1 Author General Response

Please find below our response to the referee comments. We appreciate the referees highly useful and constructive feedback, and have taken each comment on board to help improve this work.

Each of the referee comments are responded to by the author in blue text, including any actions taken. We also provide a general response section to address recurring comments. We would also like to add that, following on from the reviewers comments, we found one of our minor points of analysis was incorrect and has therefore been made redundant. This pertains to the analysis of scaled high-resolution observational uncertainties. This has had no effect on our conclusions.

Key changes include:

- Clarification for the calculations of the observational uncertainties.
- An additional section in the appendix showing figures as an example of how we calculated our observational uncertainties; as suggested by Referee # 1.
- Made the parameter table in the appendix much clearer and included descriptions of each one.
- Added additional figures to show the contribution of different parameter classes to uncertainty in GPP across the defined regions; as suggested by Referee # 2.
- Included an analysis of the effect of systematic errors in the observations; as suggested by Referee # 2. This includes additional figures in the appendix.
- Clarified and simplified, where possible, the explanation of aims and method of this study. Referee # 1 seemed to partly misinterpret what we were aiming to do and how we were doing it. In particular that we are not estimating parameter values, we are only assessing information content. We have therefore amended the text in the introduction and methods to make this clearer for readers.

Finally, can the editor please advise on how we put together supplementary material. We think that the Appendix sections A2 and A3 would be better suited in Supplementary material, but we are not sure how to do this. Thank you.

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 # 1

2.1 Summary

This paper uses satellite observations of Solar-Induced Fluorescence (SIF) in an inversion scheme (CCDAS) to reduce uncertainty in a posteriori estimates of model parameters and outputs, specifically GPP. Interestingly, no attention is given to actual parameter values or GPP estimates; the focus is entirely on how much reduction in uncertainty can be expected due to the inclusion of SIF.

The paper is reasonably well written, and uses a novel approach to attempt to reduce uncertainty in a posteriori estimates of model parameters and output. However, I feel that the paper needs clarification and perhaps some reorganization to help readers to follow the story. Furthermore, I believe that the critical issue of observational uncertainty is given too little attention and must be clarified.

The authors provide reasonably comprehensive citations for CCDAS, but the paper reads is if it were written (as it probably was) by someone who is a Data Assimilation (DA) expert. To this reviewer it seems that some details are either implied or 'skipped over'. It is likely that many readers will be DA experts themselves, but the inclusion of SIF will probably draw in readership that may not possess the DA expertise to easily understand what is going on. I may be a member of that part of the audience, so some clarification is warranted. Specifically, the relationship between covariance matrices (Cx, Cd) and standard deviation () is not entirely clear. Good point. We want readers from different audiences to be able to follow what was done easily. We have added and clarified text in the methods section to help non-DA readers relate covariance matrices to standard deviation uncertainty simplified other points where possible. We have also modified the last paragraph of the introduction to make it clearer what the specific aims are.

The description of grids used and observation area ("GOSAT grid cell"; section 2.4) needs clarification. Two grid sizes are mentioned in Section 2.4, but we don't learn much more about them until Section 2.5. Good point. We have amended this as suggested by shifting the grid resolution information to the beginning of section 2. I would like to see a more deliberate explanation of "here is what we are going to do, and here is how we are going to do it". That might fit better in Section 2.1. Some specific Issues:

- Figure 3, showing observational uncertainty, is not referred to in the section describing observational uncertainty. It needs to be. Amended.

2.2 Observational Uncertainty

Eqn 4: I see two ways that this value can be small: 1) there are many observations, and σ^2 is small. 2) There are very few observations, and Area is small. Parazoo et al. (2013) estimated uncertainty as the standard error. This has the effect of allowing a large error in regions with very few observations, like the tropics. Figure 3 in the manuscript under review shows some of the smallest observational uncertainty in the tropics, and that makes absolutely no sense to me. I've worked with the GOSAT data, and over the deepest tropics there are very few observations, which makes me suspicious that your uncertainty is small because of reason 2). Parazoo et al. did not extend their analysis to the wetter parts of Amazonia because they just didn't have enough data to justify it. Now the authors claim that this region has some of the smallest observational uncertainty on the globe! A detailed justification of how uncertainty can be very small over a region with few or no datapoints is an absolute necessity. Please refer to general response section above.

I do not think multiplying by square-root-2 is sufficient to remedy what might be unrealistically low uncertainty values. Please refer to general response section above.

When GOSAT 2010 data is aggregated onto the 1.25x1.0 degree MERRA grid, I see that the maximum number of retrievals for a given month, anywhere on the globe, is between 30-35 or so. Looking at South America, I see that very few MERRA gridcells have more than 10 retrievals in a given month during 2010, and many gridcells have 5 or fewer. Aggregating up to 7x10 (or 2x2) you are not going to get very much increase in sample size. Id like to see the authors address the sparseness of the GOSAT data and explain how this will or will not effect their method. In the amended manuscript we show an example calculation over a 7.5x10 degree grid cell including the GOSAT sub-grid cells to show how this scales across 2010. We see that aggregating from 3x3 to 7.5x10 you get actually see a big increase in sample size. For example in Jan 2010 any GOSAT sub-grid cell may have between 0-20 soundings, but aggregating to the 7.5x10 there is almost 80.

The number of GOSAT observations is invariant and does not change with grid size. The aggregation of GOSAT observations changes with grid size (Section 2.4). This should be clarified. In fact, the number of GOSAT observations does vary with grid size. With a larger grid size you capture more GOSAT soundings. You may refer to the general response section for further details. We have clarified this in the methods section.

An individual GOSAT retrieval has pixel size of around 10 km^2, I believe. OCO-2 will have a pixel size of ~ 5 km^2, and GOME-2 is a 40-80 pixel, or 3200 km2. This will have a large impact on your inversion scheme and the calculation of observational uncertainty. Since this paper only uses GOSAT, the other products probably dont need too much (or any?) explanation, but I do have questions about GOSAT and the grids used:

1. There is the possibility for (possibly) many 10km^2 GOSAT retrievals to be included in a 7.5x10 degree gridcell. For that matter there can be many of them in a 2x2 gridcell too. BETHY-SCOPE tiles in 3 PFTs; how are GOSAT retrievals registered to these PFTs? This is a good point. The observations are not separated per PFT, doing so would effectively triple the information content as there would be three times more observations, which would in fact improve the results. We compare observations at the grid cell scale. Thus information is transferred/split to PFTs through the Jacobian sensitivities, which account for PFT fractions. E.g. if a grid cell is 90% C3Gr, then the SIF sensitivity over that grid cell will be dominated by parameters relating to C3Gr, with smaller contributions from the PFTs that make up the remaining 10%. Thus, the information content of the observations is split accordingly. Are GOSAT retrievals marked with a specific land cover type, and accumulated on a per-PFT basis? What about GOSAT retrievals that are not associated with one of the 3 PFTs tiled into the BETHY-SCOPE gridcell? Are they discarded? Why or why not? We do not attempt to disaggregate observations in this way. We assume there is roughly even coverage across the PFTs, even though the absolute footprint of a GOSAT sounding is about 10km^2, it has a wide swath of around 750 km^2 with 5 footprints. Thus we assume decent coverage. This will be more important to consider in a full assimilation of the data i.e. for estimating parameter values and fluxes.

2. If all GOSAT retrievals within a gridcell are utilized, is the mean taken and used for DA with all 3 PFTs? In this case arent you 'smearing out' the information that SIF provides? Guanter et al. (2012) demonstrate that the linear relationship between SIF and individual PFTs is heterogeneous. Do you take this into account? If so, how? If not, why not? This is true for a full assimilation and parameter estimation but in this study, we do not consider the mean values of the observations, only their uncertainties as we're only interested in information content. Thus these issues are not present.

3. In August 2010 the GOSAT scan strategy was changed; the area observed was decreased, but the number of retrievals over a given region was increased. How does this effect the two questions above? Yes, good point. The observational uncertainties used in section 2.4 are standard errors (although slightly adjusted to increase the uncertainty as described in section 2.4), thus they account for the number of observations per grid cell.

The reduction in uncertainty for global GPP is dramatic (79%). However, this reduction is critically dependent upon Cd (observation uncertainty) according to equation 1. Therefore, I think it is absolutely essential that the questions surrounding the determination of this observation uncertainty are answered in a clear and categorical manner. Agreed. We have clarified our calculation of the observational uncertainties in the manuscript. Please refer to general response above.

Im not a DA expert, but I do collaborate with quite a few people who are, and I think I understand the basics. The covariance matrices are absolutely fundamental to the outcomes of a DA experiment: If the observational uncertainty is small and the model uncertainty large, the a posteriori outcome can be pulled strongly towards the observations. If the opposite is true, then it will be hard to budge the inversion away from the model prior. Is this correct? Essentially, yes this is true. However, we note that this question primarily applies to a an assimilation with real data. In this paper we assess the information content of SIF observations, i.e. only uncertainties of model parameters and GPP, not their values. We can do this because this is a linear problem, whereas the full assimilation is a non-linear problem and the subject of subsequent study. The point regarding observational uncertainty vs model uncertainty is pertinent however, and we address this in the general response section.

In this paper the first case is presented: the observational uncertainty is, to my eye extremely small and therefore results in an amazing reduction in uncertainty in the a posteriori result.

The absence of evaluation of actual posterior values of either parameter or flux values may actually hinder the analysis. If the result of the study is an outlandish value for global GPP, then that might indicate a problem. Of course, estimates of global GPP vary by about a factor of two (Huntzinger et al., 2012), so maybe this wouldnt help as much as one might hope. However, posterior parameter and flux values might offer insight, and a comprehensive evaluation of method and results (values of parameters and flux) could provide more support for the authors' conclusions. Was this considered? Why or why not? Im suspicious that posterior flux and parameter values were outlandish, and a choice was made to focus on method even though results may be untrustworthy. I suspect many readers will have this suspicion too. Assessing information content of the observations is a linear problem which can be performed independently of comparing actual values of model and observed data. This is convenient as an assessment of the information content tells us whether SIF is going to be a useful constraint on GPP before we have to go through the challenging process of fully assimilating the data. We also believe that the information content study here is substantial enough. Adding in a full assimilation to estimate parameter and GPP values is a complicated non-linear problem and adding this into the current manuscript would make for too large a study. An assimilation of the data where one actually estimates global GPP is the subject of subsequent study.

A detailed description of the construction of the observation uncertainty may detract from the papers readability, but including it in an appendix would be appropriate. Additionally, I would like to see, perhaps in supplemental material, a step-by-step description of the calculation of the observation uncertainty, perhaps in the 7x10 gridcell that contains Manaus, Brazil. Agreed. Refer to general response section.

To see such a large reduction in error sent warning bells ringing with me; I dont think it is an overstatement to say that the entire paper depends on the observation uncertainty. If the authors can demonstrate that the values shown in Figure 3 are justifiable, then the paper has merit. If not, I think the whole endeavor falls apart, as the structural underpinning would have disintegrated. In that case the paper is not worthy of publication.

2.3 Specific Comments

- Figure 2: The information here is too dense (small labels, tiny resolution on the plot) to follow. If the only pertinent information is in the lower-right- hand of the plot, why not omit the rest and enlarge this sector of the graph? Good point. We have edited this figure to make it clearer. We have removed any rows/columns that have no correlations and increased the font size.
- Figure 2: There is very little description of the graph and what it means. Again, this may be another case where the authors are assuming that their readers look at graphs like this every day and know what it is showing. Yep fair enough. We've provided a better description in the text and caption.
- Figure 3: What are the units? Amended.
- Figure 4: Absolute uncertainty annual GPP will of course correlate directly with productivity. If you standardize the time series and look at relative uncertainty I imagine that map will look very different. Have you done this? If you have, do Figures 4-6 look similar or different? We need some clarification here from the reviewer. We can do the following: prior uncertainty divided by prior GPP and posterior uncertainty divided by prior GPP. But we cannot do the following: posterior uncertainty divided posterior GPP. As this is an error propagation study we have not estimated posterior GPP.
- Table 1A: There is no description of what these parameters are and what they do. There are sporadic mentions in the text, but for the most part the reader is left to ones self to figure out what these parameters are for. I would like to see a column added (there appears to be room, as the uncertainty reduction columns could be re-formatted) with a couple of words or a phrase describing each variable. Section 3.2: line 14 on page 11 mentions that τ_W makes up 82% of the global annual uncertainty in posterior global GPP. The reader does not know what W is. At the end of Section 3.1 there are several other parameters listed, and again the reader is not told what they are. It might be helpful to have a short description in parentheses following the listing of each parameter, but I would prefer to see that information in table 1A. Good point. We will amend Table A1 and make it clear what the parameters mean if referring to them in text.

