# 1 Author General Response

## 1.1 Observational Uncertainty Calculations

As pointed out by both reviewers the calculation of the observational uncertainty requires clarification. To address these recurring comments we have done the following:

- Re-written the section in the methods on the calculation of observational uncertainties. We have gone through the calculation and justified it step by step to help readers follow what is being done and why.
- Provided a simplified equation in that section that approximates the (seemingly confusing) area-weighted uncertainty.
- Attached an additional section in supplementary material giving further details of this calculation and the exact formula. As part of this we include an example calculation for a grid cell over the Amazon with accompanying figures showing the original data and the final calculated uncertainty (for the whole 12 months).

We also clarify here. In calculating the observational uncertainties we make the assumption that the observations are independent, i.e. have uncorrelated errors. This is the same assumption made in Parazoo et al. (2013,2014).

This means, effectively, that with the aggregation of GOSAT grid cells into a larger region (i.e. the course model grid cells) there is a larger number of observations therefore the uncertainty goes down by the $1/\sqrt{n}$ law (the same occurs when calculating the standard error). This is a well-known occurrence in dealing with satellite observations and it can be surprising to see the effect of going from single sounding precision (relatively large uncertainty) to aggregated regions (relatively low uncertainty). Another way to describe this is that if you aggregate a region you're taking many independent observations (from each sub-region) and getting out just one independent observation, so to preserve the information content of those sub-regions independent observations the uncertainty goes down; this is called the Jacobian rule of probabilities.

Characterizing correlations in errors is a known problem with satellite measurements. For SIF correlated errors may be due to, for example, error in the retrieval zero-level offset. We are currently looking into the effect of the zero-level offset and will add a additional sensitivity test in the results and discussion accounting for this. If measurements have correlated errors the information content is less than without. To be on the more conservative side we scale our uncertainties by $\sqrt{2}$ which increases the uncertainty.

One reviewer also noted that the observational uncertainties over the tropics (and in particular the Amazon) in Figure 3 appear much smaller than expected. We recognise that this needs explaining. Amendments have been made to the methods section clarifying this, but we also clarify here. Again, the two main points above are relevant. Another element of the small uncertainty over the tropics in Figure 3 is that this is an "annual" uncertainty, so this accounts for the fact that during parts of the year the high-latitudes have no data, while the tropics almost always have data, therefore the tropics have more observations which leads to lower uncertainty.

#### 1.1.1 Inclusion of Structural Uncertainties

This point relates to the calculation of the covariance matrix $C_d$. Formally, this is the uncertainty covariance matrix representing observational and model uncertainty. We agree that we must specify this in the methods and have thus changed it.

There are two general types of structural uncertainties.

- First, is a structural uncertainty in the model (i.e. model structural error). This may be due to incomplete process formulation in the model equations. One can address this error by looking at statistics in the model-observation mismatch following an assimilation of the data (Kuppel et al., 2013). This is therefore only feasible following an assimilation of the data to estimate posterior SIF, posterior parameters, and posterior GPP. In the present study, we are only interested in error propagation so we do not perform an assimilation of the data.
- Second, is a structural uncertainty in the observations. This may be due to certain unknown errors in space and/or time due to (for example) systematic errors in the instrument or retrieval algorithm. One example of this for SIF is an error in the zero-level offset (Frankenberg et al., 2011;2014).

We address this issue by conducting a sensitivity test. We introduce a structural uncertainty into the error propagation system to assess the effect on the calculated posterior uncertainties. We incorporate this sensitivity test into the results and discussion to approximate the effect this extra uncertainty may produce on uncertainty in GPP.

## 2 Anonymous Referee #2

Author comments are shown in blue.

### 2.1 Summary

This study evaluates the benefit of assimilating satellite-retrieved chlorophyll fluorescence into a mechanistic land surface model, to reduce the uncertainty in model parameters and simulated gross primary production (GPP). This study indeed tackles a critical issue in the current efforts towards making the most of diverse data infromation content when building efficient carbon cycle data assimilation systems.

