The study by Li et al. entitled "Global natural wetland methane emissions (2000-2025)" uses various bottom-up and top-down estimates of methane emissions from wetlands for the period 2000-2020 in conjunction with a machine-learning approach to model global emissions for the period 2000-2025 and simultaneously project global methane emissions from wetlands for the period 2021-2025. Based on these newer estimates, wetland emissions and growth rates are presented and discussed in relation to latitude, regions and season. Their findings suggest that methane emissions from wetlands in the northern mid-to-high latitudes have increased in recent years. Hence, this machine-learning approach provides the opportunity to model more recent developments of wetland methane emissions, able to investigate possible feedback mechanism and emission changes earlier than global data networks can usually achieve.

This manuscript is a valuable contribution to the research community, and I recommend it for publication in ESSD once the authors have addressed the comments and suggestions raised in this review.

We sincerely thank you for the positive assessment of our manuscript and for recognizing the value of our machine learning framework and dataset. We appreciate your constructive feedback and have addressed each of the specific comments below.

General comments:

The authors should improve the resolution of all figures to at least 300 dpi. After improved resolution, Figures 1, 3, should be increased in size, so that small details can be seen.

Thank you. We had all the figures in 300 dpi resolution in the submission. It could be that the figure resolution was reduced when converted to PDF format or uploaded to the manuscript system. We will work with the editorial team to ensure all the figures meet the resolution requirement in the final publication.

Is it possible to write the definitions of all parameters including a "_" with small letters instead (e.g., CH4, pred)?

Thank you for the suggestion. We initially used the underscore notation (CH_4_pred) to maintain exact, searchable consistency between the manuscript text, figure legends and datasets. We will work with the journal's typesetting team to adjust the notation to meet its standard for final publication.

The gap between top-down and bottom-up approaches for wetland emissions was only recently narrowed by improvements of simulations, adding additional parameters, such as including inland freshwater systems, inundation, accounting for double counting, etc. (e.g., Saunois et al., 2025). The authors state that those are not included in the simulations (line 150). What error is expected regarding those missing emissions?

Thank you for this important comment. In the revised manuscript, we now clarify in the article title and the Methods that our target is the natural vegetated wetland category, which excludes lakes, rivers, reservoirs, coastal waters and managed sources. In addition, we also note in the Methods that the TD inversions used wetland priors from BU models that already exclude inland freshwater systems. So the key issue is not the prior definition itself, but the coarse spatial resolution of TD inversions. As noted in the GMB paper (Saunois et al., 2025), inland freshwater emissions may be implicitly included in TD estimates because wetlands and inland waters are not fully separable at coarse resolution. Because our framework is trained to reproduce the published GMB wetland estimates, it necessarily inherits this limitation. We now note this limitation in the Discussion Section 3.4 Limitation.

Key changes in the revised manuscript:

Article title now changed to: "Machine-learning-based estimates of global natural vegetated wetland methane emissions (2000-2025)"

Methods Section 2.1.1:

- First paragraph now revised to "We used monthly net CH4 flux estimates for natural vegetated wetlands from the most recent Global Methane Budget (GMB) synthesis for 2000-2020, including both process-based BU models and atmospheric TD inversions (Saunois et al., 2025; Zhang et al., 2025). In this study, the natural vegetated wetland definition excludes lakes, rivers, reservoirs, coastal waters, and managed sources, following the underlying GMB definition (Saunois et al., 2025). Detailed descriptions of each model and inversion framework are provided in the cited GMB publications; here we summarize the key elements relevant to this study."
- Added text to explain TD priors: "Wetland and inland freshwater emissions were prescribed as separate prior fluxes in the inversion framework. Wetland priors were derived from the ensemble mean of dynamic process-based BU wetland models used in the GMB estimates, whereas inland freshwater emissions were based on independent BU estimates. We used posterior flux estimates throughout this study."