- Boilley and Wald (2015) discuss a high bias in the radiation from reanalyses. Im not sure this is the same as the uncertainty mentioned in sections 2.4 and 3.4. Can you elaborate? We were not aware of the Bilolley and Wald (2015) study, so we thank the reviewer for the citation and we have included it in the manuscript. A known bias in the radiation such as this should be removed from the reanalyses data before it is distributed. We consider an uncertainty of unknown sign, as shown in Kato et al. (2012), which can be accounted for in the prior uncertainty and constrained through the error propagation system as we demonstrated. We clarify this in the manuscript.

- Page 17, lines 7-8: "...we can predict and quantify how SIF wil constrain the uncertainty of process parameters and GPP, but we cannot predict how their values will change". Why not? Can't you back the posterior values out of the a posteriori covariance matrices and the Jacobian? Isn't the whole point of DA to obtain these posterior values? The process of getting posterior parameter values and obtaining posterior fluxes is a nonlinear problem that is therefore arduous and challenging. So, before one goes down this path they can actually assess whether it is worthwhile doing by first assessing the information content, this is linear problem and therefore simpler. However, as SIF has not been used in a full DA system with a process-based model like this before it is valuable to show, in detail, what SIF may constrain, how it does it, and any caveats to this. It seems we have not made it clear enough exactly what this study is and exactly why we are doing it. We have therefore added in some extra points to the introduction and methods section to clarify this.

3 Anonymous Referee # 2

3.1 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.

3.2 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 CD, neglecting the structural uncertainty (CT , 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 CT is not taken into account, this would bear important consequences. While CT 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 CD. 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.

3.3 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 Gaussian 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 $C^{1}_{x_{post}} = C^{1}_{x_0} + H(x_{post})^T C^{1}_{d} H(x_{post})$). 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 ac- curate. 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 Cab 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 dont think they need to be on the same scale. However, we note that better labeling is required for these figures to make it clear theyre 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 $H\ GPP\ C_x\ H^T$? GPP 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 $H\ GPP\ C_x\ H^T$? Yes, we have added in an extra sentence GPP 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 Cx-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 Temperate 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 referring to the different PFT components Vcmax? Please specify. In fact were referring to the τ_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.

3.4 Technical Comments

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

P2, L23: has instead of have. has doesnt 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 its. 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 replaced 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 understand 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 PgCyr1

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.

Assimilating solar-induced chlorophyll fluorescence into the terrestrial biosphere model BETHY-SCOPE v1.0: Model description and information content

Norton Alexander J.[1], Rayner Peter J.[1], Koffi Ernest N.[2], and Scholze Marko[3]

[1]School of Earth Sciences, University of Melbourne, Australia
[2]European Commission Joint Research Centre, Ispra, Italy
[3]Department of Physical Geography and Ecosystem Science, Lund University, Sweden

Correspondence to: A.J. Norton (nortona@student.unimelb.edu.au)

Abstract. The synthesis of model and observational information using data assimilation can improve our understanding of the terrestrial carbon cycle, a key component of the Earth's climate-carbon system. Here we provide a data assimilation framework for combining observations of solar-induced chlorophyll fluorescence (SIF) and a process-based model to improve estimates of terrestrial carbon uptake, or gross primary production (GPP). We then quantify and assess the constraint SIF provides on the uncertainty of global GPP through model process parameters in an error propagation study. By incorporating one year of satellite SIF observations from the GOSAT satellite, we find that the parametric uncertainty in global annual GPP is reduced by 79%, from ± 13.0 $\mathrm{PgCyr^{-1}}$ to ± 2.8 $\mathrm{PgCyr^{-1}}$. This improvement is achieved through strong constraint of leaf growth processes and weak to moderate constraint of physiological parameters. We also find that the inclusion of uncertainty in shortwave down radiation forcing has a net-zero effect on uncertainty in GPP when incorporated in the SIF assimilation framework. This study demonstrates the powerful capacity of SIF to reduce uncertainties in process-based model estimates of GPP and the potential for improving our predictive capability of this uncertain carbon flux.

1 Introduction

The productivity of the terrestrial biosphere forms a key component of Earth's climate-carbon system. Estimates show that the terrestrial biosphere has removed about one quarter of all anthropogenic CO_2 emissions thus preventing additional climate warming (Ciais et al., 2013). Much of the interannual variability in atmospheric CO_2 concentration is also driven by terrestrial productivity. Despite this significance, understanding of the underlying mechanisms of terrestrial productivity is still lacking. This manifests in a large uncertainties in the predictive capability of terrestrial productivity and thus, future predictions of atmospheric CO_2 and temperature ~~projections~~ (Friedlingstein et al., 2006).

A key challenge is disaggregating the observable net CO_2 flux into its component fluxes, gross primary production and ecosystem respiration. Gross primary production (GPP) is the rate of CO_2 uptake through plant photosynthesis and the largest natural surface to atmosphere flux of carbon on Earth (Ciais et al., 2013). Estimating spatiotemporal patterns of GPP at the scales required for global change and modeling studies has, however, proven difficult. This is primarily due to two reasons, the complexity of the processes involved and the difficulty in observing those processes (Baldocchi et al., 2016; Schimel et al., 2015). Remote sensing observations of solar-induced chlorophyll fluorescence (SIF) offer a novel constraint on GPP and the potential to partly address these two issues (Schimel et al., 2015).

At the leaf scale chlorophyll fluorescence is emitted from photosystems I and II during the light reactions of photosynthesis. These photosystems are pigment-protein complexes that form the reaction centers for converting light energy into chemical energy. It is in photosystem II (PSII) where photochemistry, the process initiating photosynthetic electron transport and leading to CO_2 fixation, is initiated. The link between chlorophyll fluorescence and photochemistry is confounded by a third key process however, heat dissipation, also termed non-photochemical quenching (NPQ). Both photochemistry and NPQ are regulated processes, responding to changing physiological and environmental conditions (Porcar-Castell et al., 2014) . Changes in the rates of photochemistry and NPQ, and electron sinks other than CO_2 fixation, lead to a non-trivial, but direct link between chlorophyll fluorescence and photosynthetic rate ~~. However, it is because chlorophyll fluorescence emission~~ (Flexas et al., 1999; Magney et al., 2017) . Because chlorophyll fluorescence is tied in with these physiological processes ~~that~~ it has become ~~such a~~ a highly useful indicator of ~~the actual physiological state at the leaf scale~~leaf physiological state (see reviews by Baker, 2008; Porcar-Castell et al., 2014) .

At the canopy scale and beyond the link appears simpler, exhibiting ecosystem-dependent linear relationships (Guanter et al., 2013). The slope of this linear relationship can change as the light-use efficiency of either SIF or GPP changes, for example due to water stress (Daumard et al., 2010) or changing light conditions (Yang et al., 2015). SIF also seems to outperform traditional remote sensing methods, such as Normalized Difference Vegetation Index (NDVI) and the Enhanced Vegetation Index (EVI) that use reflectance to derive vegetation indices~~(e.g. NDVI, EVI) when~~ , in tracking changes in GPP at this scale (Yang et al., 2015; Walther et al., 2016). This is in part because the SIF emission originates exclusively from plants, thus the retrieval is not contaminated by background materials like soil or snow. It is expected, however, that complicating factors such as the retrieval wavelength, temporal scaling, chlorophyll content, 3-dimensional canopy structure, and stress will also play a role in the GPP-SIF link (Damm et al., 2015; Guanter et al., 2012; Rossini et al., 2015; Zhang et al., 2016). Using high-resolution spectrometers onboard satellites ~~,~~global maps of SIF have been produced. A number of existing (GOME-2, GOSAT, OCO-2, TROPOMI, SCHIAMACHY) and planned (FLEX, GEOCARB) satellite missions are capable of measuring SIF. Utilizing these remotely-sensed SIF observations directly to track changes in GPP have already proven useful even without the addition of ancillary data or model information (Lee et al., 2013; Parazoo et al., 2013; Walther et al., 2016; Yang et al., 2015).

Data assimilation enables the use of observations and model information together to produce a best estimate of the state and function of the system. ~~This~~ In the case of mechanistic models this is done by ~~providing a mechanistic model constraint based on underlying processes of terrestrial carbon cycling~~constraining the simulated processes and their parameters. Such an approach has been ~~used in the Carbon Cycle Data Assimilation System (CCDAS) to~~ applied to terrestrial biosphere models to optimize model parameters and constrain uncertainty ~~in terrestrial~~ of carbon flux estimates ~~(see Kaminski et al., 2013; Koffi et al., 2013) . The CCDAS~~ in a number of studies (see Kaminski et al., 2013; Koffi et al., 2013; Macbean et al., 2016; Peylin et al., 2016) . The Carbon Cycle Data Assimilation System (CCDAS) is one such system and has ingested observations such as atmospheric CO_2 concentration and/or the fraction of absorbed photosynthetically active radiation (FAPAR), demonstrating the benefit of combining model and observations in a regularized approach (Rayner et al., 2005; Kaminski et al., 2012). The use of SIF observations ~~in such a system~~ within a data assimilation framework may provide a highly useful, complementary constraint on GPP. ~~This may enable estimates of the spatiotemporal patterns of GPP to be improved and the parametric uncertainty of models to be reduced.~~ While one study by Parazoo et al. (2014) utilized SIF in a data assimilation system to redistribute multiple model estimates of GPP, no optimization of model process parameters was performed. Koffi et al. (2015) incorporated a mechanistic model for SIF into the CCDAS system and conducted sensitivity tests and a comparison of the model SIF and observed SIF from GOSAT demonstrating the model is capable of ingesting the data. However, SIF has not yet been used on a global scale in a data assimilation system. A key first step toward this is to quantify the potential constraint that SIF provides on ~~GPP and assess which underlying processes provide the constraint, termed here~~ the underlying processes that drive GPP and, hence, on GPP.

In this paper, we assess the ability of satellite SIF observations to constrain the parametric uncertainty of simulated GPP in a terrestrial biosphere model within a data assimilation system. This is termed an error propagation study and is similar in concept to an observing system simulation experiment or quantitative network design study (Hungershoefer et al., 2010; Kaminski et al., 2010; Koff
Parameters and simulated GPP are therefore optimized only for their uncertainty and not for their absolute quantities. Considering SIF is a novel observational constraint, this is an important first step toward a full assimilation of the data allowing us to evaluate the level of constraint SIF will impose on GPP and how that constraint is propagated through the model. ~~Here, we utilize a new modeland satellite SIF observations to determine how effectively SIF constrains model process parameter uncertainties and the uncertainty of GPP globally.~~

2 Methods

~~There is a growing literature around the problem of optimizing parameters in terrestrial biosphere models using various observational data streams.In particular, the CCDAS has provided a systematic framework for which to optimize parameters and subsequently constrain simulated quantities, namely carbon fluxes, using data such as concentration and/or FAPAR~~

We formulate this error propagation study into two key stages; (i) optimization of parameter uncertainties and; (ii) projection of parametric uncertainties onto uncertainty in diagnostic GPP. This allows us to conduct a thorough assessment of how effective SIF observations are at constraining the uncertainty of model parameters and the parametric uncertainty of model simulated GPP. Under the linear Gaussian assumption, the uncertainty of a target quantity following assimilation of the data (i.e. the posterior) is conditional only on the prior uncertainty, the uncertainty of the observations and the sensitivity of simulated observations to changes in the parameter (Tarantola, 2005) . Thus, this is a linear problem that can be performed independently of the optimization of actual parameter values. Here, ~~the quantity of interest is GPP and the observation is satellite-retrieved SIF.~~ we outline the model used to simulate the observation (SIF) and the target quantity (GPP). We also outline the model parameter set describing these processes, the uncertainty in the observations and model forcing, and general experimental setup.

2.1 Model Description

In order to ~~assimilate such an observation~~ ingest an observation into a data assimilation system, we require ~~an~~ a model or ’observation operator’ that can simulate SIF, ideally providing a process-based relationship between SIF and GPP. There are a few ways one might formulate the observation operator. Evidence shows a strong linear relationship between SIF and GPP at large spatial scales and relatively long temporal scales (Frankenberg et al., 2011b; Guanter et al., 2012), suggesting relatively simple scaling between GPP and SIF. However, it is known that the link is more complex than this, and it is expected to differ at finer spatial and temporal scales due to, for example, land surface heterogeneity or the time of day of the measurements. To ensure the model has these capabilities we have opted for a ~~more complex~~ mechanistic-based observation operator.

