

Reply to RC1:

The manuscript presents a well-structured and scientifically sound study with valuable findings on the prediction of CO₂ emissions from peatlands under warming climate and rewetting scenarios. The modeling approach and discussion are generally well executed. I recommend the authors consider the following minor revisions to improve clarity and consistency:

- **Lines 63–66:** The sentence is difficult to follow due to the use of both WTD and groundwater level, which change in opposite directions. Please consider reformulating the sentence using only WTD.
- **Lines 70–71:** Correct the double negative in “not ... neither ... nor.”
- **Table S3:** Please define the variables Alpha, N, L, and Kint.
- **Line 201:** The term KGE_{WTD} is not defined—please add a definition.
- **Line 203:** To maintain consistency with other objective functions, consider using subscripts for q_{head_amp} .

We appreciate your suggestions and will revise the sentences to enhance clarity and define the terms more precisely.

- **Lines 215–217:** It is unclear how the objective functions are combined, given their differing ranges and optimal values. Please clarify the methodology.

We understand the confusion. Below we have rewritten section “Calibration method” (previously Lines 191-245) to ensure that methodology is clearly explained:

“Calibration method

We used the Pareto Archived Dynamically Dimensioned Search (PADDs) algorithm (Asadzadeh and Tolson, 2013) available within the optimization toolkit Ostrich (Matott, 2019). PADDs is a multi-objective optimizer and obtains the pareto front across multiple objective function groups, enabling post-weighting of individual objective functions. Throughout the calibration routine, Ostrich minimized the weighted sum of squared error (WSSE) of each of the objective function groups. The PADDs algorithm was run with the user settings of maximum 1000 iterations. The period 2010-2013 was used as a calibration spin-up period and the model performance was evaluated for the 2014-2023 calibration period.

Calibration was performed against three objective function groups: $KGE_{WTD_modified}$, $r_{spatial}$ and $KGE_qME_{head}ME_{amp}$. The $KGE_{WTD_modified}$ objective group is used to optimize the model performance with respect to the WTD in peatlands. KGE is the Kling-Gupta Efficiency and consists of three terms: the Pearson correlation coefficient r , a term representing the measure of variability α and a bias term β . In KGE, β is a unitless measure of the bias specified as the ratio between the sum of simulated and observed values ($\beta = \sum sim / \sum obs$). As we use KGE to optimize the WTD (and not hydraulic head), the operational sign can be both negative (water table above surface/inundation) and positive (water table below surface), violating the idea of optimizing β as the ratio of sums of values with possibly alternating operational signs. Therefore, we are using $KGE_{WTD_modified}$ where β is replaced by the mean error (ME) (Table 1). This modification requires that the order of magnitude of the ME_{WTD} is comparable to the errors on the other terms in KGE. In our case this is ensured by the fact

that the mean observed WTD values range between approximately 0.3-0.6 m, resulting in ME_{WTD} values typically below 0.5 m. Alternatively, the ME_{WTD} term could be scaled within the KGE_{WTD} equation.

The calibration using the $KGE_{WTD_modified}$ as objective function group aims at achieving the best overall agreement between simulated and observed WTD. However, during first calibration experiments, we found that this objective function group primarily focuses on the temporal dynamics of WTD. To improve the representation of the spatial variability of the mean WTD, the correlation coefficient ($r_{spatial}$) was included as an additional objective function group (Table 1).

$KGE_qME_{head}ME_{amp}$ is an objective function group that combines three performance criteria: the Kling-Gupta Efficiency performance criterion for discharge (KGE_q), the mean error of hydraulic head in deeper aquifers (ME_{head}) and the mean error of annual amplitude of hydraulic head in the deeper aquifers (ME_{amp}). For a detailed description of the implementation of ME_{amp} as objective function see (Henriksen et al., 2020). This objective function group was included to optimize the overall water balance and streamflow dynamics expressed through the discharge at the catchment outlet (KGE_q), to match the general water level in the deeper aquifers across the catchment (ME_{head}), and to match the natural seasonal variations in hydraulic head (ME_{amp}). As the metrics of KGE_q , ME_{head} and ME_{amp} are combined into one objective function group, we need to weigh the observations, to ensure that KGE_q , ME_{head} and ME_{amp} affect the objective group of $KGE_qME_{head}ME_{amp}$ approximately equally. This was done based on WSSE from a model run with initial parameter values.

Table 1: Objective functions metrics. KGE stands for Kling-Gupta Efficiency.

