Dear Editor, Dear Reviewers,

thank you for the opportunity to re-send our manuscript. We are thankful for the detailed comments to the previous version of the manuscript, that helped us to improve the manuscript for this new version.

The largest changes include moving some parts to the supplement and shortening the introduction and the discussion sections.

We attach the revised version of the new manuscript as well as point-to-point replies to the reviewer's comments.

With kind regards,
Tea Thum
(on behalf of all the co-authors)

Replies to the reviewer comments

The reviewer comments are in magenta, the replies in black and the manuscript text in italics.

The efforts made by the authors to address my (too) numerous comments and to improve their manuscript are greatly appreciated, as well as their detailed responses. I find that the paper much improved over the original version. However, I still believe that additional revisions are necessary for it to become more impactful.

We thank the reviewer for such a detailed new reading of the manuscript and addressing several issues that will help us to improve the quality of the manuscript.

Firstly, the paper is still overly long (31 pages without the references and Supplementary Materials). It would benefit from a clearer and more concise organization to better highlight its key developments and findings.

We have worked on condensing the text.

Secondly, certain aspects of the model parameterization and data processing remain unclear, despite their potential impact on the interpretation of the results (while the informations may be included in the paper, they are obscured by the paper's presentation).

We have now made clear in the current version that we did not do any specific parameterization for the model.

Structure of the paper

- Including "QUINCY terrestrial biosphere model" and "simulations" in the title would better define the study s focus and increase the visibility of the authors model.

According to the Biogeoscience journal's guidelines, the whole name of the model must be in the title, not only the acronym (because of this, we removed the model's name from two publications' titles last year). Therefore, using the model's name in the title would lead to a very long title, that would make it quite difficult to understand.

The point about simulations is good, and we modified the title now to:

"*Using different radiative transfer schemes for solar-induced chlorophyll fluorescence (SIF) in evergreen coniferous forests with a terrestrial biosphere model*"

so that the fact that this is a modelling study becomes clear.

- In several places, topically connected sections are disjointed, which makes the reading process more difficult. For instance:
- l137 and l150 related to GPP and gap-filling

We moved the sentence from L137 to the paragraph that described the eddy covariance measurements.

- l177 and l319: regarding the QUINCY PFT at the three sites

The first occurrence is connected to the general model presentation of QUINCY and the latter to the modelling protocol (i.e., describing how the simulations were done in this study). For our mind, this seems like a valid reason to have PFT mentioned at two places.

- l210, l240, l320 and l327: informations on soil optical properties

L210 belongs to the general model description of QUINCY and L240 to the model description of mSCOPE. This would seem sensible to be separated. L320 included an attempt to further clarify the beginning of the modelling protocol paragraph to help to understand what having one PFT means, but we removed that now. L327 specifically states which values for soil reflectance were used for this PFT. It would not make sense to have this within the general model description of QUINCY within the current organization of the manuscript.

- l216-220 and §2.5.3 on the use of the fesc parameterization to calculate SIF with one of the modelling approach

We separated the equation from section 2.5.3 based on the earlier round comment from the reviewer. We thought that it made sense, as the equation for fesc is very general for overall SIF modelling and often even presented in the introduction section of manuscripts. We now added a reference to the equation in section 2.5.3.

- in §3.3, l434: sentence related to Figure 7 when the analysis of the figure is provided later…

It is the next figure in line, and it is here used to explain some of the features in Fig. 6.

- The Introduction and Discussion sections can be improved and condensed.

The introduction was extended in the previous round because the reviewer wished for much more references to it. We have now worked on condensing the introduction. We also removed some points of minor relevance from the Discussion and condensed its text.

The description of the study objectives (from line 93) needs clarifications. QUINCY should be introduced in line 93 already; Given that the authors have mostly followed the approach of Raczka et al. (2019) to represent NPQs, is "how to accounting for it" truly a research question?

We modified the specific research question to include “*across different sites*”.