There are, however, a few important issues in this manuscript, some of them critical. They are listed in the general comments below, followed by specific remarks/corrections.

## 2.2 General Comments

First, while the manuscript is often fairly written, on numerous occasions sentences are redundant, strangely formulated, thus logical progression of arguments is hard to follow. Frankly, it sometimes feels as if the authors did not read themselves again before submitting the manuscript. It could be just be a matter of style, but in some occasions it simply results in a lack of clarity. While I tried to list specific parts in the Specific comments and Technical comments section, I suggest a strong effort of rewriting in general. That will also make the manuscript much more accessible to modellers/data experts outside the field of CCDAS or even data assimilation at large.

Second, and perhaps more importantly, the way the observation uncertainty used in Eq. 1 is defined is quite vague. Judging from the elements presented in Sect. 2.4, it seems that only the 'measurements' uncertainty of GOSAT retrievals of SIF is accounted for in $C_D$, neglecting the structural uncertainty ($C_T$ , using the notation of Tarantola (1987)) of the BETHY-SCOPE model. If structural uncertainty is considered, that should be detailed in Sect. 2.4. If $C_T$ is not taken into account, this would bear important consequences. While $C_T$ is hard to estimate explicitly (although some diagnostic methods exist, e.g. see Desroziers et al. (2005), applied to land surface models by Kuppel et al. (2013)), its magnitude and structure might be comensurate or even dominant over measurement uncertainties when building $C_D$. Not including it in Eq. 1 would then largely underestimate the posterior uncertainty of parameters and, by propagation that of modelled GPP. As noted for another reviewer, this would constitutes a serious theoretical flaw in the scope of this study and make it unsuitable for publication. Please refer to general response section.

## 2.3 Specific Comments

P2, L11-12: This sentence is rather vague, can the authors be more precise and add references to support this assertion? We have re-worded this and provided references.

P2, L27-28: Data assimilation is not only used with mechanistic models nor for ter- restrial carbon cycle modeling. I suggest to reformulated, for example:

"In the case of mechanistic models, this is done by constraining the simulated underlying processes.". Good point, we have amended this as suggested.

P2, L28-32: In this review of the state of the art, efforts from other groups to build "mechanistic" CCDAS might deserve to be cited as well, e.g. (Peylin et al., 2016) and the discussion/review by MacBean et al. (2016). Absolutely. We have amended this in the manuscript.

P3, L4-6: Some references would be necessary to back these assertions. We have added references to these points.

P5, L8: The last sentence of this paragraph feels rather clumsy, it should reformulated. We have re-written this last sentence.

P5, L9: Table A1 is rather long and that is fair game given the number of parameters, yet to make it more reader-friendly I would suggest to:

- include a description column for each type of parameter,
- add the corresponding PFT between brackets for all PFT-dependent parameters, as is done for Vcmax,
- add "subsection rows" with parameter categories (leaf growth, ecophysiology etc.).

Good point. We have updated the table as suggested.

P6, L2-3: It is because the PDFs of parameters and observations is treated as Gauss- sian that it can be described by their first two moments, mean and standard deviation (taken here as the metric of uncertainty, that might need to be specified here already well), not the other way around. Yes that is correct, we have amended this.

P6, L1-4: The definition of observations here should be more precise; the reader (especially if not familiar with the data assimilation vocabulary) would assume it relates to measured observations (as the previous paragraph uses "SIF observations" to designate measurements), while in a rigorous probabilistic framework it should refer to quantities in the observation space (including measurements and model outputs, see General comments). We thank the reviewer for the clarification. This section has been more explicit here to make it clear what observational information is, in particular reference to $C_d$.

P6, L12-13: I guess that the authors meant with this sentence that a) in a linear world H is independent from x, but b) this is an oversimplication, therefore c) bringing limitation in accuracy to a method relying on H(x0) to approximate H(xpost). It is not clear at all from the current formulation, which even almost suggest that because of linearity the choice of x0 can influence the results (through a changing H)... We thank the reviewer for pointing out this possible misinterpretation. This has been re-formulated to ensure it is clear.