Objective function group	Observations	No. of observation wells	Metric	Abbreviation	Equation	Range	Optimum value
$KGE_{WTD_modified}$	Daily WTD in shallow wells (in peat)	22	Modified KGE on WTD	$KGE_{WTD_modified}$	$1 - \sqrt{(r_{WTD} - 1)^2 + (\alpha_{WTD} - 1)^2 + (ME_{WTD})^2}$ Where, $ME_{WTD} = \frac{1}{n} \sum_{i=1}^n WTD_{sim_i} - WTD_{obs_i}$	$[-\infty; 1]$	1
$r_{spatial}$	Mean WTD over the calibration period	22	Spatial correlation of the mean WTD	$r_{spatial}$	$r(WTD_{sim}, WTD_{obs})$	$[-1; 1]$	1
$KGE_qME_{head}ME_{amp}$	Discharge	1	KGE on discharge	KGE_q	$1 - \sqrt{(r_q - 1)^2 + (\alpha_q - 1)^2 + (\beta_q - 1)^2}$	$[-\infty; 1]$	1
	Hydraulic head in deep wells (in mineral soil)	66	Mean error on hydraulic heads	ME_{head}	$\frac{1}{n} \sum_{i=1}^n head_{sim_i} - head_{obs_i}$	$[-\infty; \infty]$	0
		8	Mean error on yearly amplitude of hydraulic heads	ME_{amp}	$\frac{1}{n} \sum_{i=1}^n A_{sim_i} - A_{obs_i}$	$[-\infty; \infty]$	0

WTD: water table depth [m], q: discharge [m/s], head: hydraulic head [m], A: amplitude [m]

A local sensitivity analysis based on initial parameter values from Table S4 was performed and values of composite scaled sensitivity (CSS) were obtained. Selection of free calibration parameters were based on the criterion that parameters were included if their CSS was larger than 0.05*CSS of the parameter with the highest CSS. The resulting 11 free parameters are indicated with grey in Table S4. Other parameters were kept at the values listed in Table S4 or tied to the calibration parameters.

“

- **Tables S3 and S4:** Both tables present water retention parameters for peat, but with differing values. Could you explain the reason for this discrepancy?

We apologize for this discrepancy. Some values originate from an earlier model calibration, and the inconsistent values will be corrected.

- **Line 287:** Remove the unnecessary line break.

Certainly, we will.

- **Line 433:** The temperature sensitivity of soil fCO₂ is commonly reported as non-linear. While the selected model and alternatives are well addressed in the discussion, I’m curious—did you explore temperature sensitivity within WTD bins? The model appears to miss many of the observed high fCO₂ values; perhaps a non-linear temperature dependence could improve performance?

I am unsure whether you meant exploring temperature sensitivity within **WTD bins** or within **temperature bins**. In the paper we explored the WTD sensitivity to fCO₂ within temperature bins. Figure R1 illustrates the relationship between fCO₂ and temperature within temperature bins (as defined in the article), and Figure R2 presents the relationship between fCO₂ and temperature within WTD bins. I agree that the observed data clearly indicates a non-linear relationship between fCO₂ and temperature.

However, the relationship between fCO₂ and temperature in our developed WTD-Tair model is also non-linear. In Figure R3, I have plotted the observed fCO₂, the fCO₂ simulated using the WTD-Tair model, and the fCO₂ simulated using the Lloyd-Taylor model (Rigney et al., 2018, Eq. 5) against temperature. The results clearly show that the WTD-Tair model follows the Lloyd-Taylor model but provides a better match with the observed fCO₂, as it also accounts for WTD sensitivity ($r_{\text{WTD-Tair model}} = 0.78$, $r_{\text{Lloyd-Taylor model}} = 0.63$).

The very high observed values, e.g. in Figure R3, will not be possible to capture with a simple model based solely on Tair and WTD, since we also observe much lower fCO₂ rates on other days with similar Tair and WTD (see Figure S5), so the model becomes a compromise that captures some variability while preserving a sound mean.

I agree that it would be valuable to investigate whether the non-linear relationship between fCO₂ and temperature within bins could be incorporated into our emission model to improve its performance especially for high values of fCO₂. However, I also believe that our developed WTD-Tair model is simple (which is a strength in itself) and performs well, even though it underestimates the highest fCO₂ values. Therefore, we consider the WTD-Tair model sufficient for the purpose of this study.