2.2 ~~Model Description~~

In this section we describe the newly developed terrestrial biosphere model for simulating and assimilating SIF. The model is an integration of the existing models BETHY (Biosphere Energy Transfer Hydrology) (Rayner et al., 2005; Knorr et al., 2010) and SCOPE (Soil Canopy Observation, Photosynthesis and Energy fluxes) (Van der Tol et al., 2009) and builds upon the developments ~~by~~ of Koffi et al. (2015). The coupling of BETHY and SCOPE enables spatially explicit, plant-type dependent, global simulations of GPP and SIF. This model may be run on a computationally efficient, low-resolution spatial grid of 7.5° × 10° or a high-resolution spatial grid of 2° × 2°.

BETHY is a process based terrestrial biosphere model at the core of the Carbon Cycle Data Assimilation System (CCDAS) (Rayner et al., 2005; Scholze et al., 2007). Full model description details can be found elsewhere (e.g. Rayner et al., 2005; Scholze et al., 2007; Knorr et al., 2010). Briefly, BETHY simulates carbon assimilation and plant and soil respiration within a full energy and water balance. The version used here also incorporates a leaf area dynamics module for prognostic leaf area index (LAI) as described in Knorr et al. (2010). This module includes parameters for leaf development, phenology and senescence processes ~~to determine LAI~~ (hereby collectively termed leaf growth) to determine LAI in a scheme that incorporates temperature, water and light limitations on growth and is capable of representing the major global phenology types (Knorr et al., 2010). This scheme also enables the representation of subgrid variability in leaf growth, representing the likely variability in growth triggers across a grid cell and necessary for differentiability between process parameters and state variables. The full BETHY model consists of four key modules: (i) energy and water balance; (ii) photosynthesis; (iii) leaf growth and; (iv) carbon balance. It represents variability in physiology and leaf growth of plant classes by 13 plant functional types (PFTs) (see Table 1) originally based on classifications by Wilson and Henderson-Sellers (1985). Each model grid cell may consist of up to three PFTs as defined by their grid cell fractional coverage.

SCOPE is a vertical (1-D) integrated radiative transfer and energy balance model with modules for photosynthesis and chlorophyll fluorescence (Van der Tol et al., 2009). At present it is the only process-based model capable of simulating canopy-scale chlorophyll fluorescence. SCOPE incorporates current understanding of chlorophyll fluorescence processes including canopy radiative transfer, re-absorption of fluorescence within the canopy, and the non-linear relationship between chlorophyll fluorescence quantum yield and other quenching processes (Van der Tol et al., 2009, 2014). Leaf level chlorophyll fluorescence

Table 1. PFTs defined in BETHY and their abbreviations.

PFT #	PFT Name	Abbreviation
1	Tropical broadleaved evergreen tree	TrEv
2	Tropical broadleaved deciduous tree	TrDec
3	Temperate broadleaved evergreen tree	TmpEv
4	Temperate broadleaved deciduous tree	TmpDec
5	Evergreen coniferous tree	EvCn
6	Deciduous coniferous tree	DecCn
7	Evergreen shrub	EvShr
8	Deciduous shrub	DecShr
9	C3 grass	C3Gr
10	C4 grass	C4Gr
11	Tundra vegetation	Tund
12	Swamp vegetation	Wetl
13	Crops	Crop

is coupled to the commonly used Farquhar and Collatz models for C3 and C4 photosynthesis, respectively (Van der Tol et al., 2009). A current limitation of SCOPE is that there is no link between leaf level biochemistry and soil moisture. This is compensated by changes in LAI as provided by BETHY.

The canopy radiative transfer and photosynthesis schemes of BETHY have been replaced by the corresponding schemes in SCOPE, including the components required for calculation of chlorophyll fluorescence at leaf and canopy scales. The spatial resolution, vegetation (PFT) characteristics, leaf growth, and carbon balance are handled by BETHY. SCOPE therefore takes in climate forcing (meteorological and radiation data) and LAI from BETHY, and returns GPP. BETHY calculates the canopy water balance, leaf growth, and net carbon fluxes, which will prove useful in future when assimilating other data streams (e.g. atmospheric CO_2 concentration). Importantly, SCOPE provides a process-based link between SIF and GPP allowing the transfer of information from observations of SIF to simulated GPP. Subsequently, information from SIF may also be transferred to carbon fluxes resulting from GPP such as net ecosystem productivity.

2.2 Model Process Parameters

In this error propagation system, ~~model process parameter uncertainties are the quantities which SIF constrains~~information content from SIF observations are used to constrain the uncertainty in model process parameters. Parameters can be either global or differentiated by PFT. Global parameters apply to plants or soils everywhere while PFT-dependent parameters enable differentiation between physiological and leaf growth traits. Some key parameters for this study such as the maximum carboxylation capacity (V_{cmax}) and chlorophyll ~~a/b~~ a/b content (C_{ab}) are considered PFT-dependent. From an ecophysiological perspective, there are other parameters ~~from~~ specific to SCOPE that may be considered PFT-dependent such as the vegetation height and leaf angle distribution parameters. However, we have assumed them to be global to simplify the problem. GPP is relatively insensitive to these parameters, so this is not expected to impact ~~results. In future~~ the GPP uncertainty reduction results. Despite this, in a full assimilation with the SIF data it may be ~~worthwhile considering them~~ necessary to make these PFT-dependent ~~anyhow~~to improve the model-observed fit.

We expose 72 parameters from BETHY-SCOPE to the error propagation system (see Table A1). As stated above, each of these is represented by its PDF, assumed to be Gaussian. The mean and standard deviation for the prior parameters is shown Table A1. Choice of the prior mean and uncertainty for parameters follow those used in previous studies (Kaminski et al., 2012; Knorr et al., 2010; Koffi et al., 2015). For new parameters that are not well characterized (e.g. SCOPE parameters) we assign relatively large prior uncertainties, and mean values in line with the default SCOPE parameters and with Koffi et al. (2015). The choice of the prior may be considered important here considering we are using a linear approximation of the model around x_0 and that the model is known to be non-linear. Therefore, sensitivities can differ depending upon the choice of x_0 (Koffi et al., 2015).

There are seven SCOPE parameters exposed, one of which is PFT-dependent. These parameters were chosen due to their importance in simulating SIF or GPP, and to sensitivity tests such as those performed by Verrelst et al. (2015). They include ~~chlorophyll a/b content (~~C_{ab}~~)~~, leaf dry matter content (C_{dm}), leaf senescent material fraction (C_s), two leaf distribution function parameters ($LIDFa$, $LIDFb$), vegetation height (hc) and leaf width. V_{cmax} is a parameter shared by BETHY and SCOPE.

2.3 Uncertainty Calculations

~~Past studies utilizing the CCDAS framework have typically formulated the assimilation problem into two stages: model calibration and diagnostic or prognostic simulations (Kaminski et al., 2013) . Diagnostic refers to simulations over the calibration period (e.g. Rayner et al., 2005) , and prognostic outside of the calibration period (e.g. Scholze et al., 2007) . As we are investigating the usefulness of a relatively new observation with a newly coupled model we conduct an investigation into the potential level of calibration and constraint SIF can provide. We therefore formulate the problem into two slightly different stages; (i) optimization of parameter uncertainties and; (ii) projection of uncertainties onto uncertainty in diagnostic GPP.This means that we optimize model parameters and simulated GPP only for their uncertainty and not the absolute quantities. This process of assessing the information content of the SIF observations in uncertainty space is a useful first step towards a full assimilation of the data, allowing us to evaluate the level of constraint SIF is expected to impose on GPP and how that constraint is propagated through the model.~~

To calculate the uncertainty in parameter values following the constraint provided by the observational information of SIF (i.e. the posterior uncertainty) we propagate uncertainty from the observations onto the parameters. In order to perform this, we utilize a probabilistic framework where the state of information on parameters and observations is expressed by their corresponding probability density functions (PDF) ~~. Parameters and observations are therefore described~~ (see Tarantola, 2005) . The probability density of the errors in these quantities is assumed to be Gaussian, thus they are describable by their mean and uncertainty~~, and treated as Gaussian. Their respective means~~. The prior information on parameters is quantified by a PDF in parameter space and the observational information by a PDF in observational space. The mean values for the parameters and observations are denoted by x and d, ~~and their respective covariance matrices~~ respectively. The uncertainty covariance matrices in parameter space and observational space are denoted by C_x and C_d~~.~~, respectively. Formally, C_d represents the errors in the observations and in the model simulated counterpart (i.e. model error) (Scholze et al., 2016) . We only consider the contribution of observational errors to C_d, however we also perform a sensitivity test to investigate the effect of incorporating similar structural uncertainties described further below.

For linear and weakly non-linear problems we can assume that Gaussian probability densities propagate forward through to Gaussian distributed simulated quantities (Tarantola, 2005). This ~~allows~~ permits linear error propagation from the input parameters to the model outputs. Estimating posterior uncertainties of the parameters for these types of problems can therefore be performed independently of the parameter estimation~~. Therefore, we can calculate the expected posterior covariance matrix following constraint by observations using equation 1~~, in other words without the need to constrain ~~parameters mean values . This approximation~~ the mean values of the parameters (Kaminski et al., 2010, 2012) . This requires a matrix of partial derivatives of a target quantity with respect to its variables, also called a Jacobian matrix (H)~~, for which we calculate around our prior parameter values (x_0)~~. This matrix represents the sensitivity of a simulated quantity (e.g. SIF~~or~~, GPP) to the parameters. With ~~this assumption of linearity the choice of x_0 can influence the results considering it determines the point in model space where~~ the linear approximation, H ~~is calculated.Use~~ is calculated around the prior parameter values (x_0). This simplification of the model sensitivity brings limitations to the accuracy of the method. However, with the aggregation of subgrid variability across a model grid cell, sudden shifts in model sensitivity (e.g. step functions) are less likely or realistic; the present model incorporates these effects (Knorr et al., 2010) . Additionally, because the parameter space can be very large, the use of prior knowledge on ~~process parameter values helps~~ x_0 helps to limit the effect of this problem as H at x_0 likely provides a decent approximation of the true H that would occur at the global optimum (Tarantola, 2005) . The simplification is also useful considering the high computational cost of calculating H.

~~The first step is to utilize uncertainty in the observations to constrain the uncertainty in the process parameters. The information content of the observations and the parameters can be expressed by their respective inverse covariance matrices, C_d^{-1} and $C_{x_0}^{-1}$. Constraint of $C_{x_0}^{-1}$ using the observations and the Jacobian matrix allows us to~~ To calculate the posterior parameter covariance matrix ($C_{x_{post}}$) ~~as in equation 1~~ following constraint by observational information, C_d, we use Eq. 1 (Tarantola, 2005) .

$$C_{x_{post}}^{-1} = C_{x_0}^{-1} + H^T C_d^{-1} H \tag{1}$$

Where ~~$C_{x_{post}}$ expresses the posterior parameter covariance in matrix form, while~~ H expresses the Jacobian for SIF and H^T the Jacobian transposed. Comparing parameter uncertainties in the prior (C_{x_0}) and the posterior ($C_{x_{post}}$) allows us to quantify the improvement in parameter precision following the observational constraint. The parameter uncertainties in C_{x_0} and $C_{x_{post}}$ may be expressed as standard deviations (σ) by calculating the square root of their diagonal elements. We can therefore assess the relative uncertainty reduction in parameter following SIF constraint, or 'effective constraint', with 1 - ($\sigma_{posterior}/\sigma_{prior}$). This quantifies the effective constraint of the prior uncertainty and may be represented as a percentage decrease in σ uncertainty.

The observational constraint introduces ~~some~~ correlations into the posterior parameter distributions, thus posterior parameter uncertainties are not wholly independent. Strong correlations in $C_{x_{post}}$ indicate parameters that cannot be resolved independently in an assimilation~~while~~ , however their linear combinations can be. If large enough, these correlations can contribute significantly to the overall constraint of the target quantity (Bodman, 2013). We calculate correlations in parameters by expressing the covariances as correlations as in ~~equation 2 (see Tarantola, 2005, pg.71)~~ Eq. 2 (see Tarantola, 2005, p.71) . As a result, diagonal elements have a correlation equal to one while off-diagonals elements can range between -1 and 1.

$$R_{i,j} = \frac{C_{i,j}}{\sqrt{C_{i,i}}\sqrt{C_{j,j}}} \tag{2}$$

Using the parameter covariance matrix we can assess how parameter uncertainties ~~project~~ propagate forward through the model onto uncertainty in GPP using the Jacobian rule ~~for~~ of probabilities, the same method outlined in Rayner et al. (2005). This is the second stage of our error propagation ~~problem~~study. Using C_{x_0} we estimate the prior uncertainty in a vector of simulated target quantities (i.e. GPP). Similarly, using $C_{x_{post}}$ we estimate the posterior uncertainty in a vector of simulated target quantities. We calculate the uncertainty covariance of GPP ~~using equation~~ (C_{GPP}) using Eq. 3.

$$C_{GPP} = H_{GPP} C_x H_{GPP}^T \tag{3}$$

Where H_{GPP} is the Jacobian matrix of GPP with respect to the parameters. With this we can quantify the improvement in precision of simulated GPP by using either C_{x_0} or $C_{x_{post}}$ in ~~equation~~ Eq. 3. Therefore, using the forward model, a statistical estimation scheme and a set of observational uncertainties we can assess the information content of ~~those~~ the SIF observations in the context of the model, ~~it's~~ its parameter set, and ~~the quantity of interest~~ simulated GPP taking explicit consideration of uncertainties.