P6, L13-21: I am not sure how "the use of prior knowledge limits the effect of this problem": is it because we assume that the posterior parameters values will be close enough to the prior set, so that H(x0) is anyway similar to H(xpost) even if the model is not linear? In addition, the authors should give a reference for Eq. 1 (e.g., Tarantola, 1987) and explictly state that because linearity is assumed it takes the formed expressed in this manuscript (while the general equation is $C_{x_{post}}^{-1} = C_{x_0}^{-1} + H(x_{post})^T C_d^{-1} H(x_{post})$). Yes, in part it is because we assume $x_0$ is close the the global optimum that would be obtained in a full assimilation i.e. $x_{posterior}$. Considering the parameter space is vary large, the use of prior knowledge places the parameters into a reasonable physical range. Subsequently the sensitivities calculated in H are more reasonable. We also assume that these functions are smooth. The BETHY-SCOPE model also has no step functions, (which would cause large differences in H even for a slightly different $x_0$). In fact, even if there were step functions, Knorr et al. (2010) points out that a population of plants that, individually, have step functions, average up to a smooth function across a grid cell. We thank the reviewer for the clarification, we have amended the methods section to reflect this point.

P7, L6: "those observations" is at best vague and at worst confusing, since it seems to relate to "observational uncertainty" (rather than "uncertainties") but again, observational uncertainties normally also includes the model component. We have re-written this sentence to be more precise.

P7, L27 - P8, L8: In this whole paragraph (and the derived results and discussion), it would be important to mention which uncertainty is dealt with (random or systematic). Since only the random error can be studied this kind of framework, the potential impact of a systematic error (a bias) should be discussed as well, or at least mentioned. Agreed. The error we can consider is a random error of unknown sign, which would still in fact be systematic as we apply a scaling factor to all of the forcing data. As another reviewer pointed out, we do not consider a known bias (i.e. systematic error of known sign) as this should be corrected for in the data already. We have now clarified this section.

P8, L10-11: Any proof/reference this it is sufficient? Even if it is expert knowledge, the authors should at least state it. Perhaps "sufficient" is the wrong word here. In using a low-resolution grid, this model equations are the same as a high resolution grid such that $H$ relates SIF and GPP to parameters in effectively the same way. And considering Gaussian uncertainties propagate linearly in this study (i.e. with associated assumptions), the model grid resolution does not matter so much. We have re-worded this.

P8, L22: "Effective constraint" rather than "constraint", might be more accurate. Yes, this might be more accurate. We have changed "constraint" to "effective constraint" where ever necessary.

P9, L9: Which global physiological parameters are the authors referring to? Rows 37-68 in Table A1? See earlier comment on making Table A1 clearer. We have amended table A1 as suggested so this should be more clear now.

P9, L10-17: The values of constraints in the text do not correspond to those shown in Table A1. Please update.Good catch! Thanks. We have updated these values.

P10, L3: Maybe add between brackets than the chlorophyll parameters are $C_ab$ components. Yep, thanks. This should be shown as $C_{ab}$ rather than worded "chlorophyll", so we have amended this.

P10, L3-4: "During the assimilation" comes a bit abruptly. I guess the authors are talking about prospective data assimilation efforts with BETHY-SCOPE and SIF, please expand to make easier for the reader to understand. We have amended this to say " Thus, during a full assimilation with the SIF data only the sum..." as is done in other parts of the paper.

P10, L9: This is a somewhat confusing formulation to say that uncertainty (and its subsequent reduction) is quantified as one standard deviation. Maybe giving this reference metric already in the methods would be helpful. Okay. We have added in this reference metric to the methods section under Uncertainty Calculations.

P10, L10-15: I suggest to have Fig. 3 (not mentioned in the text, maybe already in Sect. 2.4.?) on the color same scale as Figs. 4 and 5. We have now referred to it in the text. Figures 3 and Figures 4/5 are different quantities (SIF and GPP, respectively) so we don't think they need to be on the same scale. However, we note that better labeling is required for these figures to make it clear they're different quantities, so we have done this.