- Are all (numerous) figures and tables really necessary? Figure 1 could be removed (or moved to Supplementary Materials). Informations of some tables could be added in the corresponding figures or in the text. For instance the informations on computational

efficiency provided in Table 4 could be added directly in the text in §3.1.3.

The number of figures is high, because we show the same results for all the three sites. A model developer would find this relevant to see the generalizability of the results. Most of these figures are in the Supplement and therefore including them should not bring extra burden to the reader, who can read the manuscript by only concentrating on the CA-Obs results.

The reason we started the results section with figure 2, was to show the “end result” of the modelling to help the reader to see what the story will lead into. The radiative transfer simulations already include the NPQ-formulation that will not be introduced in the first part of the Results-section.

We have removed Table 4 according to the reviewer’s suggestion.

Clarifications

- Several definitions of the red and far-red spectral ranges co-exist, depending on the measurement / observation characteristics: 680-686 nm vs 687 nm for the red region (American sites vs Finish site), 745-758 nm vs 760 nm (in situ data) vs 740 nm (TROPOSIF) for the far red. It is still not clear whether these distinct spectral samplings are considered or not for the model simulations and comparisons with respect to the observation data.
All the more the presentation of the radiative transfer calculations in QUINCY (§2.4.1, §2.5.3) mentions two large spectral ranges (300-700 nm for the visible and 700-3000 nm for the near- infrared parts). What spectral ranges were considered to prescribe leaf and soil optical properties for simulating SIF in the two narrow spectral windows? Averaged values over 300-700 nm / 700-3000 nm? If so, does it have an impact on the consistency of the model–data comparison?

As we tried to convey in the earlier version of the manuscript, the visible and near infrared regions of the QUINCY model each have a specific value that is used for the whole range. Therefore, these values are not averaged for the specific regions, but it is one value per region. For the regions that had lower wavelength than 700 nm the visible region values were considered and for the regions that had higher wavelength than 700 nm the near infrared value was considered. This has now been clarified in the text in section 2.5.

- There is a misunderstanding regarding the wavelength at which TROPOSIF SIF data are provided with the fitting windows from which they are derived. The authors refer to the product user manual (p11 – description of the variables) to justify the use of "743 nm". The indication of 743 in the variable name is to distinguish the SIF estimates from the two fitting windows. Whatever the fitting window, SIF was scaled at 740 nm (see §3.1 of the same document and the "long_name" attributes of the variables p.12, as well as Guanter et al. (2019)). For the model-data comparison, simulations at 740 nm (instead of 743 nm) would be more relevant (although the impact is likely small).

Thank you for bringing this topic up again. Unfortunately there was some miscommunication between modellers and remote sensing experts among the authors and this information had

not been clear. Indeed, as the reviewer is stating, the impact will be small, but we re-made the figures and related calculations with this correct value.

As I understood, the authors consider different temporal samplings when comparing QUINCY SIF simulations with TROPOSIF data: mid-day values for the simulations and daily averages for TROPOSIF. If so, these different sampling characteristics likely contributes to the observed mismatch in the amplitude of SIF time series between model simulations and observations (with higher SIF values around noon for the simulations). For a more coherent comparison, the authors should consider 1) daily averages both for simulations and observations, and 2) the "daily corrected" SIF estimates from TROPOSIF (SIF_Corr_743 variable) to limit the impact of uneven temporal distribution of observations within a given day.

The reviewer raises valid points and we have added a point about this to the discussion. The reason why we wanted to use these “midday” values here, even though they can be criticized, is that we were interested in the seasonal cycle and for these sites, the influence of day length would force a seasonal cycle that has nothing to do with the seasonal development of the vegetation. We wanted the logic behind these values to follow the logic of the other analysis shown in this study. Our simulation approach takes the latitude and the incoming radiation into account, and this is why we considered it to be more correct to use the instantaneous values. We also considered removing the remote sensing part of this manuscript since the length was criticized, but it has quite small weight, so we decided to keep it for now.