2.4 Uncertainty in Observations and Model Forcing Variables

~~Observational uncertainties in SIF~~ The observational uncertainty is a critical component in assessing the potential impact of an observing system on the estimation of carbon fluxes. Observational uncertainties of SIF used here are calculated from the GOSAT satellite observations for 2010. ~~These are interpolated to the model grid resolution as demonstrated below. To get the variance of a target grid cell at the model grid resolution (ylat,xlon)we first determine the area-weighted variance of each GOSAT grid cell (ilat, jlon) within that target grid cell. The area-weighting per GOSAT grid cell~~ ~~($\hat{Area}_{ilat,jlon}$)~~ ~~isealculated as the area divided by the total area of the target grid cell. This enables us to account for different grid cell sizes considering SIF is in physical units per unit area. We then sum the area-weighted variances and scale this uncertainty~~ This data is obtained from the ACOS (Atmospheric CO_2 Observations from Space) project at a grid resolution of 3° × 3°. As the model simulations are performed on a low-resolution grid (7.5° × 10°), we aggregate the observational uncertainties to this resolution using Eq. 4 as described below.

We assume the observations are independent and have uncorrelated errors, that is, they are distributed randomly. Assuming uncorrelated errors is, however, likely to overestimate the information content particularly if using the standard error as the uncertainty. Although it has been used in recent studies with satellite SIF (e.g. Parazoo et al., 2014) , the standard error is likely to be an overly optimistic approximation of the information content. For this study, we take a slightly conservative approach, scaling the calculated standard error by the square root of two ~~(see equation 4). Scaling the uncertainty in this way~~ as shown in Eq. 4. This effectively doubles the variance in an independent dimension and reduces the information content to compensate for the assumption of uncorrelated errors.

$$\sigma^2_{ylat,xlon} = \sqrt{2} \sum (\hat{Area}^2_{ilat,jlon} \cdot \sigma^2_{ilat,jlon})$$

~~We assume uncorrelated uncertainties in the observations. To ensure our results are not sensitive to the method of calculating the observational uncertainty we conduct some simple sensitivity tests with varied observational uncertainties within~~ Through aggregation of GOSAT grid cells to the model grid resolution the number of independent observations is reduced. To account for this the uncertainty in a given model grid cell is, approximately, divided by the square root of the ~~range expected.~~ number of GOSAT grid cells with SIF data that fall within that model grid cell (N). More precisely, we apply an area-weighting term in the equation (see Supplementary material Eq. A1). This has the effect of scaling the uncertainty by the $1/\sqrt{N}$ law, but takes into account the fact that SIF is in physical units per units area (i.e. $W\ m^{-2}\ \mu m^{-1}\ sr^{-1}$) and that grid cells have different areas over different latitudes. A full description of this calculation and detailed example is shown in supplementary material.

~~With the use of low-resolution observations the constraint of parameter uncertainties is actually underestimated. This is expected as with a high-resolution setup the number of observations will increase while the number of parameters will remain constant, resulting in stronger uncertainty reductions. Considering this , we also approximate the expected parameter uncertainty reductions from higher resolution observations(2°x 2°). Compared with~~ Therefore, the calculation of the observational uncertainties of SIF used here is approximated by Eq. 4 (for further details see Supplementary material Section A2). For a given model grid cell, the variance (σ^2) is approximately equal to the sum of the ~~low-resolution grid used here (7.5°x 10°) , a 2°x 2°grid has approximately 19-times more observations~~ standard error of each individual GOSAT grid cell (σ_i) squared, then scaled by the number of individual GOSAT grid cells with data and the square root of two.

$$\sigma^2 = \sqrt{2} \left[\frac{1}{\sqrt{N}} \sum_i \sigma_i^2 \right] \tag{4}$$

The resulting annual observational uncertainties, shown in Figure 3, appear to be much smaller than the uncertainties of individual GOSAT grid cells. In part this is due to the aggregation of multiple independent observations. ~~Therefore, to estimate the high-resolution observational uncertainty we divide the low-resolution GOSAT observational uncertaintiesby $\sqrt{19}$. Using~~

~~this scaled observational uncertainty estimate, parameter uncertainty and uncertainty in GPP is estimated.~~ Regions with more soundings across the year (e.g. the tropics) will also have smaller annual uncertainties.

Where feasible, systematic uncertainties in the SIF observations should also be considered in error propagation analyses. While systematic errors in the model cannot be assessed prior to a full assimilation of the data (Kuppel et al., 2013) , systematic errors in the observations can be. To incorporate this into our analysis, we investigate one source of structural uncertainty due to potential errors in the zero-level offset. The zero-level offset correction is done to prevent biases in the SIF retrieval (Frankenberg et al., 2011a) . Based on previous analyses, systematic uncertainties in the SIF retrieval may be considered small (Frankenberg et al., 2011a, 2014) . Here, we provide a more detailed assessment and characterization of the in-orbit systematic uncertainties. This is performed by assessing zero-level offset corrected GOSAT SIF soundings over the non-fluorescent regions of Antarctica and central Greenland during January and July, respectively. Systematic errors appear quite small ($\pm$ $0.06\ W\ m^{-2}\ \mu m^{-1}\ sr^{-1}$) (see Appendix Figure 10) and may vary seasonally. We therefore assess the effect of a conservative systematic random error of size $\pm$ $0.1\ W\ m^{-2}\ \mu m^{-1}\ sr^{-1}$ in the zero-level offset seasonally. This provides a sensitivity test of incorporating this systematic uncertainty into the error propagation system.

An additional source of uncertainty in model estimates of GPP is climate forcing. As mentioned by Koffi et al. (2015), while uncertainty in forcing such as incoming radiation is not considered in the current CCDAS setup~~. As might be expected however~~, it is considered to be an important variable in driving SIF (Verrelst et al., 2015) and GPP (Farquhar et al., 1980) . Without consideration of uncertainties in forcing variables the uncertainty in GPP may be underestimated. Studies that use process-based models or empirically-derived relationships do not explicitly consider such uncertainties (e.g. Beer et al., 2010). One such forcing variable is downward shortwave radiation (SWRad). Monthly means of SWRad are suggested to have ~~an uncertainty~~ a random error of 12 Wm^{-2} due mostly to uncertainty in clouds and aerosols (Kato et al., 2012). We therefore investigate SWRad uncertainty may be considered in GPP estimates. Furthermore, as SIF responds strongly to SWRad, there is the potential to utilize SIF observations as a constraint on the uncertainty of the forcing. We therefore conduct an additional experiment that incorporates the uncertainty in SWRad in the error propagation system. For this experiment an additional parameter representing SWRad is added to the inversion, which acts as a scaling factor for SWRad globally. We investigate the level of constraint SIF provides on this scaling factor, and the subsequent effects of incorporating uncertainty in SWRad in this inversion on uncertainty in GPP.

2.5 Model and Data Setup

In this study BETHY-SCOPE is run for the year 2010 on ~~a computationally fast~~the computationally efficient, low-resolution ~~grid scale~~ spatial grid (7.5° × 10°)~~, sufficient for investigating error propagation~~ . As the dynamical equations are the same for either low-resolution or high-resolution scales, use of the low-resolution setup is appropriate for an error propagation study as long as careful consideration is taken with observational uncertainties. Climate forcing in the form of daily meteorological input fields for running the model (precipitation, minimum and maximum temperatures, and incoming solar radiation) were obtained from the WATCH/ERA Interim data set (WFDEI Weedon et al., 2014). Photosynthesis and fluorescence are simulated at an hourly time step but forced by the respective monthly mean diurnal cycle. Leaf growth and hydrology are simulated daily.

SIF is simulated at 755 nm, the wavelength corresponding to the GOSAT retrieval frequency and near to the OCO-2 retrieval frequency (757 nm). We focus upon the constraint by SIF measurements at 1:00 p.m. local time as it closely corresponds to the local overpass time of the SIF-observing satellites GOSAT and OCO-2. However, we also investigate the effect of using alternative SIF-observing times (e.g. the GOME-2 satellite overpass time) and multiple observing times simultaneously on the constraint of GPP. ~~SIF is simulated at 755 nm, the wavelength corresponding to the GOSAT retrieval frequency and near to the OCO-2 retrieval frequency (757 nm).~~

3 Results

3.1 Parameter Uncertainties

A key metric for assessing the relative uncertainty reduction, or 'effective constraint', is defined as 1-($\sigma_{posterior}/\sigma_{prior}$~~. The~~). The effective constraint for all 72 parameters following constraint by SIF is shown in Figure 1 and in Table A1. We define weak, moderate and strong effective constraint as the relative uncertainty reduction from 1-10%, 10-50%, and >50%, respectively.

Parameters describing leaf composition (C_{ab}, C_{dm}, C_{sm}) generally achieve strong effective constraint from SIF. For eleven of the thirteen C_{ab} parameters the uncertainty is strongly constrained, between about 50% and 85%. SIF is highly sensitive to C_{ab} and we assign a relatively large prior uncertainty on these parameters, so ~~a~~ considerable constraint is expected. For the tropical broadleaved evergreen tree PFT however, the effective constraint on C_{ab} is much lower at 8%. For other leaf composition parameters C_{dm} and C_{sm} SIF effectively constrains the uncertainty by 2% and <1% respectively.

Varied effective constraint is seen for the leaf growth parameters (parameters 37-53 in Table A1) that control phenology and leaf area. Four out of the seventeen leaf growth parameters exhibit strong uncertainty reductions. These parameters pertain to a variety of processes including the temperature at leaf onset, day length at leaf shedding, leaf longevity, and the expected length of dry spell before leaf shedding (τ_W) (see Table ~~1~~A1). The parameter τ_W is important in controlling leaf area and it sees strong effective constraint from SIF, from 44-69% depending upon which class of PFT it pertains to. For the parameters that are PFT-specific, there is generally a larger constraint seen when they relate to the C3Gr, C4Gr and crops. For example, uncertainty in τ_W for grasses and crops (τ_W^{Gr}) is effectively constrained by 70%.

Leaf physiological parameters (parameters 1-36 in Table A1) see a weak to moderate level of effective constraint. Of particular importance for simulating GPP is the PFT-specific parameter V_{cmax}. ~~Constraint~~ Effective constraint on V_{cmax} varies from <1% up to 30% depending upon the PFT of interest. Six PFTs that, combined, represent about 70% of the land surface have their V_{cmax} parameters constrained by >10%. Some global physiological parameters receive a weak constraint from SIF. However, for most of these (E_{R_d}, E_{K_O}, E_k, K_O, $a_{J,V}$) there is only a minor reduction of uncertainty as SIF is weakly sensitive to them. The parameters that do see a weak effective constraint include the Michaelis-Menton enzyme kinetics constant for carboxylation (K_C; 2.3%), the corresponding activation energy for carboxylation (E_{K_C}; 2.5%) and for V_{cmax} ($E_{V_{max}}$; 7.6%), and quantum efficiency parameters (α_q; 5.2% and α_i; 1.1%).

Global canopy structure parameters (parameters 69-72 in Table A1) also see a weak to moderate constraint from SIF. In particular the structural parameters $LIDFa$ and $LIDFb$ see their uncertainty reduced by 23% and 16%, respectively. The

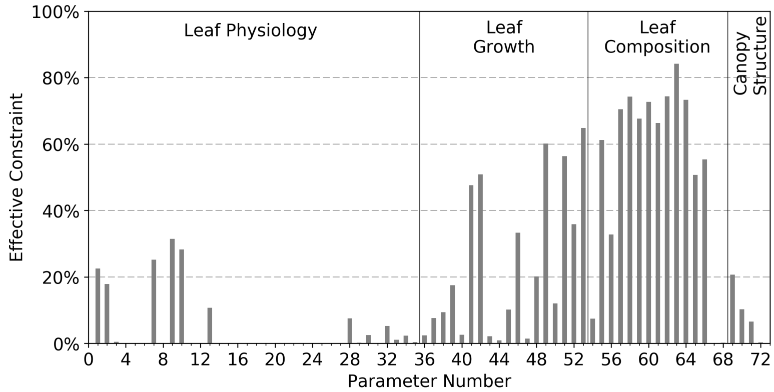

Figure 1. ~~Relative reduction in uncertainty (1σ) for~~ Effective constraint of BETHY-SCOPE model process parameters ~~following~~ from SIF ~~constraint~~observations. ~~The parameters~~ Only the parameter numbers are ~~defined in~~ given, for the corresponding descriptions see Table A1.

parameters for vegetation height and leaf width, which are used to calculate the fluorescence "hot-spot" variable (see Van der Tol et al., 2009), are effectively constrained by 9% and <1%, respectively.