New Section 3.4 Limitation in Discussion:

"A limitation concerns source attribution in the TD estimates used for training. Although the TD inversions adopted BU wetland priors that exclude inland freshwater systems, the coarse spatial resolution of TD inversions can still limit the separation of natural vegetated wetland emissions from nearby open-water or inland-freshwater emissions in mixed floodplain environments. This uncertainty cannot be quantified explicitly in the present framework because no independent source-resolved posterior diagnostic is available. Future progress is likely to come from higher-resolution atmospheric inverse frameworks. Two recent studies performed 25 km resolution atmospheric inversions using satellite observations in regional and global applications and improved spatial attribution (East et al., 2025; Hancock et al., 2025).

Another limitation is that our framework predicts wetland emissions from climate predictors only. ENSO-related impacts can be captured only indirectly, while dynamic inundation and atmospheric chemistry perturbations are outside the predictor space of the model."

The uncertainties of the emissions are reported with a CI value of 95%. Highest uncertainties for wetland CH_4 emissions correlate with highest emitting regions (e.g., Zhu et al., 2025). This is not reflected by the CI, as can be seen in Figure S1 (0-10%). Did the authors consider uncertainties from the spread of ensemble simulations? Did all BU and TD estimates provide global coverage, or is there a bias in spatial data coverage, influencing uncertainties? Remapping coarser products on finer resolution; have the authors just replicated the mean and uncertainties, or did they used a scaling factor (particularly important for the uncertainty)? The manuscript would benefit from more discussions regarding uncertainties and uncertainty propagation.

Thank you. We agree that the uncertainty propagation should be better clarified. We derived emission uncertainty from the spread across the 35 GMB ensemble members and 10 ERA5 ensemble members. The emulator skill uncertainty is evaluated differently as emission uncertainty, using R^2 and RMSE. We now explicitly distinguish these two uncertainty concepts in the Methods (new section 2.4). Regarding spatial coverage, all BU and TD estimates used in this study were provided at global gridded fields on 1 degree resolution from the GMB archive (except two estimates from CLASSIC and JSBACH, as stated in the manuscript, were remapped to 1 degree resolution). Therefore, we don't anticipate that the reported uncertainty is caused by unequal spatial coverage among models.

New Section 2.4 Uncertainty analysis reads:

"Emulator skill uncertainty was evaluated using held-out test period R^2, RMSE, NRMSE and prediction residuals relative to the GMB fields (Sections 2.2.3 and 3.1). Uncertainty in reconstructed wetland CH_4 emissions was estimated from the ensemble spread across all reconstructions generated from the 35 GMB estimates and 10 ERA5 ensemble members. This propagated two sources of uncertainty through the framework: (i) spread in the underlying GMB estimates and (ii) uncertainty in the predictors represented by the ERA5 ensemble. We use 95% confidence interval (CI) to represent this emission uncertainty."

Methane experienced a remarkable growth during 2020-2022, coinciding with an unusual La Nina event (triple-peak, Hasan et al., 2022, Nisbet et al., 2023), also supported by observations in the tropics (e.g., Ort et al., 2026; Balasus et al., 2026). Figure 5 shows a strong increase of GMB emission estimates towards 2020, mostly visible in tropical regions, which the XGBoost predictions cannot reproduce. Simultaneously, the Covid-19 pandemic is discussed to had a global influence on CH4 oxidation capacity though the decreased NOx emissions (e.g., Stevenson et al., 2021). However, as Covid-19 influences most likely are not considered by the XGBoost predictions, neither are unusual ENSO phenomena, aren't they? Sections 3.3, and 3.4 would benefit from more detailed discussions about limitations of the XGBoost predictions, and comparisons of the results with other studies, to better quantify the value of the results provided by this machine-learning approach.