P11, L3: A figure showing the uncertainty reduction Could the authors briefly detail how they assessed the relative contribution of covariances to the total uncertainty in GPP? By summing the non-diagonal terms in $H_{GPP}C_xH_{GPP}^T$? Using equation 3 we assess the constraint with the full covariance matrix $C_x$ (i.e. including off-diagonal terms). Then we assess the constraint with off-diagonal terms set to zero in $C_x$. The difference between these two cases is the contribution of correlations. We have now outlined this in the manuscript.

P11, L7: Could the authors briefly detail how they assessed the relative contribution of covariances to the total uncertainty in GPP? By summing the non-diagonal terms in $H_{GPP}C_xH_{GPP}^T$ ? Yes, we have added in an extra sentence explaining what we did as follows "we can assess the contribution of these correlations to the constraint of GPP by setting all off-diagonal elements in $C_{x_{post}}$ to zero in Equation **??**, the difference between this and the standard case that uses the full $C_{x_{post}}$ equates to the contribution of correlations".

P11, L17: As the authors state in the discussion, the fact that GPP is relatively insensitive to Cab derives from the lack of a mechanistic link in the model between chlorophyll content and carboxylation rate. I suggest therefore to remove the "discussive" end of this sentence here and leave for the discussion where it is explained. Okay, good point.

P11, L23-24: I disagree with the last part of this sentence: it seems to me that the increase in relative uncertainty contribution of physiological processes only says that they are less constrained than other processes, therefore the stated "limitations" is just relative to other well-constrained parameters. Without looking at the absolute value of uncertainty in GPP arising from each group of parameters (from which is then calculated the relative contribution), no statement can be made about how really "limited" is the constraint of SIF in ultimately reducing the uncertainty of a given parameter to simulate GPP. That is correct and a good point to make. We have amended this statement to say "Uncertainties in physiological parameters are constrained less than the leaf growth parameters which results in them contributing more in relative terms to the posterior uncertainty of GPP".

P11, L27 to P12: I feel that an additional figure would be needed here, to show how the constraints in GPP from given parameters groups changes across the year in Tem- perate and Boreal regions. It could be for example a monthly-binned boxplot, each box corresponding to the range of constraint GPP for a given group of parameters, using colors or panels to separate regions. That would help the reader to support all the description given in the main text. Okay, good suggestion. We have added another figure here to show the contributions of parameter groups to uncertainty across the year for each region.

P12, L4: "exaggerated" seems quite subjective. Okay, we have re-phrased this to say "Similar differences between seasonal constraint is seen for the Temperate North, although with a smaller seasonal variation in SIF constraint that ranges between 74% and 87% across the year".

P12, L8: The parameter Vcmax is mentioned, then "these parameters", I guess refer- ring to the different PFT components Vcmax? Please specify. In fact we're referring to the $\tau_W$ parameters, so we have now specified in the manuscript.

P14, L10-14: This might be suited for the discussion section. Agreed. Amended.

P15, L10: How did the authors get this number? As its effectively treated as a parameter, we can assess the relative uncertainty reduction by the same equation used for parameters (i.e. 1-$\sigma_{post}/\sigma_{prior}$). We have specified how this is done in the manuscript now.

P16, L810: I would move this sentence to the next paragraph, where diurnal dynamics are discussed. Amended.

P16, L31-35: And additional figure showing the relative contribution of each parameters to modelled GPP uncertainty would make the results clearer. Perhaps using the same barplot setup as Fig. 1, except that y-axis would relative contribution to GPP uncertainty, and prior and posterior results could be shown using mirroring bars (2 y-axis would be needed then, one going upwards and the other downards). This is a good point, an additional figure will help readers follow results+discussion. We have created a figure similar to described: with classes of PFTs and their contribution to uncertainty in GPP (in Pg C yr-1) across the year. We thank the reviewer for this suggestion.