Parameters that pertain to more dominant PFTs in terms of land surface coverage (e.g. C3 grass) tend to see stronger uncertainty reductions. This is due to them being exposed to more SIF observations.

~~As expected, with high-resolution observations there is stronger constraint of parameter uncertainties (see Table A1). Strong constraint is seen for 27 parameters compared to 16 in the low-resolution tests. This includes strong constraint of five~~ $\cancel{V_{cmax}}$ ~~parameters, and moderate constraint of seven other physiological parameters.~~

With the observational constraint correlations are introduced into the posterior parameter distributions. We assess these correlations using 2, shown in Figure 2. We find strong (R $\geq$ 0.5) positive correlations between nine of the PFT-specific ~~chlorophyll~~ C_{ab} parameters. These are also negatively correlated the leaf angle distribution parameter $LIDFa$. Thus~~during the assimilation~~, during a full assimilation with SIF data only the sum of C_{ab} and $LIDFa$ can be resolved, not their individual values. ~~Some~~ Two leaf growth parameters are also ~~significantly correlated, including~~ strongly correlated, T_ϕ with T_r ~~and~~ $\cancel{\xi}$ ~~with~~ $\cancel{\tilde{A}}$. Smaller correlations are also present between the subset of parameters shown in Figure 2.

To assess the effect of incorporating a systematic error from the observations into this analysis we apply a seasonal, systematic random σ error of 0.1 $W\ m^{-2}\ \mu m^{-1}\ sr^{-1}$. This is incorporated as four additional parameters, one for each season, that scale the SIF signal across the globe. We find that the inclusion of this systematic error has a negligible effect on posterior uncertainties of the parameters. The difference in effective constraint between this sensitivity test case and the standard case above is <1% for any given parameter.

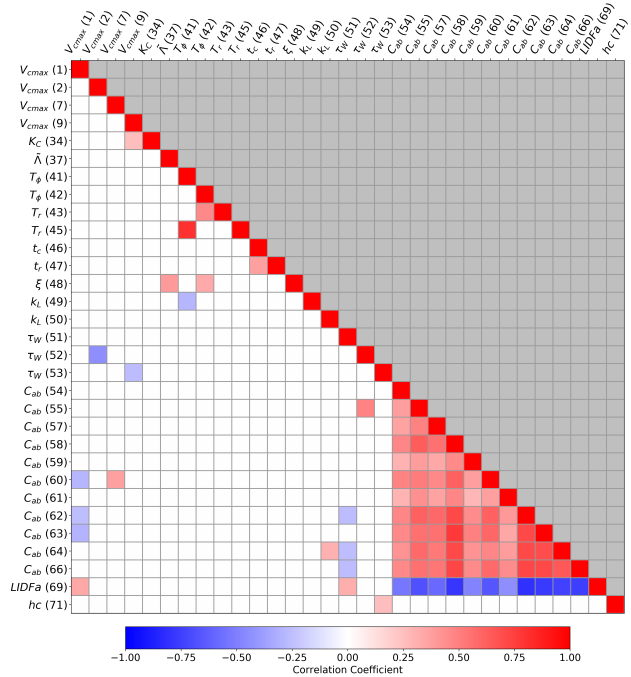

Figure 2. ~~Correlations~~ Correlation coefficients (r-value) in the posterior parameter covariance matrix ($C_{x_{post}}$). This shows the magnitude and sign of correlations in posterior parameter uncertainties following constraint with SIF data. Only ~~values~~ parameters with an absolute correlation coefficient >~~0.1~~ 0.25 with one or more other parameters are shown. Values above and below the diagonal are identical, therefore those above are coloured grey. The axes labels show the parameter symbol and number as defined in Table A1.

3.2 Uncertainty in GPP

To assess the constraint imposed by SIF on simulated GPP we compare the prior and posterior uncertainty in GPP as calculated using ~~equation~~ Eq. 3. Similar to the assessment of parameter uncertainty reductions, to assess the effective constraint of SIF on GPP we use a metric that measures the relative uncertainty reduction in ~~1~~σ from the prior to the posterior.

Utilizing SIF observations at 1:00 p.m. results in the uncertainty in global annual GPP to decrease from 13.0 PgCyr^{-1} to 2.8 PgCyr^{-1}, constituting a 79% reduction of the prior uncertainty. Spatially, the prior uncertainty in GPP varies across the globe, with particularly large uncertainties in regions with high productivity as might be expected (Figure 4). In the posterior, it is clear that uncertainty in GPP is strongly reduced across the globe (Figure 5). The relative uncertainty reduction (Figure 6) appears to show smaller constraint of uncertainty in the boreal regions, however this is because prior uncertainty is already relatively low (Figure 4). As with the parameter uncertainty reductions, we expect that with the use of higher resolution observations there will be stronger constraint of the uncertainty. When utilizing the high-resolution (2°x 2°) observational uncertainties, uncertainty in global GPP is reduced to 1.3 PgCyr^{-1}, constituting a 90% reduction in uncertainty relative to the prior.

To assess which parameters contribute to the uncertainty in GPP for the prior and posterior, we can conduct linear analysis of the uncertainty contributions. Typically this technique can only be used for the prior as the correlations in posterior parameter uncertainties, excluded from the linear analysis, also contribute toward the overall constraint. However, we can assess the contribution of these correlations to the constraint of GPP by setting the off-diagonal elements in $C_{x_{post}}$ to zero and using it in Eq. 3; the difference between this and the standard case that uses the full $C_{x_{post}}$ equates to the contribution of correlations. We find that the contribution of these correlations to the constraint of GPP is small ($0.12\ \mathrm{PgCyr}^{-1}$ or <1%), thus we can assume the linear analysis technique holds for the posterior as well. This finding is supported by the correlation analysis in posterior parameter uncertainties which showed few significant correlations in parameters relevant for GPP. This result is encouraging as it indicates that the parameters in a SIF assimilation system contributing most to the constraint of GPP are capable of being resolved independently.

Using linear analysis of the uncertainty we find that uncertainty in global annual GPP in the prior and posterior ~~stem~~ stems from different processes. For the prior we see that the uncertainty in GPP is dominated, at 91%, by parameters describing leaf growth processes. Of these, a single parameter, τ_W for C3 grass, C4 grass and crops (τ_W^{Gr}) makes up 82% of the uncertainty in global annual GPP. Parameters representing physiological processes account for about 6% of prior uncertainty, most of which stem from the V_{cmax} parameters. Parameters for C_{ab} only account for 2.5% of the uncertainty~~, as may be expected considering GPP is relatively insensitive to C_{ab}~~.

For the posterior, ~~following an overall reduction of the~~ which has a lower overall uncertainty in GPP, uncertainty is dominated by parameters representing physiological processes. Physiological parameters account for 53% of the uncertainty in posterior annual GPP, with V_{cmax} parameters alone accounting for 40%. The relative contribution by leaf growth parameters is reduced to 45%, and for τ_W^{Gr} to 25%. For C_{ab} the relative contribution is smaller than the prior at 1.3%. This shift in which parameters contribute to the relative uncertainty in GPP between the prior and the posterior demonstrates how effectively SIF constrains leaf growth processes. ~~Although there are uncertainty reductions~~ Uncertainties in physiological parameters ~~, the increase in the relative uncertainty contribution of these processes in the posterior GPP demonstrates the limitations in SIF constraining leaf physiology~~are constrained less than the leaf growth parameters which results in them contributing more in relative terms to the posterior uncertainty of GPP.

Regionally, we split the land into three regions, the Boreal region (above 45° North), the Temperate North (30° to 45° North) and the Tropics (30° South to 30° North). SIF constraint on annual GPP varies substantially across different regions of the globe, with relative uncertainty ~~reductions~~ reduction in of 48%, 82%, and 79% for the Boreal, Temperate North and Tropics regions, respectively. ~~For the~~ In Figure 7 we show the contribution of parameter classes (leaf physiology, leaf growth, leaf composition and canopy structure; see Table A1 for details) to the parametric uncertainty of GPP across the year for each of these regions. From Figure 7 it can be seen that the Boreal and Temperate North regions ~~there are also~~ exhibit seasonal differences in total uncertainty and in the constraint SIF provides. This is caused by seasonal dependencies in the sensitivity of SIF and GPP to certain processes (e.g. leaf development versus leaf senescence) as well as seasonal differences in the density of observations in ~~this region~~these regions. There are far fewer GOSAT satellite observations during Boreal autumn and winter, thus there are fewer observations to constrain processes controlling GPP during this time.

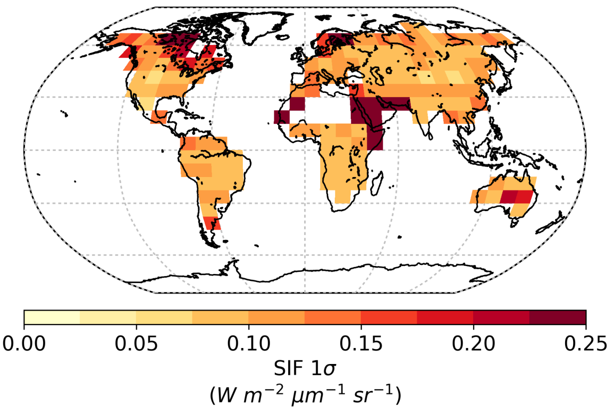

Figure 3. ~~SIF~~ Annual observational uncertainty ~~(1σ)~~ of SIF interpolated from GOSAT ~~SIF~~ observations for 2010.

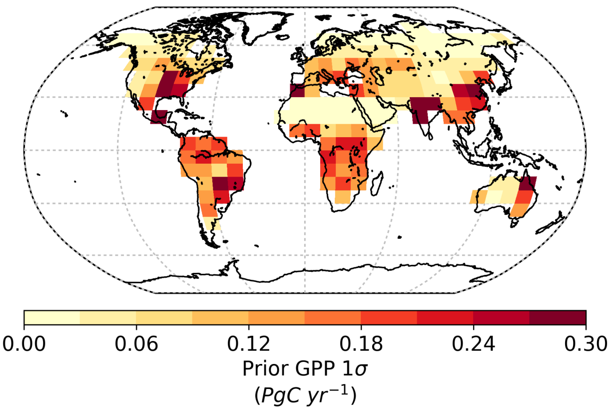

Figure 4. Prior parametric uncertainty in annual GPP~~(1σ)~~.

During the start of the growing season leaf physiology, in particular photosynthetic rate constants (V_{cmax}), play a larger role whereas later in the growing season during the warmest months leaf growth, via water limitation on leaf area (~~via~~ τ_W^{Gr}) of grasses, plays a larger role. Therefore in the Boreal region, where the strongest seasonality in constraint is seen, from July through to January SIF constrains GPP by >60%. Uncertainty in GPP during these months is dominated by the leaf growth parameters τ_W^{Gr} and k_L along with C_{ab} (for EvCn) all of which receive considerable constraint from SIF. From February to June however, SIF constrains GPP by less than 50%, as a large proportion of the uncertainty ~~is made up by the less constrained~~ arises from the less-constrained V_{cmax} parameters. Following SIF constraint, uncertainty in Boreal GPP stems mostly from uncertainty in ~~V_{cmax}~~ leaf physiology, particularly for the EvCn PFT. Similar differences between seasonal constraint is seen

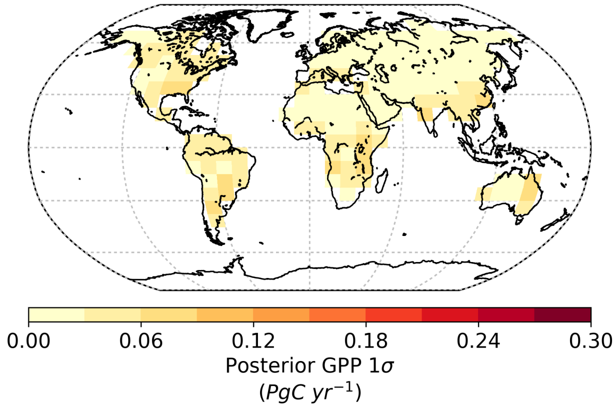

Figure 5. Posterior parametric uncertainty in annual GPP~~(1σ)~~.