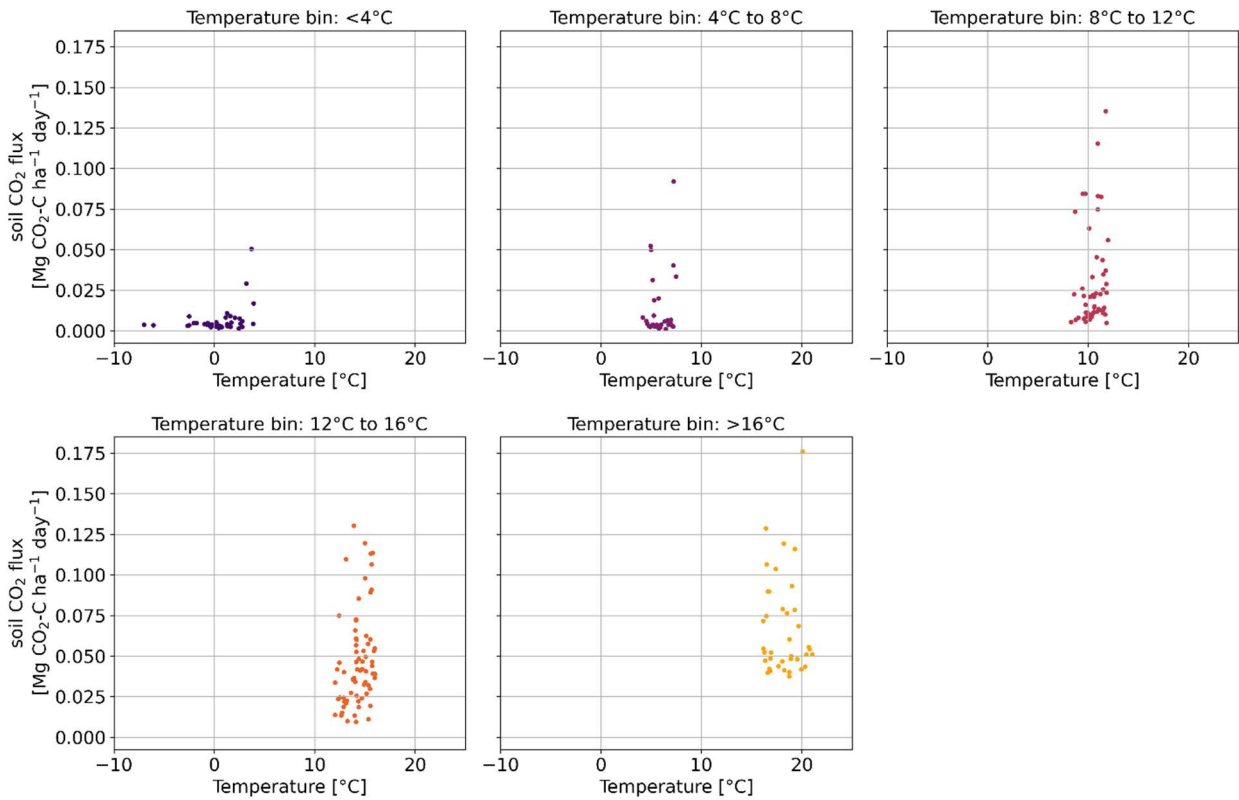


Figure 1: temperature sensitivity within temperature bins

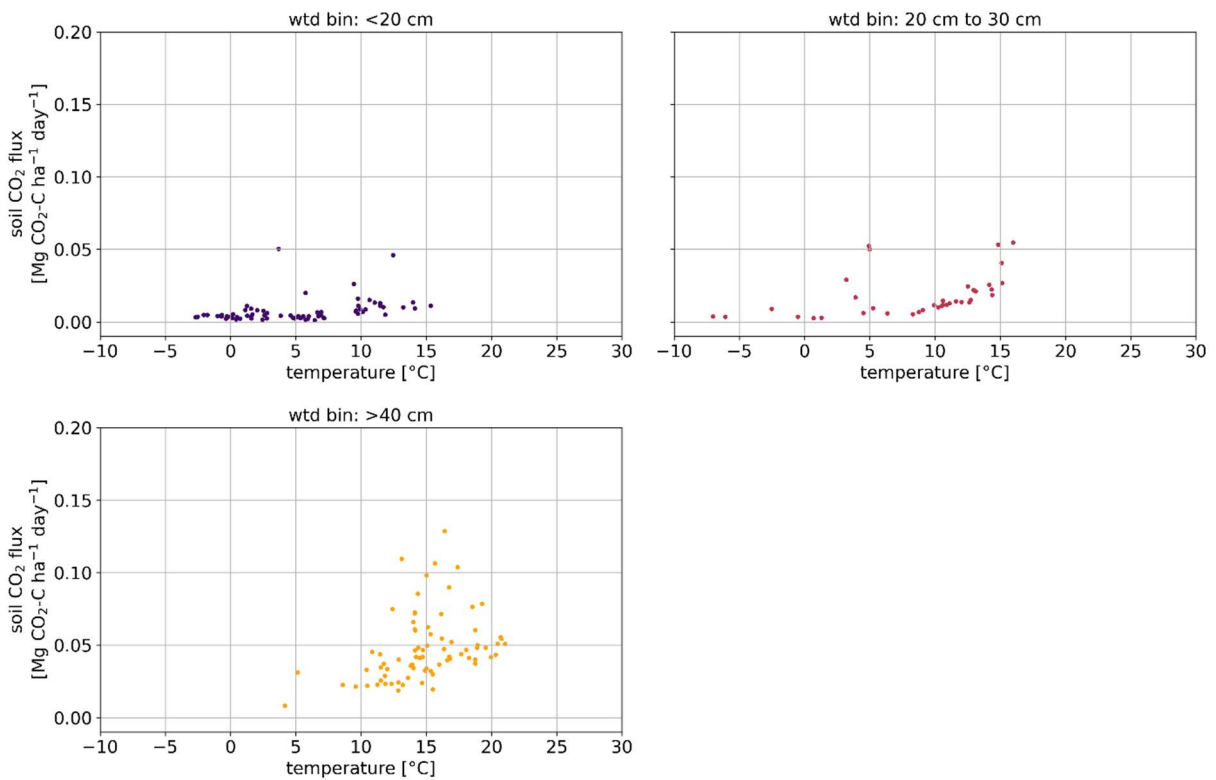


Figure 2: temperature sensitivity within WTD bins

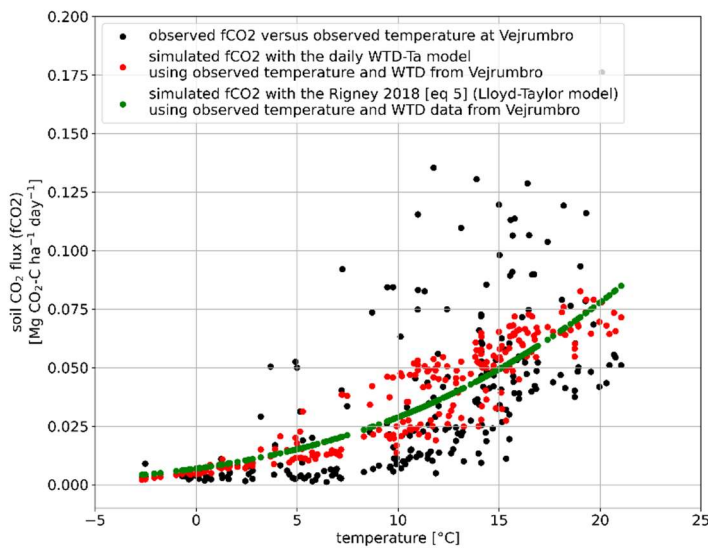


Figure 3: observed and simulated $f\text{CO}_2$ versus observed temperature for the observed $f\text{CO}_2$, $f\text{CO}_2$ simulated with the WTD-Ta model and $f\text{CO}_2$ simulated with the Lloyd-Taylor mode.

- **Lines 479–487:** While I understand the need to limit $f\text{CO}_2$, the rationale behind the chosen thresholds (WTD 62.5 cm and Tair 25°C) is unclear. I expect these values to strongly influence the comparison of modeled fluxes for extreme years (Lines 516–517). Please elaborate.

In the Daily WTD–Tair model, fluxes are sensitive to the WTD threshold but less affected by the temperature threshold. As you also note, we needed to constrain $f\text{CO}_2$ in the Daily WTD–Tair model. Our rationale for the selected threshold is explained in lines 479–487, where we aimed to choose a reasonable value based on data and comparability to the sensitivity range of the annual model. Currently, we lack a more accurate approach to define the WTD threshold because daily $f\text{CO}_2$ data for WTD values deeper than 47 cm are unavailable (line 475).