Thank you for this insightful comment. We agree that the pronounced increase in GMB estimates towards 2020 is not fully reproduced by our emulator and deserves clearer discussion. We have now added discussion about the potential impacts of indirect capture of ENSO and changes in OH (such as COVID) could impact the model predictions. Specifically, we:

- added in Section 3.3: "We cautiously interpret this mismatch as the limitations of the emulator and uncertainty in the underlying GMB ensemble. The climate predictors used in this study can only indirectly capture ENSO (El Niño-Southern Oscillation) effects. Although the training period includes several ENSO events, the unusual persistence of the 2020-2023 triple-dip La Niña may still be underrepresented. In addition, most TD estimates from GMB use a similar prescribed OH (hydroxyl radical) distribution, possibly causing OH-related uncertainty to be underrepresented, particularly during anomalous periods such as COVID."
- added in Section 3.4: "Another limitation is that our framework predicts wetland emissions from climate predictors only. ENSO-related impacts can be captured only indirectly, while dynamic inundation and atmospheric chemistry perturbations are outside the predictor space of the model."

With this machine-learning approach, predictions from 2021-2025 were performed. Why did the authors decide for this time period only, and not included climatological predictions further into the future?

Thank you. We extended the record only through 2025 because our goal is to provide a low-latency continuation of the GMB estimates using routinely updated, observation-constrained predictor data. Extending future projections would need predictors from projected data and introduce additional scenario and model uncertainty that has not been evaluated here. We remain interested in predicting future wetland emissions as you suggested when we are able to solve these limitations and uncertainty issues.

Specific comments:

Title: The title should be reconsidered to match the content of the manuscript and the journal requirements more. The machine-learning approach should be somehow mentioned. What about “A machine-learning approach to estimate global wetland methane emissions to recent years (2000-2025)” or “Extending global wetland methane emissions to more recent years using a machine-learning approach (2000-2025)”?

Thank you. We agree that the machine-learning approach should be highlighted. Now we have changed the title to “Machine-learning-based estimates of global natural vegetated wetland methane emissions (2000-2025)”.

Abstract: I would remove the “z-scores” in the Abstract (too detailed).

Thank you, z-scores are removed from the Abstract.

Introduction: Wetlands for BU and TD approaches have the largest uncertainties of all methane emissions. I miss this important information in the introduction, alongside with suitable literature (e.g., Zhu et al., 2025).

Thank you, now the text reads “Wetlands are the largest natural source of methane (CH4), a potent greenhouse gas, yet they remain one of the most uncertain components of the global CH4 budget (Saunois et al., 2025; Zhu et al., 2025).”

Line 143: Please specify this statement: “sparsity of training data”. Do you mean from satellites, in-situ observations, both? Then, this is only true for some regions, e.g., the tropics and high-latitudes. Wouldn’t this be also an issue for the data used for this study, as it is based on in-situ and satellite data?

Thank you. The data we used in this study have global spatial coverage. Therefore, it shouldn’t be an issue for this work, but for other machine-learning approaches that use satellites or in-situ observations. We now clarify that the sparsity of training data is particularly an issue for the tropics and high latitudes.

Lines 140-158: What makes the machine-learning approach presented in this study superior, or how does it differ from other data-driven modeling approaches (e.g., Yuan et al., 2024)?

Thank you. Our intention is not to claim that our approach is superior to other models, but rather that it serves a different purpose (building a low-latency framework) and performs well.

Lines 190-194: The GMB dataset used is a combination of BU wetland biogeochemical models and TD inversion products. The TD inversions include observations from satellites and surface observations, which were then used to estimate posterior wetland emissions. However, the authors state here that they used anthropogenic emission inventories for each inversion. It is not clear to me if the authors therefore only include wetland emissions, or also emissions from anthropogenic or other natural sources. Please specify more, how the posterior flux estimates were achieved.

Thank you for pointing out this confusion. The anthropogenic emission inventories were only used as priors to represent anthropogenic emissions in TD inversions. We have now clarified that in the Methods.

Line 196: Do all BU and TD approaches provide global data coverage? If not, please discuss possible data biases.

All BU and TD estimates provide global data coverage.

Lines 251-253: What is the reason for this defined cut-off? Is it the highest Xth percentiles?

This cutoff was not defined from a percentile threshold, but was chosen pragmatically to exclude grid cells with persistently negligible annual mean fluxes and thereby reduce computational burden and noise in the emulator training.