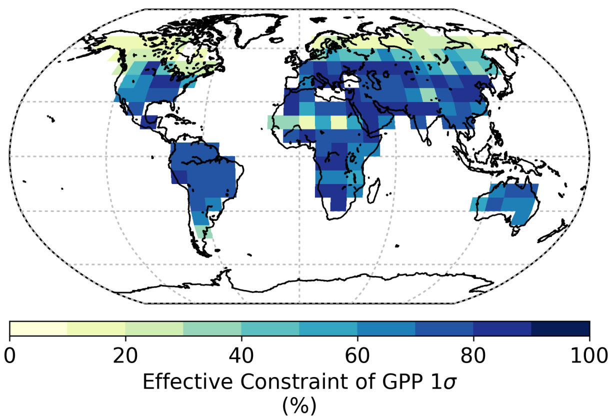

Figure 6. Relative uncertainty reduction (i.e. effective constraint) of parametric uncertainty in annual GPP from prior to posterior.

for the Temperate North, ~~albeit not as exaggerated with SIF-constraint ranging~~ although with a smaller seasonal variation in SIF constraint that ranges between 74% and 87% across the year.

For the Tropics uncertainty reduction in GPP is about 80% across the year. Uncertainty in the prior is dominated by the leaf growth parameters and in particular the τ_W parameters controlling water-limited leaf area. SIF constraint is primarily propagated through ~~these~~ the τ_W parameters onto GPP resulting in a well-constrained posterior with a ~~1~~σ uncertainty of 1.6 PgC yr-1 ~~on the Tropics annual GPP~~ in annual GPP of the Tropics. Although moderate constraint is seen in the key PFT-specific parameter V_{cmax} for the dominant ~~Tropical PFTs~~ tropical PFTs (see Figure 1), in the posterior these parameters ~~make up~~ contribute to roughly 35% of the uncertainty in annual GPP.

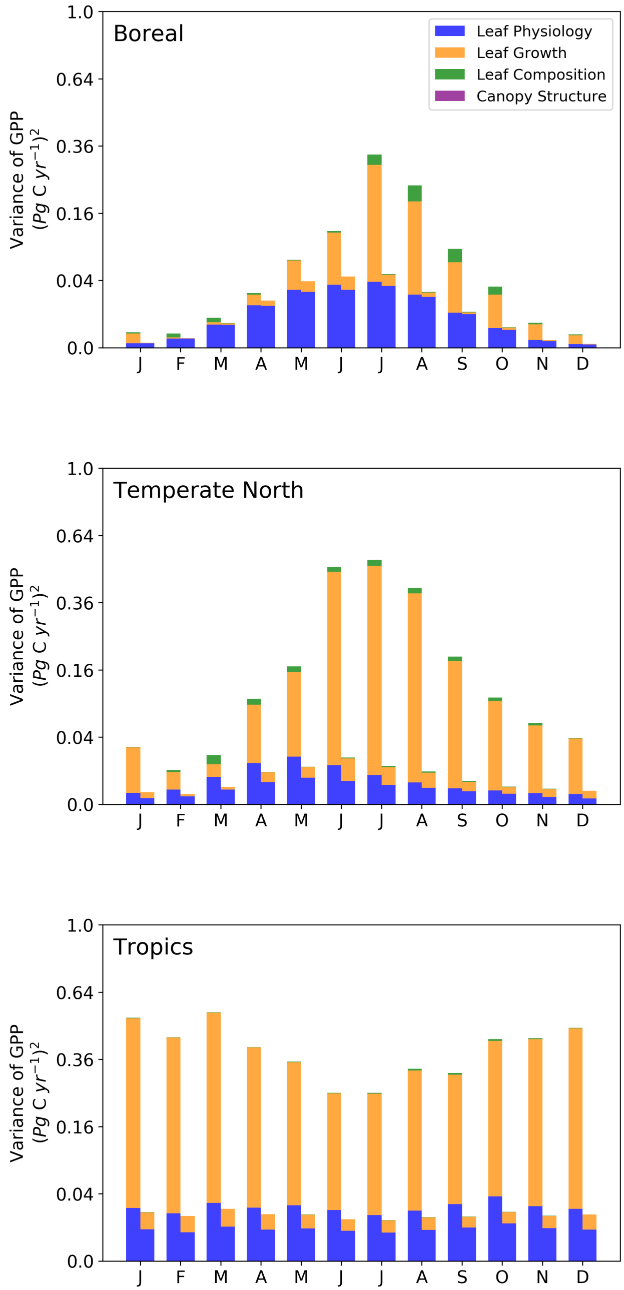

Figure 7. Contribution of parameter classes to parametric uncertainty in monthly GPP for three regions (see Table A1 for details on these parameter classes). For each month, the bar on the left is the prior and the bar on the right is the posterior. Uncertainties are represented as variances, thus the units are in $\mathrm{PgCyr^{-1}}$ squared and, for clarity, the y axes are on a quadratic-transformed scale.

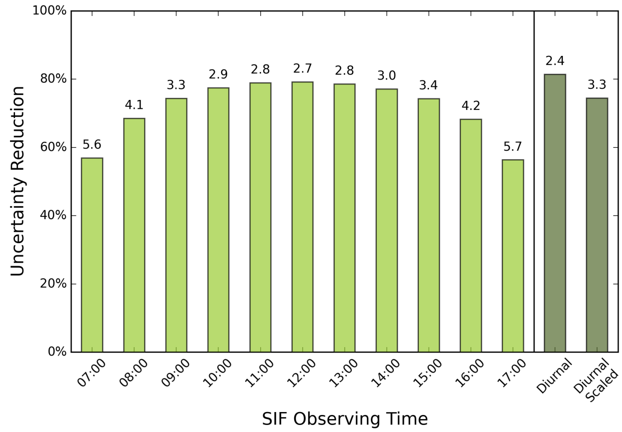

Figure 8. Relative uncertainty reduction in global annual GPP for different observing times and the two diurnal cycle configurations. Values at the top of the bars correspond to the posterior uncertainty ($\cancel{1\sigma}\sigma$) in global annual GPP.

3.3 Diurnal SIF Constraint

With this setup it is possible to test how the SIF-constraint on GPP might change with alternative observational times. Considering this, we test how the constraint on GPP changes when assimilating observations of SIF from alternative times of the day, assuming the same number of observations and the same observational uncertainty as used above. From this we see that different observing times yield differences in the posterior uncertainty and the relative constraint of GPP (see Figure 8). The constraint on global annual GPP when using SIF-observing times between 9:00 a.m. and 3:00 p.m. is quite similar, with the posterior uncertainty in global annual GPP ranging from 2.7 $\mathrm{PgCyr^{-1}}$ (constraint of 79%) to 3.4 $\mathrm{PgCyr^{-1}}$ (constraint of 74%). The most significant constraint on GPP is obtained when using SIF observations at between 11:00 or 13:00, nearest to the peak in the diurnal cycle of both GPP and SIF.

We also test the effect of utilizing SIF measurements at multiple times of the day simultaneously. We select the times 8:00 a.m., 12 noon, and 4:00 p.m., replicating a theoretical geostationary satellite. For this experiment we first test the effect of increasing the number of observations by a factor of three, assuming the same uncertainty for the three observation times. Second, we also increase the number of observations by a factor of three, but scale the variance of these observations by one third. Using this second test we can assess whether differences in parameter sensitivities of SIF and GPP at the different times of the day adds value in the overall constraint.

Using a diurnal cycle of observations results in a posterior uncertainty of 2.4 $\mathrm{PgCyr^{-1}}$, or a relative reduction of 81% as in Figure 8. This is an extra 2% constraint on the uncertainty in GPP compared with observations at 12:00 noon alone. If we use a diurnal cycle of observations with scaled uncertainties, we see a slightly reduced constraint on GPP where the posterior uncertainty is 3.3 $\mathrm{PgCyr^{-1}}$ equivalent to a 74% reduction in uncertainty (Figure 8). ~~Therefore the difference in model~~

~~sensitivities across the diurnal cycle is not sufficient to outweigh having additional observations at midday. The constraint is worse with these scaled observational uncertainties as we are effectively removing some useful observational information at midday, which is the most sensitive time of day, and getting extra observational information at the lower-sensitivity times of 8:00 a.m. and 4:00 p.m..~~

3.4 Incorporating Uncertainty in Radiation

In order to assess the effects of incorporating uncertainty in SWRad we conduct three experiments. First is a control run, equivalent to using SIF at 1:00 p.m. as before. Second includes uncertainty in SWRad by adding it into the posterior uncertainty calculation, what might be done normally when accounting for uncertainty in forcing. Third is incorporating uncertainty in SWRad into the error propagation system with SIF, ~~effectively treating it as a model parameter such that its~~ such that it's uncertainty may be constrained. This third experiment effectively treats SWRad as a model parameter by adding an extra row and column to C_x.

Including the uncertainty in SWRad in the calculation of posterior uncertainty in GPP results in an additional 0.02 PgCyr^{-1} to the prior uncertainty in global annual GPP. This is a small effect relative to the parametric uncertainties. Moreover, if we incorporate SWRad uncertainty into the error propagation system we see that this additional uncertainty is mitigated by the SIF constraint. With SWRad uncertainty included, the posterior uncertainty in GPP remains at 2.8 PgCyr^{-1}, equivalent to the case without accounting for uncertainty in SWRad, ~~with both providing about a~~ in both cases resulting in a relative reduction of the GPP uncertainty by 78.6~~% relative reduction in uncertainty of GPP~~%. This mitigation of the additional uncertainty from SWRad is possible because both SIF and GPP are strongly sensitive to it, thus any constraint on SWRad from SIF is also propagated through to GPP.

Table 2. ~~Experiments with~~ Parametric uncertainty and effective constraint for each of the SWDown ~~Uncertainty~~experiments. Prior and posterior values shown are the one standard deviation (σ) uncertainty in global annual GPP.

Experiment	Prior GPP ~~1σ~~ (PgCyr^{-1})	Posterior GPP ~~1σ~~ (PgCyr^{-1})	~~Relative Uncertainty Reduction~~ Effective Constraint
Control	13.04	2.79	78.59%
Control+SWRad	13.05	2.87	78.01%
With SIF Constraint	13.05	2.80	78.57%

By assessing the prior and posterior uncertainty in SWRad in $C_{x_{prior}}$ and $C_{x_{post}}$, respectively, we can assess the effective constraint following use of SIF in the error propagation system. We find that SIF constrains the SWRad uncertainty by about

28%. This gain in information on SWRad naturally results in less information being available for other parameters. The relative uncertainty reduction for most parameters decreases by just a few percent. For example most C_{ab} parameters see a decrease in constraint of around 1%, and five of the V_{cmax} parameters between 0.5-3%. With GPP exhibiting low sensitivity to C_{ab} parameters and strong sensitivity to SWRad, the transfer of information from C_{ab} to SWRad results in an overall mitigated effect of SWRad uncertainty on GPP.

4 Discussion

The results presented show that with one year of satellite SIF data observed at the GOSAT and OCO-2 satellite overpass time and SIF retrieval wavelength we can constrain a large portion of the BETHY-SCOPE parameter space and ultimately ~~constrain~~ yields a parametric uncertainty in global annual GPP ~~to~~ of $\pm$ 2.8 PgCyr^{-1}. This constitutes an uncertainty reduction of 79% in global annual GPP relative to the prior. Although this data-driven constraint is model dependent, it is much improved on the often-reported uncertainty of $\pm$ 8 PgCyr^{-1} from the empirical-model-based upscaled product of Beer et al. (2010).

We note that our analysis is an underestimate of the constraint, as it is performed with relatively low-resolution observations. ~~We demonstrated, however, that with~~ With the use of higher resolution observations ~~the constraint gets~~ such as those from OCO-2 the constraint will get stronger. Similarly, with a longer time-series of data there will be stronger constraint. This occurs because the number of observations increases while the number of parameters remain constant. ~~With an approximation of observational uncertainty in higher resolution observations(2°x 2°) we see that uncertainty in global annual GPP is reduced by 90% to 1.3 .~~ We also find that the effect of incorporating a systematic error in the observations, for example due to a seasonal error in the zero-level offset, is negligible on posterior parametric uncertainties. This may be negligible because, for a given season, this systematic uncertainty applies across all data points and may act to scale all SIF values and therefore the sensitivities as well. In any case, any systematic error in the zero-level offset corrected data assessed here (Appendix Figure 10) appears small.

The constraint on global GPP is similar when assimilating SIF at any time between 9:00 a.m. and 3:00 p.m.. Assimilating observations at the daily maximum of SIF and GPP provides the strongest constraint as both quantities exhibit the strongest parameter sensitivities at these times. Depending upon the state of the vegetation and the environmental stress conditions, maximum SIF and GPP may occur anywhere between mid-morning and early afternoon. Therefore, we expect that effective use of different satellite-retrieved SIF observations for assimilation studies will depend not so much on their observing time but more on the spatiotemporal resolution, measurement precision, and subsequent uncertainty.