Below we have tried to update the reasoning behind the selected thresholds for WTD and Tair (previously Lines 479–487):

“The Vejrumbro dataset used for fitting the *Daily WTD-Tair model* was limited to a maximum WTD of 47 cm and maximum Tair of 21°C (Figure S7). Outside this range, the predictions of the *Daily WTD-Tair model* exhibits increased uncertainty. At the same time, it is generally understood that the upper portion of the peat layer drives the net CO_2 emissions observed at the surface. Therefore, the extrapolation of WTD in the *Daily WTD-Tair model* must be constrained. The *Daily WTD-Tair model* should be sensitive within a WTD range comparable to the expected daily variation in the Annual WTD model, which also reaches an $f\text{CO}_2$ asymptotic at deeper water tables. In the *Annual WTD model*, the Annual NECB reaches 90% of its maximum asymptotic level at a mean annual WTD of 30 cm (Figure 5). The mean annual WTD results from intra-annual (within year) WTD variation described by the annual amplitude. The mean annual amplitude (based on monthly means) is 65 cm, across the 22 observed WTD time series in the Tuse Stream catchment used for calibrating the hydrological model. We assume that a mean annual WTD of 30 cm originates from an annual WTD variation with a similar amplitude. Therefore, we assume that the WTD range of the *Daily*

*WTD-Tair model is $30 + 65/2 \text{ cm} = 62.5 \text{ cm}$. For the Tair range, it is assumed that the sensitivity continues until 25°C , which is a daily average value very rarely occurring, even in future climate projections. Thus, when applying the *Daily WTD-Tair model*, daily WTD values and Tair values were truncated, setting WTD and Tair to 62.5 cm and 25°C , respectively, when exceeding those thresholds.”*

We carried out a sensitivity analysis to evaluate the impact of the WTD threshold (Figure R4). The impact of the WTD threshold is clear, and as expected the WTD threshold has a greater influence on the extreme year than the average years. On average the emission estimate changes by app. 10% for average years and 14% for the extreme dry year as a function of WTD threshold with a reasonable range of thresholds. It should be noted that these differences would apply to all scenarios for climate and rewetting, e.g., in Figure 7 and 9 and therefore would not change the conclusions regarding impact of model selections, only the total emission estimates. The sensitivity to the Tair threshold is much smaller, mainly because daily average temperatures above 20°C very rarely occur.

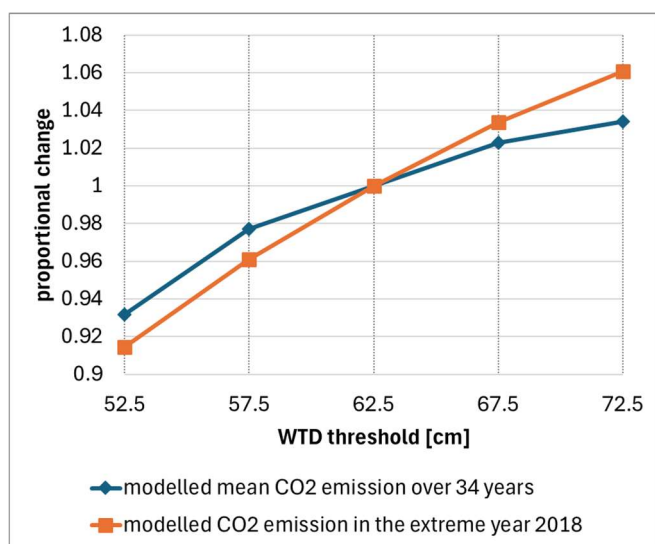


Figure R4: Proportional change in mean CO_2 emission from Daily WTD-Tair model aggregated to annual for different WTD threshold values. 62.5 cm is the threshold value in the preprint. Blue: modelled mean CO_2 emissions across 34 years. Orange: modelled CO_2 emissions for the extreme year 2018.

- **Lines 488–491:** The statement that “both CO_2 flux models exclusively account for the CO_2 emissions from the peat soil” is misleading. The Annual model includes NEBC, and thus GPP, as discussed earlier in the manuscript. GPP is also used as a reason for downscaling the daily model for fCO_2 . Please revise this statement for accuracy.

You are right, we appreciate your input and will adjust the sentence as recommended.

- **Figure 6:** Consider combining panels c and d, as they contain overlapping information.

We agree, there is overlapping information. However, the repetition of the daily model aggregated to annual values serves as reference in both plots, both to the variations in Tair and WTD and to the alternative Annual model. In order to not make panel c too “busy” by adding the annual model, we prefer to keep both panels.

Discussion: The section is generally well structured. However, it would benefit from additional citations, e.g., regarding the advantages of hydrological models and process-based emission models.

In the revised manuscript, we will include a more profound discussion on the use of hydrological and emission models, along with relevant citations.