Lines 263-269: What is the exact difference here between the two test data periods and the validation dataset? Please specify their roles more. And have you tested different choices of test and training data periods? Did those choices make a big difference on the outcome for the 2020-2025 period?

The training period (2003-2017) was used to fit the XGBoost models. Within this training period, the final 24 months (2016-2017) were reserved as a validation set for early stopping and hyperparameter selection only. The two test windows (2000-2002 and 2018-2020) were fully withheld from both fitting and tuning and were used only for out-of-sample evaluation. We chose two temporally separated test periods to assess model skill under climatic conditions both earlier and later than the training window. We did not perform a systematic sensitivity experiment over multiple alternative train/test splits in the present study.

Now the Methods section includes a clarification that reads:

“Each model-ensemble pair dataset was temporally split into a training set and two testing sets to assess reconstruction capability across different climatic periods at monthly scale. The training data covers January 2003 to December 2017. The testing data covers two periods: January 2000-December 2002 and January 2018-December 2020. Two test windows were used only for out-of-sample evaluation and were not seen during fitting or tuning. The final 24 months of the training period (January 2016 to December 2017) were reserved as a validation set for early stopping and hyperparameter selection. This splitting strategy is to test the model's ability to capture long-term climate variabilities in early/late periods and reduce overfitting.”

Line 295: Were latitudinal differences considered by calculating the grid-cell areas?

Yes, latitudinal differences were considered when converting fluxes to emissions.

Line 338: Is R^2=0.65+-0.003 here correct for the global monthly mean? Comparing this value with Figure S3a, the mean would lie more around 0.9, or are there strong outliners reducing the mean? Then, it might be better to mention this, and where they are geographically, or adding the median and standard deviation/percentiles as well, so that this skewness is clear. Or what is the difference in data shown between Fig. 1a and Fig. S3a? Same for the following R^2 mentioned in the text. Is the CI for the mean value here an averaged over all CI's? What about the spread across the means?

Thank you. Fig 1a presents grid-cell reconstruction skill. Monthly test-period R2 are first computed at the grid-cell level and then collapsed to one R2 per grid cell, after which the global mean and its 95%CI are calculated across grid cells (R2=0.65+-0.003). However, fig S3a shows monthly R2 for spatially aggregated latitude bands, where model predictions and GMB emissions are first pooled within each latitude band and month before R2 is calculated. Because this aggregation reduces grid-scale noise, the monthly R2 values in fig 3a are expected to be higher than the mean grid-cell R2 reported for fig 1a. We agree that this distinction was not clear enough in the original manuscript, now we clarify it explicitly in the revised manuscript and figure captions.

Lines 369-371: Please include here the global mean wetland CH4 emissions for the predicted and the GMB CH4 over the test periods. What is their estimated CI/uncertainty? Does the global difference of 2.27 Tg yr-1 exceed the associated uncertainties?

Thank you for this suggestion. We have revised the text to include the global mean wetland CH4 emissions during the test periods for both the predicted emissions and CH4_GMB, together with their 95% confidence intervals. During the test periods, the global mean predicted wetland CH4 emission is 158.48 ± 2.09 Tg CH4/year, compared with 160.75 ± 2.24 Tg CH4/year for CH4_GMB, corresponding to a difference of 2.27 Tg CH4/year (1.41%). This difference is small and does not exceed the combined uncertainty of the two estimates, so it should be interpreted as being within the estimated uncertainty range. We now revise the text to: "During the test periods, the global mean predicted wetland CH4 emission is 158.5 ± 2.1 Tg CH4/year, which is 2.3 Tg CH4/year lower than the global mean of CH4_GMB (160.7 ± 2.2 Tg CH4/year), corresponding to a 1.4% underestimation (Figure 2c). This difference is within the estimated uncertainty range of the two global means."

Line 392: Define "z-score" briefly and what this value represents. Or refer to literature defining the z-score.