- The informations on leaf clumping and leaf angle distribution help better understanding the modelling approach. However, "hemispheric" leaf angle distribution is not a common term (no reference is provided); Do the author mean "spherical" distribution?

Apologies, this should have been spherical. It has now been corrected.

The range of variation of the clumping index (Figure S1) could be discussed in the Discussion section; is it consistent with values found in the litterature for this type of vegetation?

Sure, we added the following text to lines 504-506:

“*The simulated clumping index at FI-Sod is close to observed based estimates (Chen et al., 2005; Schraik et al., 2023) and the seasonal cycle is in line with observations that clumping index increases with increasing solar zenith angle (Chen and Cihlar, 1996).*”

- Isn t the radiative transfer scheme of mSCOPE also based on a two-stream approach?

According to our understanding, mSCOPE would be a four-stream approach.

- With an RMSE of 0.77 units wrt GPP simulations using two QUINCY versions (with / without mSCOPE), the authors conclude that the simulations are similar (l254). Table 2 indicates anRMSE between simulations and observations of about 0.88 units at a daily scale

for the same site. Do I understand correctly that the modelling error is similar to the model data mismatch?

Sorry, would the RMSE for the daily simulations compared to observations be 1.47 µmol m^{-2} s^{-1}? 0.88 is the r^2 value of simulations compared to the observations.

- Quantifying the different components of the error budget between model simulations and observations significantly improved the interpretation of the results. However for SIF, as the main outcome is model overestimation, the slope of the regression line may be the most relevant parameter for this diagnosis. The values could be provided in Table 3.

The slope of regression is shown in Figure 4. It could perhaps be thought that the bias (in Table 3) and the magnitude of overestimation would convey a similar message.

- Is there a specific reason why NPQs was not considered at FI-Sod? Its accounting would likely be beneficial for some of the analyses, in particular in §3.4.

Now that the TROPOSIF comparison has been added to the FI-Sod site, it can be seen that the early increase predicted by the NPQs formulation is likely wrong for this site. Because of this, we did not want to add in further complexity to the model, having there a feature that is not supported by the observations.

- The computational efficiency of the radiative transfer models is one of the evaluation criterion. The authors identified limitations in the use of mSCOPE for large scale simulations (§4.2). Later, they discussed the limitations of 1D RT schemes and identified the accounting of the 3D structure of the canopy within QUINCY and RT schemes as a way to improve the realism of the simulations. Wouldn t this in turn increase the computational cost?

Yes, it would and we had that point elsewhere in the manuscript. Now in line 604.

Figures and tables

- The diagrams (Figure S2 to Figure S4) really help understanding the different implementations of SIF calculations within SCOPE. However, it is regrettable that the quality of these figures does not match that of the other figures in the article.

We removed the colors from these figures, as that seemed to be part of the problem.

- Several of the figures are likely not colorblind friendly (the figures mentioned above as well as Figure 6 or Figure 12 for instance).

(We assume that the reviewer is referencing to Fig. S12).

The figures mentioned earlier (S2 to S4) do not have colors any more.

For Figures 6 and S12 we changed the orange color to pink.

- On Figure 3, with the color chosen, mSCOPE and L2SM simulations are hard to distinguish.

Could the difficulty stem from the fact that the lines are close to each other? This is difficult to circumvent for this figure. We checked the figure with the Coblis (Color Blindness simulator) and did not see clear issues with this figure. However, we made the L2SM line brighter for this figure at CA-Obs so that it would hopefully be more easily distinguishable.

- Please homogeneise the notation for "red / far-red" (Figure 3 and others) or " R/FR" (Figure 4).

We have used R/FR in figures where it was necessary in order to fit in enough large label names. In those cases we have always clarified in the caption its definition. We hope that this would be enough for the reader

- Figure S13: It is not clear if the distinct y-axis for the simulations apply to all three sites or not. Why do the simulations at CA-Obs do not cover the same air temperature range as the observations?

