Comment on bg-2021-316
Anonymous Referee #1

Referee comment on "Evaluating alternative ebullition models for predicting peatland methane emission and its pathways via data–model fusion" by Shuang Ma et al., Biogeosciences Discuss., https://doi.org/10.5194/bg-2021-316-RC1, 2022

This study compared two algorithms of peatland methane ebullitive fluxes and examined the impact of data availability on simulation reliability. The EBG algorithm was found to be more accurate than the ECT in simulating both methane fluxes and pore water methane concentrations. Using both methane fluxes and pore methane concentrations in parameter calibration led to better constrained parameter values. Overall, the paper is well-written, and the experiments are solid. The conclusions can be potentially helpful for methane modelers. However, there remain doubts in the motivation and aim of the study, and more information and clarifications are needed regarding the methods and the data analysis.

Major comments

How available are pore water methane concentration measurements in wetland sites generally? More observations would always help to better constrain parameter values and obtain more accurate simulation results, but if such variables are quite rare or hard to be measured, how will the finding inform other modelers on how to improve their model performances?

L31: This sentence could be misleading since the paper didn’t actually investigate the relative contributions of each pathway but only presented the modeled results which were not validated against measurements.

Table 1: Why did you select growing season only? How long did each measurement last?
What does the ‘CH4 fluxes’ include here (from ebullition only)? Since ebullition is quite spatially heterogeneous, how many observation sites were deployed at the same time?

Eq 1: Why were fstp, fph, fred not calibrated? Did you do any sensitivity tests to determine what parameters to calibrate?

Method: More information about the modeling approach is needed. How was the model structured horizontally? For how long and with what time step were the models run? What was the spin-up period? How did you deal with data with different frequency and time span for model calibration?

Table 4: Is there a quantitative criterion of well and poorly-constrained parameters? Why do you choose mean+-sd instead of median and q1-q3? For Figure 2 e,f, the posterior distributions of EBG-F are skewed so mean+-sd may reflect a wide range but the values are still relatively concentrated. However, you claim the parameter values of EBG-F in these two figures to be poorly constrained, which doesn’t seem to be a fair judgement.

Figure 4: How does Tveg affect the uncertainty range of ECT?

L440: As mentioned in Introduction, the ECT is assumed to overestimate ebullition, which is the opposite of the finding here. This needs more explanation.

Minor comments

Figure 4: the legend for observations should be dots with ranges.