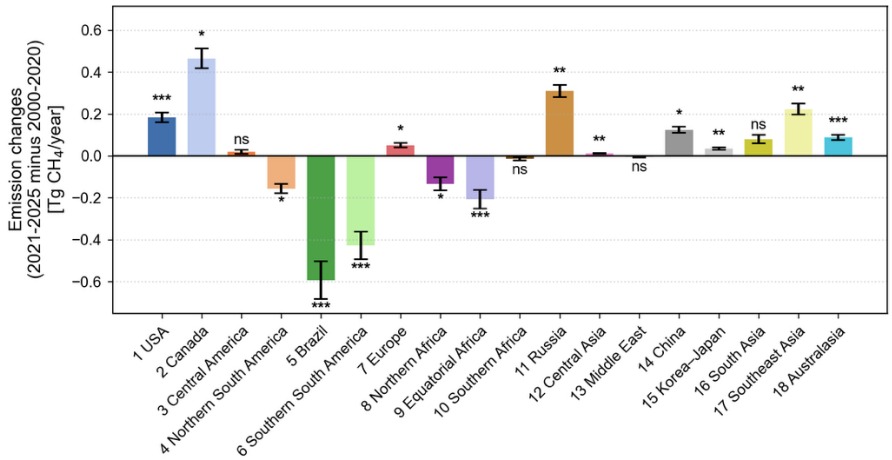
Thank you for the suggestion. We have now defined z-score in the manuscript.

Lines 414-420: Explain the definition of significance in the text as well.

Done.

Figure 4: Is there a reason, why the total emissions are plotted in logarithmic scales? It does not seem to be needed regarding their values ranging from mostly 0-30 Tg yr-1. Would it be possible to include the emission changes between 2021-2025 compared to 2000-2020, for all regions on, e.g., a second y-axis? As the authors are mainly referring to those differences in the text.

Thank you. The total emissions were plotted on log scale because some regions have relatively low emissions, and their emissions won't be clear to read if we apply the same scale for all regions. We agree with your suggestion that emission changes are the focus of the corresponding section, and they should be clearly indicated in the figure. We have now revised the figure (see below) to visualize the emission changes in recent years (2021-2025 vs 2000-2020) for each region. We believe the new figure provides a clearer and more direct comparison across regions.



Lines 440-452: The authors are referring to two La Nina events (within the training data) and one El Nino event (in the prediction) here. How certain are the authors that the XGBoost prediction can represent the variability of ENSO phenomena frequencies?

Thank you. Our intention was not to suggest that our model can independently reproduce ENSO event frequencies, but rather to assess whether it can capture methane emission responses associated with ENSO-related hydroclimatic anomalies when such conditions are reflected in the predictor variables. In this sense, the model's performance is conditional on the predictor-emission relationships learned from the training data. We agree that the limited number of ENSO events during training and test periods introduces uncertainty. We have added a limitation discussion on this point in Section 3.4 and Section 3.3 where we discuss the 2019-2020 mismatch.

Figure 5: Could linear regression lines be included here to verify the trends (slopes)? Furthermore, linear regression lines for the separated time periods (2000-2020; 2021-

2025) could show the difference in trend development for each period more visually, supporting the discussions in the text.

Thank you. We agree that the slopes should be presented for separate time periods. Figure 5 already contains multiple elements, thus adding regression lines would make the figure overly crowded. To address it, we revised Figure 6a to show the emission trends separately for three time periods, which provides a clearer visual comparison of trend development and better supports the related discussion.

Conclusion: At present, the conclusion is limited mainly to the results of the seasonal analysis of the methane growth rate, which, in my opinion, is not the only aspect of this work. The model structure and performance evaluation should also be briefly mentioned here to place the results in a broader context. Furthermore, it should be discussed whether there is room for improvement or whether this model is also suitable for other applications, apart from CH4 emissions from wetlands.

Thank you for this helpful suggestion. In the revised manuscript, we now expand the conclusion to briefly summarize the emulator's overall performance and the main limitations and opportunities for improvement. We also clarify that the framework is not limited conceptually to wetland CH4 emissions, but could potentially be adapted to other low-latency Earth system applications where a robust gridded target ensemble and routinely updated predictor fields are available.