P16, L33: "Free" sounds a bit odd here, what do the authors want to say? Just that it is the only process parameter that is optimizable (i.e. not a fixed parameter). We have removed "free" and changed this to be "process parameter".

P16, L34-35: I assumes that by "[. . . ] only other free parameter controlling leaf area index other [. . . ]" the authors mean that the model is highly sensitive to this parameter (i.e., large values in H), so adding to little prior parameter knowledge results indeed in large propagated uncertainty. The first aspect is however not quite clear from the current formulation. Since this separate consideration of sensitivity and parameter knowledge is essential when considering output uncertainty, here in the discussion I suggest detailing a bit more these aspects. Useful supporting references are, e.g., discussions in Dietze et al. (2014) and Kuppel et al. (2014). This is helpful. We have clarified this in the manuscript.

P17, L1-2: This sentence ("The prevalence [. . . ] global scale") is rather general and does not add much to the following one (which gives numbers). I suggest removing the former. Agreed. Amended.

## 2.4 Technical Comments

P2, L16: Definition of NDVI and EVI acronyms, first introduced here . Amended.

P2, L23: has instead of have. "has" doesn't sound right to me.

P2, L35: It is not the process that provides the constraints, rather the latter being constrained! Amended.

P6, L9: Replaces "equation 1" by "Eq. 1". It also applies to L17, to "equation [2,3,4]" on [P6;L26], [P7;L2-L4-L14] and [P10;L8]. Amended.

P6, L10-11: Strange formulation, I would suggest: "[. . . ] a Jacobian matrix (H ), which is calculated around [. . . ]" Amended.

P6, L26: "p. 71" instead of "pg. 71". Amended

P7, L6: "its" instead of "it's". Amended.

P7, L10: "described" would be more accurate than "demonstrated" This section has been re-written, "demonstrated" is no longer present.

P7, L27-29: "As might be expected" is quite subjective. I suggest to connect the two sentences: "[. . . ] while uncertainty in forcing such as incoming radiation is not considered in the curret CCDAS setup, it is considered to be an important variable driving SIF (Verrelst et al., 2015) and GPP (rference needed)." Amended.

P9, L2: "Table A1" instead of "Table 1". Amended.

P10, L7: If ”as” refers only to the posterior uncertainty in GPP, it should then be re- placed by ”the latter being”. It refers to both prior and posterior, so we have left this as it is.

P11, L12: ”stems” instead of ”stem”. Amended.

P12: ”made up by” (L2) and ”make up” (L8) are somewhat colloquial/vague here, it could be respctively replace by ”arises from” and ”contribute to”.Amended.

P14, L4: Changing with ”Second, we also increase [. . . ]” might help the reader under- stand you are describing the other experiment. Good point, amended.

P15, L7: I suggest ”[. . . ]SWRad, in both cases resulting in a relative reduction in the GPP uncertainty by about 78.6%”. Amended as suggested.

P15, L17-18: ”constraints” is repeated a lot here, I suggest: ”[. . . ] ultimately yields a global annual GPP estimate within ± 2.8 PgC.yr1.”.We have altered this sentence already to specify that it is ”parametric uncertainty” that is reported. It reads: ”..and ultimately yields a parametric uncertainty in global annual GPP of ± 2.8 $PgCyr^{-1}$”

P16, L18: ”however” seems somewhat redundant. Agreed. Amended.

P16, L18: ”PSII” should be defined on L11. Amended.

P17, L9: ”feasible with” feels odd. Maybe ”acheviable using”? Amended.

P17, L23-24: I suggest rephrasing as follows: ”This in line with Koffi et al. (2015) who found limited sensitivity of simulated SIF to Vcmax.” Amended.

P18, L7-8: The meaning is not clear, I assumed the authors meant ”While including this forcing uncertainty increases the prior GPP uncertainty, incorporating the former within SIF uncertainty itself mitigates the downstream effect on GPP.” Yes, this sentence is a little confusing. We have re-worded in the manuscript similarly to suggested.

P18, L16: Maybe replace ”can also be” by ”will also be”.Amended.