) and bulk oxygen availability, soil microbial community changes,Âă mineral

sorption, and more broadly temperature and moisture changes. (Note that the actual peat

depths are UK-Amo with 10 m of peat while RU-Fyo has around 1 m.)

In both cases the errors in ER are offset by a grossly overestimated GEP. I can understand why

there is little difference in GEP if the conditions are general the same at the peat surface. This

section needs some more thinking – how much of the ER comes from autotrophic respiration?

We have included new plots, reproduced below, that display GPP, AR, and HR for

comparison of their relative magnitudes.

Interactive comment on Geosci. Model Dev. Discuss., 8, 10089, 2015.