Technical comments:

Please change throughout the whole manuscript all units Tg/year into Tg a-1, or Tg yr-1 (Tg CH4/year, accordingly).

Done, now all units are in Tg CH_4/year.

Line 82-85: Shorten sentence for easier reading flow. Exchange "and applying the framework to" with "applied on". Put " at monthly 1°x1° resolution" into next sentence (you anyhow repeat this information here).

Thanks. We have revised the sentences accordingly.

Line 86 – 88: Make a new sentence out of the content of what you write within the brackets (R^2 and RMSE).

Done.

Line 88-89: Change to "model estimates, including 22 process-based, and 13 atmospheric inversion estimates, paired with ..."

Done.

Line 90: Change “While” to “Our results show that”.

Done.

Line 92: Change “, this” to “. This”.

Done.

Line 96: I would suggest “The predicted emissions are able to capture …”.

Done.

Line 99: Don’t you mean 2000-2020 instead of 2000-2025?

Revised.

Line 101: Include “modeled” before “dataset”.

Done.

Line 115: Change “lags” into “a delay”.

Thank you for the suggestion. We have considered this change, but we prefer to retain “lags” because it more accurately describes the recurring multi-year time gaps associated with successive synthesis updates.

Line 122: This sentence has a lot of “ands”. Maybe “that regulate methane production, alongside with oxidation within the sediment and water column, to be eventually released to the atmosphere.

Done.

Line 185: Change to “The TD inversion products”.

Done.

Line 209: missing bracket after “index”

Done.

Line 209: Please define LAI precisely before using this abbreviation.

Done.

Line 223: Split into two sentences: “… transport. Their relevance …”

Done.

Line 231: “XGBoost” is used here the first time. Please define it properly. Or move this whole part to Section 2.2.2.

Thanks. This part is moved to Section 2.2.2.

Line 235: Correct “section” to “Section”.

We moved this part to Section 2.2.2 and therefore removed the text “see section 2.2.2 Model Architecture”.

Line 245: Please adjust Eq. 1 to a proper equation form (e.g., use mosin and mocos; use , etc.). Why aren’t you use just “m” for months, as it is more common in literature.

Thank you. We have adjusted the terms to m, m_{sin}, and m_{cos}.

Line 252: Change “of 1x20-15 kg CH4/m2/s“ to “⩾ 10 - 15 kg CH4 m-2 s-1”. And remove “or greater”.

Done.

Line 264 + 265: Change “cover” to “covers”.

Done.

Lines 286-292: Almost the exact same sentence is repeated twice.

Done, repeated sentence is removed.

Line 294: change unit to “kg CH4 m-2 s-1”.

Done, we changed it to kg $CH_4/m^2/s$ to be consistent with other units.

Line 295: “m^2”

Done.

Line 319: Include “them” after “compared”

Done.

Line 320: Please state earlier in the manuscript, that you defined the 18 geographic regions similar to Saunois et al., 2025. Best when you mention them the first time.

We now stated it in the earlier section 2.2.3.

Line 332: Don’t you mean Figures 5, S6, S8?

Thank you, you are right, we have corrected it in the updated manuscript.

Line 398: Remove “this study”.

Done.

Line 435: Make sure the caption of the Figure is on the same page than the Figure.

Thank you. We will work together with the journal’s editorial team to ensure the final publication meets all requirements.

Line 449: Aren't you referring to Figures S6, S7 here?

Revised.

Literature:

Zhu et al., 2025, DOI: 10.1088/1748-9326/adad02

Yuan et al., 2024, https://doi.org/10.1038/s41558-024-01933-3

Hasan et al., 2022, https://doi.org/10.3389/fclim.2022.1001174

Nisbet et al., 2023, https://doi.org/10.1029/2023GB007875

Stevenson et al., 2021, DOI: 10.5194/acp-2021-604

Ort et al., 2026, DOI: 10.22541/essoar.177248180.00166712/v1

Balasus et al., 2026, https://doi.org/10.5194/egusphere-2025-6251