Subplots a and c have the same y-axis as the observations. We have now clarified this in the caption.

As mentioned in the caption, the range of the subpanel displaying results for CA-Obs has been cut to show better the variation within the higher temperatures. As in subpanel b, the simulated values get very high, as the model can have very small estimates for GPP during cold temperatures. We have now added extra clarification about this to the caption.

- Figure 8 and Figure S16: The months on the x-axis should be provided.

Figure 8 has two years and also some other plots having two years also have only years denoted. For the Fig. S16 we added the months as well as the model result with NPQs.

- Figure S12: Different symbols (or colors) should be used for the results with/without NPQs.

Changed.

- Table 2: I understood that the "upscaled" approach was not considered anymore…

Thanks for noticing this. We removed the line for the “upscaled” from this table and other tables in the supplement.

- Table S1: What is the point of indicating "r.u." if no value is provided?

This is quite a regular way to express the unit for this variable in the literature so we would rather keep it this way.

- Is Table S5 really necessary? R^2 and RMSE could be provided in Figure 6 and I do not see what the values of a and b add.

For some people reading the r^2 and RMSE values in a table is more clearer than adding these (altogether six values for each subpanel) to a plot. It is quite common in scientific literature to show the values of the fitted parameters, so we'd like to keep them.

Other comments:

- l15: What does "seasonal development" mean?

"Seasonal development" was here referring to GPP and SIF. Since this was too unclear, we replaced this with "seasonal cycle".

- l21: What are those "numerous applications"?

We removed that sentence, since the reviewer asked to shorten the text.

- l41: "The variability in radiative transfer through the canopy": What does that mean?

Sentence removed.

- l55: The sentence remains unclear.

The sentence was:
"*The model by Johnson and Berry (2021) has a tight coupling between photosynthesis and ChlF and allows for two-directional modelling, estimating SIF from GPP and vice versa.*"

We changed this to:
"*The model by Johnson and Berry (2021) has a tight coupling between photosynthesis and ChlF and allows estimating SIF from GPP and vice versa*."

- l63: Data assimilation studies have been conducted with other models than BEPS (some of the related studies are already cited in the paper).

Many of the studies mentioned in the paragraph (ORCHIDEE and BETHY) also did data assimilation. We clarified this now in the text.

- l108-l109: What are those satellite missions?

FLEX added.

- Title of §2.1: The observations are actually described in §2.2.

We removed the word "observations" from the title.

- l117: Please correct the link for Ameriflux.

Corrected.

- l134: Do the authors mean "over all three scans" (instead of "over all the observations")?

Yes, we changed to that.

- l164: Why not present the surface extent evenly for FI-Sod, given the site is considered to assess the impact of spatial averaging (Figure 7)?

Sure, this is 56 km x 22 km and has now been added.

The presentation of the pixel composition in terms of biomes better fits in the Discussion.

The reason we wanted to place it in the Methods was to make the reader clear why we are not using ENF occurrences from the data.

- l198: "ten layers" was already presented in l178.

We removed the latter occurrence.

- l210: "constrained to values larger than 10°": What does that mean?

When the value would be lower than ten degrees, it is given a value of ten degrees.

However, we had a mistake here, as lower than 10 degrees is the threshold for the solar elevation angle and 80 degrees would then be the threshold for the solar zenith angle.

We added an explanation to the text:

"*i.e. it will receive value of 80 in case of a lower zenith angle*"

- l215: "...was caused by the different Farquhard et al. model formulation...". Not clear. Does this relate to something described earlier in the manuscript?

Yes, it refers to the model description earlier in the manuscript:

"*Photosynthesis is calculated according to Kull and Kruijt (1998). This approach is based on the biochemical model of Farquhar et al. (1980), but instead of the regular implementation of having the minimum of the two branches limiting photosynthesis (light-limited rate of photosynthesis and carboxylation capacity limited rate), the amount of light-saturated region in the leaf is taken into account.*"

The way the implementation has been done was mentioned in S1.4.