A confounding factor to this expectation is the uncertain role of physiological stress on the diurnal cycle of SIF and GPP and subsequent modeling capabilities of these processes.

Multiple studies have shown that various forms of environmental stress result in downregulation of ~~photosystem II~~ PSII and changes in the fluorescence yield, particularly evident across the diurnal cycle (Carter et al., 2004; Daumard et al., 2010; Flexas et al., 1999, 2000, 2002; Freedman et al., 2002). By ingesting SIF observations at multiple times of the day we hypothesized that there could be improvements in the overall constraint on GPP as the SIF observations would capture the vegetation in different states of stress. ~~However, we~~ We saw only minor improvements in the constraint and less constraint if we assumed no additional information in the observations (i.e. with scaled uncertainty). Thus, the difference in model parameter sensitivities of SIF and GPP at other times ~~during~~ across the diurnal cycle were not sufficient to add value to the constraint. Additionally, the constraint is worse with these scaled observational uncertainties as we are effectively removing some useful observational information at midday, which is the most sensitive time of day, and getting extra observational information at the lower-sensitivity times of 8:00 a.m. and 4:00 p.m.. This is likely due to limitations of the model~~however~~. Although BETHY-SCOPE simulates light-induced downregulation of PSII, there is no mechanism present to simulate other forms of stress that might be expected to emerge across the diurnal cycle. However, even with a perfect model, the spatial footprint and spatiotemporal averaging of satellite observations may smooth over stress signals. Considering these factors, there is no technical reason, other than computational requirements, why a data assimilation system such as this could not ingest individual soundings of SIF observations to remedy the problem.

The constraint of SIF on GPP occurs via multiple processes including leaf growth, leaf composition, physiology, and canopy structure. For the prior, uncertainty in global GPP is dominated by leaf growth processes. There is a clear and direct link between leaf growth processes and GPP (Baldocchi, 2008) as the dynamics of leaf area influences canopy APAR which in turn strongly influences GPP. Leaf growth parameter uncertainties are relatively large in the prior, with coefficients of variation up to 50%. It is perhaps no surprise then that these parameters project a large uncertainty onto GPP. Regardless, both GPP and SIF respond similarly to the leaf growth parameters so information from observations of SIF can provide direct constraint on GPP in this way. Many leaf growth parameters, particularly for grasses, crops, and deciduous trees and shrubs, receive constraint of >40% from SIF thus the overall contribution of leaf growth parameters in the posterior is considerably reduced.

Of particular importance is the parameter describing water limitation on leaf growth (τ_W), which accounts for about 80% of the prior uncertainty in global GPP. ~~While~~ Model SIF and GPP are highly sensitive to this parameter hence there are large values in H and H_{GPP} pertaining to τ_W. This relates to the model formulation as many of the leaf growth parameters determine phenological processes such as temperature or light dependent growth triggers (i.e. temporal evolution of leaf area), while τ_W is the only ~~free~~ process parameter controlling leaf area other than intrinsic maximum LAI ($\tilde{\Lambda}$) (Knorr et al., 2010). ~~Considering this, and that~~ Additionally, as we assume little prior knowledge ~~on~~ for τ_W (i.e. it is highly uncertain) ~~,~~ it projects a ~~considerable~~ relatively large uncertainty onto GPP.

At the global scale, τ_W for crops, C3 grasses and C4 grasses (τ_W^{Gr}) is particularly important. ~~The prevalence of these three PFTs across all biomes means they can have strong effect at the global scale.~~ Combined, these three PFTs cover about 47% of the land surface and account for 58% of global annual GPP in the present model setup. Although this contribution to global GPP may seem high, it is based on the prior estimate. In a recent study by Scholze et al. (2016) where atmospheric CO_2 concentration and SMOS soil moisture were assimilated into BETHY, the posterior value for τ_W^{Gr} shifted approximately three standard deviations away from the prior, the result of which would have been a large change in the GPP of these PFTs. This exposes a limitation to the present study as we can predict and quantify how SIF will constrain the uncertainty of process parameters and GPP, but we cannot predict how their values will change.

The constraint SIF provides on leaf growth processes is also perhaps ~~feasible~~ achievable from other remote sensing products such as FAPAR (e.g. Kaminski et al., 2012). A direct comparative study would be required to assess the advantages and disadvantages of each observational constraint. Nevertheless, issues arise with these alternative observations when observing dense canopies (Yang et al., 2015) or vegetation with high photosynthetic rates such as crops as they are near saturation (Guanter et al., 2014). Information on maximum potential LAI ($\tilde{\Lambda}$) and parameters pertaining to understorey shrubs and grasses are therefore also limited (Knorr et al., 2010). A strong benefit of SIF is that it shows minimal saturation effects (e.g. Yang et al., 2015), especially beyond 700 nm where most current satellite SIF measurements are made.

The strong constraint SIF provides on leaf growth processes indicates that it is likely to provide improved monitoring of key phenological processes such as the timing of leaf onset, leaf sensescence and growing season length. This will be highly useful in interpreting results from a full assimilation with SIF as the posterior process parameter values can be compared with independent ecophysiological data, taking consideration of spatial scale issues.

Beyond observing LAI dynamics SIF can also provide critical insights into physiological processes (e.g. Walther et al., 2016). We see here that SIF provides weak to moderate constraint on a range of physiological parameters, including up to 30% constraint on V_{cmax} parameters. The limited constraint on these parameters results in the posterior being dominated by uncertainty in the parameters representing physiological processes. This is in line with Koffi et al. (2015) ~~considering they~~ who found limited sensitivity of simulated SIF to V_{cmax}. We note that under certain conditions, where other key variables are well known, SIF can be used to retrieve V_{cmax} (Zhang et al., 2014). The ability of SIF to inform on physiological processes at all will provide researchers with a powerful new insight into spatiotemporal patterns of GPP. As was shown by Walther et al. (2016) and Yang et al. (2015) this is particularly important for evergreen vegetation as changes in photosynthetic activity are not always reflected by changes in traditional vegetation indices.

Chlorophyll content here constitutes a classic nuisance variable. A nuisance variable is one that is not perfectly known, impacts the observations we wish to use but not the target variable (Rayner et al., 2005). However, exploiting the well-documented correlation between leaf nitrogen content, V_{cmax}, and C_{ab} may help curtail this problem (Evans, 1989; Kattge et al., 2009). Houborg et al. (2013) demonstrated that by including a semi-mechanistic relationship between these variables in the Community Land Model and using satellite-based estimates of chlorophyll to derive V_{cmax}, there is significant improvement in predictions of carbon fluxes over a field site. Implementing such a semi-mechanistic link in a data assimilation system would enable the strong constraint that SIF provides on C_{ab} to feed more directly onto GPP. However, in this study it is assumed C_{ab} and V_{cmax} can be resolved independently which may not be the case considering ecophysiological studies have shown the two parameters are commonly correlated.

Almost all terrestrial carbon cycle models use down-welling radiation at the Earth's surface as an input variable. Any uncertainty in this forcing will translate into uncertainty in carbon fluxes including GPP, and few studies consider such uncertainties. A known systematic error in forcing variables (e.g. Boilley and Wald, 2015) cannot be considered in the present error propagation system, however, in such a case a correction to the data should be performed as it will bias carbon flux estimates. For random errors that cannot be removed however, they may be considered in the uncertainty of carbon flux estimates using error propagation. At the global scale, Kato et al. (2012) used a perturbation study, along with modeled irradiance and remotely sensed measurements to compute ~~an uncertainty (1~~a random error (σ) of 12 Wm^{-2} for monthly gridded downward shortwave radiation over the land. We considered this uncertainty by incorporating it into the error propagation system with SIF. ~~By including this uncertainty as the prior uncertainty in GPPincreases, however, when incorporating it in with SIF the effect is mitigated~~While including this forcing uncertainty in the prior increases the prior uncertainty of GPP, incorporating the former in error propagation analysis with the SIF constraint mitigates the downstream effect on GPP. SIF can therefore provide useful information on the SWRad forcing via a data assimilation system. The consideration of uncertainties in forcing variables such as SWRad on terrestrial carbon fluxes is important when estimating the uncertainty in GPP. However, the effect on uncertainty in GPP may be strongly reduced by using SIF observations.

The results presented here demonstrate how SIF observations may be utilized to optimize a process-based terrestrial biosphere model and constrain uncertainty of simulated GPP. These results are, however, model dependent. The assumption is that the model simulates the most important processes driving SIF and GPP. Some key, remaining unknowns include how processes such as environmental stress, 3-dimensional canopy structure effects, or nitrogen cycling may affect the SIF signal. As better understanding is developed on the role that these processes play, modeling capabilities ~~can~~ will also be improved. Additionally, a different set of prior parameter values will alter the results due to changes in the Jacobian. Use of prior knowledge, based on ecophysiological data and its probable range, is critical to curtail this problem. The choice of how to spatially differentiate the parameters will also affect results (Ziehn et al., 2011). Selecting an optimal parameter set that has the fewest degrees of freedom, yet provides the best fit to the observational data is outside the scope of this study however. Implementation of a parameter estimation scheme in a full data assimilation system with SIF and other observational data will help address these challenges. Earlier work by Koffi et al. (2015) demonstrated that the model can simulate the patterns of observed satellite SIF quite well, indicating the model can ingest the data. Further work will be needed to assess how well the model can simulate patterns of SIF with an optimized, realistic parameter set.

5 Conclusions

We assessed the ability of satellite SIF observations to constrain uncertainty in model parameters and uncertainty in spatiotemporal patterns of simulated GPP using a process-based terrestrial biosphere model. The results show that there is strong constraint of parametric uncertainties across a wide range of processes including leaf growth dynamics and leaf physiology when assimilating just one year of SIF observations. Combined, the SIF constraint on parametric uncertainties propagates through to a strong reduction of uncertainty in GPP. The prior uncertainty in global annual GPP is reduced by 79% from 13.0 PgCyr^{-1} to 2.8 PgCyr^{-1}. Although model dependent, this result demonstrates the potential of SIF observations to improve our understanding of GPP. We also showed that a data assimilation framework with error propagation such as this allows us to account for uncertainty in model forcing such as SWRad. Surprisingly, by including it into this framework with SIF observations there is a net-zero effect on uncertainty in GPP due to the sensitivity of both SIF and GPP to radiation. This study is a crucial first step toward assimilating satellite SIF data to estimate spatiotemporal patterns of GPP. With the addition of other observational constraints such as atmospheric CO_2 concentration or soil moisture there is also the possibility of accurately disaggregating the net carbon flux into its component fluxes, GPP and ecosystem respiration. Indeed, with these additional, complementary observations of the terrestrial biosphere further constraint could be gained as other regions of parameter space ~~may~~ can be resolved (Scholze et al., 2016).

6 Code availability

The BETHY-SCOPE model code is available in the repository at https://github.com/NortonAlex/BETHY-SCOPE-Interactive-Phenology. The GOSAT satellite SIF data used in this paper is from the ACOS project (version b35).

Appendix A

A1 Model Process Parameters

Table A1. BETHY-SCOPE process parameters along with their prior and optimized uncertainties following SIF constraint, represented as one standard deviation. ~~Uncertainty~~ Relative uncertainty reduction (i.e. effective constraint) is reported for the error propagation with low-resolution and high-resolution SIF observations. Units are: V_{cmax}, μmol(CO_2) m^{-2} s^{-1}; $a_{J,V}$, dimensionless ratio; activation energies E, J mol^{-1}; $\tilde{\Lambda}$, m^2 m^{-2}; T_ϕ, °C; T_r, °C; t_c, hours; t_r, hours; ξ, d^{-1}; k_L, d^{-1}; τ_W, days; C_{ab}, μg cm^{-2}; C_{dm}, g cm^{-2}; C_{sm}, dimensionless fraction; hc, m; leaf width, m.

