BG review of Tikkasalo et al

Overall:
Scientific significance: excellent, scientific quality: excellent, Presentation quality: good
Overall this is an excellent paper. Many of my comments are requesting clarification or more details. Aside from these minor points, I have two major issues to point out:

The first about the lack of a spatially explicit flux modeling for CO2 as was done for CH4 and N2O. Much of the paper is justifiable building expectations for the impacts of spatial heterogeneity after clear cutting, and it is surprisingly absent in the results and discussion for CO2. In comparison to CH4 and N2O, I would expect CO2 to be easier to model given its strong relationships to variables already reported in the gap-filling discussion. The authors could take the GPP and respiration models used with gap-filling and apply the same spatial disaggregation technique as they did with CH4 and N2O.

The second issue is about the methane flux results. The flux estimates from the plant-covered ditch surface-type are extremely large, almost unbelievably so. These results need to be justified and put in context of other methane emissions. Given that the areal contributions of this surface type and therefore their weights within the footprint, are so small, it could be very difficult to have confidence in these results. In addition to comparisons to chamber fluxes or other studies, I would suggest investigating the robustness of the methane surface-type model with a simulation. Generate a flux for each surface-type based on your equations 3 and 4, calculate the theoretical EC observation after multiplying by the pixel footprint weight and summing, then add some reasonable random noise. Then apply your dissaggregation model and see if you can recover the original parameters you used to generate the fluxes. This is a straightforward way to test whether your dataset is under-determined or not. If you do not have enough variability in footprints weights from surface-types to recover your simulated fluxes, then you will have to reduce the complexity of surface-types or use a longer time series of data.

General comments:

- In figure 1 and throughout when color-coded landcover types are displayed: It is difficult to distinguish similar colors. The greens in particular all look the same. A more divergent color scheme would improve readability throughout the paper.

- Model predicative performance for the gap-filling ML models and the spatially explicit footprint flux models is evaluated and reported using R-squared. Whenever R-squared is reported, the slope and intercept of the regression should also be reported. R-squared describes the variance around the fit, but the slope and intercept describe model bias which is equally important. I also suggest providing the RMSE as a more useful metric than R2 because it is in comparable units.
- Section 2.8:
    - The methods described for surface-type modeling are the same as those used by Ludwig et al. 2024 from your introduction, and it should be cited here as well.
    - Can you please provide some justification for your choice of prior distributions.
    - Please describe your tests for convergence and their outcomes.
    - Please clarify that only non-gap-filled data were used in the surface-type modeling analysis
    - Why use LOO cross validation for the surface type modeling, when you already have withheld data in artificial gaps created for the gap-filling ML models?
- Figure 4 and 5: include slope and intercept on the fit depicted in panel c.
- Figure 6: The bold line for the central quartile is hard to distinguish, can you make it bigger?
- Table 3: I understand that the gap-filled budgets in the second and third column are agnostic to the area and make-up of the footprint. How are the surface type modeled fluxes summarized to comparable numbers to the gap-filled EC data, given that each observation has a different distribution and weight of surface types? The modeled fluxes can be weighted by footprints before summarizing to a budget, but due to gaps, there are timepoints without footprints. It would make more sense to me to use your surface-type models to calculate the budgets for the entire domain in your Figure 1, and then similarly apply the gap-filled time series of fluxes to the same area when summarizing, rather than reporting on a per area (ha-1) basis. By controlling the areal extent of this comparison it might also reveal interesting agreements or discrepancies between the surface-type model budgets and the footprint-agnostic gap-filled budgets.
- Section 4.1 first paragraph:

    - The spatial heterogeneity is generally put in context of similar ecosystems and other clear-cutting studies. But what is lacking is a quantitative comparison of the magnitude of these fluxes determined here (figure 7) to other studies. For example, is your exposed peat flux typical of peat ch4 fluxes? While I am not surprised by a slight uptake of methane in some surface types, it is surprising to see methane uptake in the ditch surface water. Similar features in polygonal tundra are large methane sources. The methane flux from plant covered ditches, the vast majority of all methane at this site, is alarmingly large, as in, it is similar to methane fluxes measured by eddy covariance at active landfills in warm climates. This result needs to be put in context of other fluxes and justified.
    - In table 2, you set up an investigation of scenarios to determine the level of complexity to use in the spatial disaggregation of fluxes. This is a great tool for supporting the robustness of your surface-type model results. You present results from the best model of the set described in the table. I would like to see more results on all scenarios. Specifically, how do the surface type flux estimates change in each version in table 2? In two of the five versions, your highest flux type is lumped with your lowest flux type, and discussing how the fluxes turn out in these scenarios would help provide confidence in the model results.
- Figures S7 and S8: Please include the panel letter designations in the legend as well for clarity.

Specific comments:

Line 88: Need space after period at the end of sentence
Line 107: Missing word. "[The] likely reason for this...
Line 247: missing space in citation for (Kljun et al 2015)
Line 565: Should cite Ludwig et al. 2024 here as well.
Line 581: Typo 'emissionsdd,'