- Section 2.6: Most of the detailed informations could be moved into Supplementary Materials.

The details were added here because the reviewer requested them on the earlier reviewer round. As requested, we've moved information to the supplement.

- l325: Is the fact that no spinup is needed related to the fact that LAI and leaf N content are prescribed?

In the same paragraph we said that the canopy model does not require spinup. Because there is no soil described, some canopy properties must be described.

- l337-l339: Is that sentence really needed in this "evaluation methodology" section?

We thought that this would be a suitable point to bring this point up, as not all the modelling papers bring in aspects of GPP modelling, but as the reviewer has pointed out, the manuscript is too long and we dropped this.

- l354: The first sentence should be moved to the introduction of §3.

Sentence was removed and content added to the introduction.

- Title of §3.1.3: "Performance comparison" instead?

Sure, we modified the title according to the suggestion.

- l420: Why not showing these APAR data on the figure?

They were in the figure, but the legend had not been revised. We now fixed the legend.

- l501: "close to...": But above or below?

It seems to be below. This point is based on looking at the figure 2 of the Parazoo et al. (BG, 2019 study. The summertime midday values range between 0.4 to 2.1 W m^{-2} μm^{-1} sr^{-1}, and the model mean is around 1 W m^{-2} μm^{-1} sr^{-1}. The midday value for US-NR1 for the far-red region with QUINCY would be around one also, a bit below. As this comparison is based only on visual comparison of the published plot, we rather keep the text only with “close to”.

- l543: "leaf" optical properties?

We removed this sentence in order to shorten the discussion.

- l567-l568: The reference to the figure is missing.

Thanks, reference is added now.

This "linear relationship" at FI-Sod does not align well with the fitting curve chosen in Figure 6.

No, it doesn’t fit and we’ve discussed that.

- l586: This uncertainty is consistent with the retrieval error already indicated in the TROPOSIF paper…

We added here also the reference to that paper.

- l588: How may site heterogeneity influence the accuracy of TROPOSIF data? Do the authors mean that interpreting the comparison of different datasets at different scales has to be balanced depending on spatial heterogeneity?

The sentence has been removed.

- l601: Are the studied sites concerned by such water-stressed conditions?

Yes, one can see this very clearly in the long time series from the Sodankylä site, where drought years have considerable impact on GPP. As that one is not yet published (a manuscript is in preparation by M. Aurela), we added here a reference to Thum et al. (2007), that shows a milder drought event.

- l653-l654: "misrepresentation of conifer leaves"... Do the authors refer to their structural description/organisation or their optical properties?

We refer to the structural description of needles and their spatial clumped arrangement. We have rephrased the statement (lines 600-604) as follows to clarify

"We hypothesize that the consistent overestimation might arise from the misrepresentation of conifer needles and canopy. The leaf plate-theory-based radiative transfer models does not reproduce the cylindrical structure of needles. At the same time, the 1D canopy radiative models do not incorporate the strong clumping of conifer needles. However, no fluorescence emission has been implemented in needle-like leaf radiative transfer models, and 3D canopy transfer modules are too computationally demanding for TBMs."

- l660: "The TROPOSIF product from satellite", the formulation is awkward.

Apologies for the awkward formulation. We had changed that as a response to the earlier comment round. We changed this into “The TROPOSIF product”.

References

Chen, J. M. and J. Cihlar, "Quantifying the effect of canopy architecture on optical measurements of leaf area index using two gap size analysis methods," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 33, no. 3, pp. 777-787, May 1995, doi: 10.1109/36.387593.

Chen, J. M., Menges, C. H., & Leblanc, S. G. (2005). Global mapping of foliage clumping index using multi-angular satellite data. *Remote Sensing of Environment, 97*(4), 447–457. https://doi.org/10.1016/j.rse.2005.05.003