Class	#	Description	Parameter	Prior Mean	Prior Uncertainty	~~Uncertainty Reduction~~ Effective Constraint (%)~~Low-Res. Uncertainty Reduction (%) High-Res.~~
LEAF PHYSIOLOGY	1	Maximum carboxylation rate at 25°C	V_{cmax} (TrEv)	60	12	29.3 ~~53.5~~
	2		V_{cmax} (TrDec)	90	18	~~25.0~~ ~~52.6~~22.3
	3		V_{cmax} (TmpEv)	41	8.2	~~0.8~~ ~~10.8~~17.8
	4		V_{cmax} (TmpDec)	35	7	~~<0.1~~ ~~0.5~~0.4
	5		V_{cmax} (EvCn)	29	5.8	~~0.2~~ ~~3.4~~<0.1
	6		V_{cmax} (DecCn)	53	10.6	<0.1 ~~0.2~~
	7		V_{cmax} (EvShr)	52	10.4	~~32.8~~ ~~54.3~~25.1
	8		V_{cmax} (DecShr)	160	32	<0.1~~0.6~~
	9		V_{cmax} (C3Gr)	42	8.4	~~38.4~~ ~~57.1~~31.5
	10		V_{cmax} (C4Gr)	8	1.6	~~36.3~~ ~~62.9~~28.3
	11		V_{cmax} (Tund)	20	4	~~0.2~~ ~~3.8~~0.1
	12		V_{cmax} (Wetl)	20	4	<0.1~~0.1~~

LEAF PHYSIOLOGY	27	Activation energy for dark respiration	E_{R_d}	45000	2250	<0.1 ~~0.3~~
	28	Activation energy for V_{cmax}	$E_{V_{max}}$	58520	2926	~~11.5 32.1~~7.6
	29	Activation energy for K_O	E_{K_O}	35948	1797	~~0.3 1.1~~0.2
	30	Activation energy for K_C	E_{K_C}	59356	2967	~~3.7 12.9~~2.5
	31	Activation energy for k (PEPcase CO_2 specificity; C4)	E_k	50967	2548	<0.1~~<0.1~~
	32	Quantum efficiency (C3)	α_q	0.28	0.014	~~8.4 25.7~~5.2
	33	Quantum efficiency (C4)	α_i	0.04	0.002	~~1.9 17.4~~1.1
	34	Michaelis-Menten constant of Rubisco for CO_2	K_C	460e-6	23e-6	~~2.8 7.6~~2.3
	35	Michaelis-Menten constant of Rubisco for O_2	K_O	0.33	0.0165	0.4 ~~1.1~~
	36	Temp. dependency of CO_2 compensation point	gt	1.7e-6	8.5e-8	~~3.6 10.2~~2.4
LEAF GROWTH	37	Max. leaf area index	$\tilde{\Lambda}$	5	0.25	~~10.8 36.6~~7.6
	38	Temperature at leaf onset	T_ϕ (4)	10	0.5	~~15.0 47.2~~9.3
	39		T_ϕ (5,6,11)	10	0.5	~~24.4 48.4~~17.5
	40		T_ϕ (8)	8	0.5	~~4.7 31.8~~2.5
	41		T_ϕ (9,10,12)	2	0.5	~~52.0 71.9~~47.6
	42		T_ϕ (13)	15	1	~~55.9 70.5~~50.9
	43	Spatial range (1σ) of T_ϕ	T_r (4,8,13)	2	0.1	~~3.6 25.1~~2.2
	44		T_r (5,6,11)	2	0.1	

LEAF COMPOSITION	54	Chlorophyll ab content	C_{ab} (TrEv)	40	20	~~9.1 24.3~~7.4
	55		C_{ab} (TrDec)	15	20	~~64.6 74.5~~61.2
	56		C_{ab} (TmpEv)	15	20	~~41.4 67.3~~32.8
	57		C_{ab} (TmpDec)	10	20	~~73.4 80.7~~70.5
	58		C_{ab} (EvCn)	10	20	~~75.9 81.1~~74.3
	59		C_{ab} (DecCn)	10	20	~~71.2 79.8~~67.6
	60		C_{ab} (EvShr)	10	20	~~74.9 81.3~~72.7
	61		C_{ab} (DecShr)	10	20	~~70.1 79.1~~66.3
	62		C_{ab} (C3Gr)	10	20	~~76.0 81.5~~74.3
	63		C_{ab} (C4Gr)	5	20	~~85.2 88.5~~84.2
	64		C_{ab} (Tund)	10	20	~~75.1 80.9~~73.3
	65		C_{ab} (Wetl)	10	20	~~57.9 77.0~~50.7
	66		C_{ab} (Crop)	20	20	~~58.4 68.4~~55.4
	67	Dry matter content	C_{dm}	0.012	0.002	<0.1 ~~0.1~~
	68	Senescent material content	C_{sm}	0	0.01	~~0.3 2.1~~0.2
CANOPY STRUCTURE	69	Leaf inclination distribution function parameters	$LIDFa$	-0.35	0.1	~~24.1 42.2~~20.7
	70		$LIDFb$	-0.15	0.1	~~16.0~~

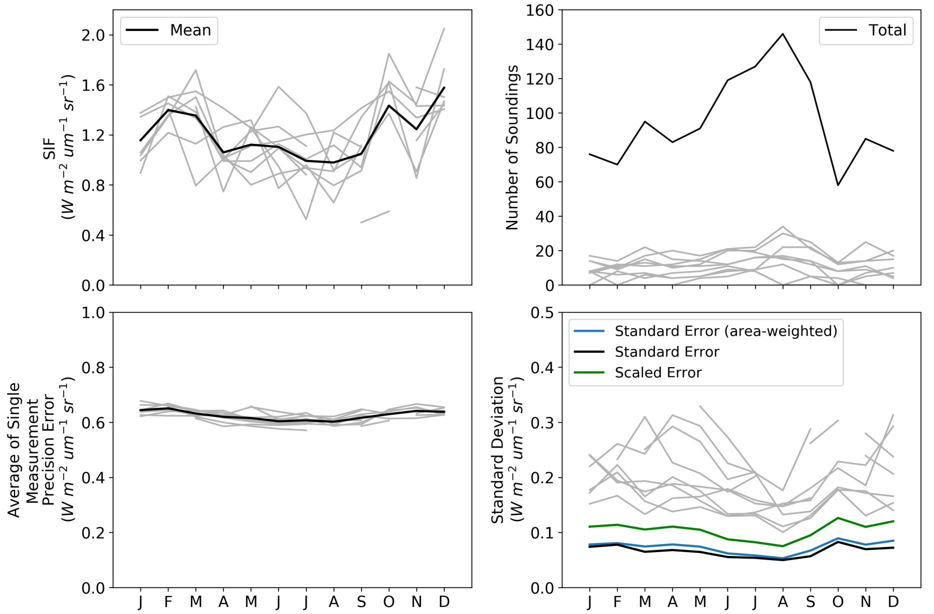

Figure 9. An example of the GOSAT SIF data and uncertainty calculations over a low-resolution model grid cell centered over the Amazon forest at 3.75°S and 65°W. Grey lines show individual 3° × 3°GOSAT grid cells. Black lines show the aggregated data for the 7.5° × 10°model grid cell. Bottom right shows the calculated uncertainty (standard deviation) at the model grid resolution in black, blue and green. The black line is the standard error calculated using Eq. 4; the blue line is the standard error calculated using Eq. A1; the green line is the same as the blue but scaled by $\sqrt{2}$ to account for correlated errors which is used in this study.

A2 GOSAT SIF Uncertainty Calculations

To get the variance of a target grid cell at the model grid resolution (ylat,xlon) we first determine the area-weighted variance of each GOSAT grid cell (ilat,jlon) within that target grid cell. The area-weighting per GOSAT grid cell ($\hat{Area}_{ilat,jlon}$) is calculated as the area divided by the total area of the target grid cell. This enables us to account for different grid cell sizes considering SIF is in physical units per unit area. We then sum the area-weighted variances and scale this uncertainty by the square root of two (see equation 4). Scaling the uncertainty in this way effectively doubles the variance in an independent dimension.

$$\sigma^2_{ylat,xlon} = \sqrt{2}\sum(\hat{Area}^2_{ilat,jlon}\cdot\sigma^2_{ilat,jlon}) \qquad \text{(A1)}$$

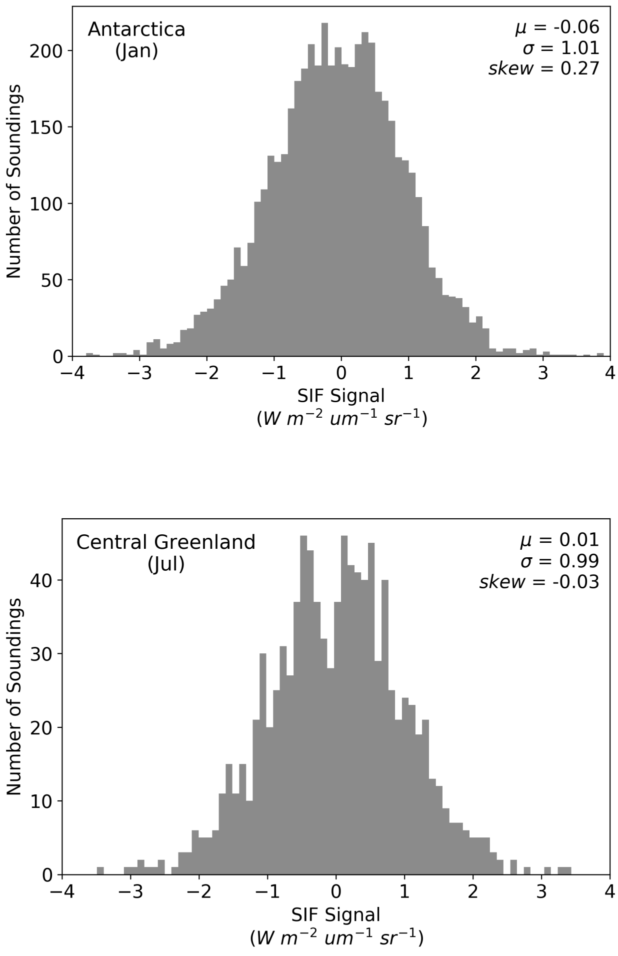

Figure 10. Analysis of systematic errors in the GOSAT SIF observations. We assess the zero-level offset corrected GOSAT SIF soundings over two ice-covered and therefore non-fluorescent regions. The first is Antarctica in January, between latitudes 70°S to 80°S and longitudes 75°W to 155°E. The second is central Greenland in July, between latitudes 73°N to 80°N and longitudes 30°W to 52°W. With no systematic error the mean (μ) value of the distribution should be on zero.

A3 Systematic Error in GOSAT SIF Observations

Acknowledgements. A. Norton was partly supported by an Australian Postgraduate Award provided by the Australian Government and a CSIRO OCE Scholarship. The research was funded, in part, by the ARC Center of Excellence for Climate System Science (grant CE110001028).

References

Baker, N. R.: Chlorophyll fluorescence: a probe of photosynthesis in vivo., Annual review of plant biology, 59, 89–113, doi:10.1146/annurev.arplant.59.032607.092759, 2008.

Baldocchi, D., Ryu, Y., and Keenan, T.: Terrestrial Carbon Cycle Variability, F1000Research, 5, 2371, doi:10.12688/f1000research.8962.1, 2016.

Baldocchi, D. D.: 'Breathing'of the terrestrial biosphere: lessons learned from a global network of carbon dioxide flux measurement systems, Australian Journal of Botany, 56, 1–26, doi:doi:10.1071/BT07151, 2008.

Beer, C., Reichstein, M., Tomelleri, E., and Ciais, P.: Terrestrial gross carbon dioxide uptake: global distribution and covariation with climate, Science, 329, 834–838, doi:10.1126/science.1184984, 2010.

Bodman, R. W.: Uncertainty in temperature projections reduced using carbon cycle and climate observations, Nature Climate Change, 3, 725–729, doi:10.1038/nclimate1903, 2013.

Boilley, A. and Wald, L.: Comparison between meteorological re-analyses from ERA-Interim and MERRA and measurements of daily solar irradiation at surface, Renewable Energy, 75, 135–143, doi:10.1016/j.renene.2014.09.042, 2015.

Carter, G. A., Freedman, A., Kebabian, P. L., and Scott, H. E.: Use of a prototype instrument to detect short-term changes in solar-excited leaf fluorescence, International Journal of Remote Sensing, 25, 1779–1784, doi:10.1080/01431160310001619544, 2004.

Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., Chhabra, A., DeFries, R., Galloway, J., Heimann, M., Jones, C., Le Quéré, C., Myneni, R. B., Piao, S., and Thornton, P.: Carbon and Other Biogeochemical Cycles, in: Climate Change 2013: The Physical Science Bases. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by Stocker, T. F., Qin, D., Plattner, G. K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., pp. 465–570, Cambridge University Press, Cambridge, 2013.

Damm, A., Guanter, L., Paul-Limoges, E., van der Tol, C., Hueni, A., Buchmann, N., Eugster, W., Ammann, C., and Schaepman, M.: Far-red sun-induced chlorophyll fluorescence shows ecosystem-specific relationships to gross primary production: An assessment based on observational and modeling approaches, Remote Sensing of Environment, 166, 91–105, doi:10.1016/j.rse.2015.06.004, 2015.

Daumard, F., Champagne, S., Fournier, A., Goulas, Y., Ounis, A., Hanocq, J.-F., and Moya, I.: A Field Platform for Continuous Measurement of Canopy Fluorescence, IEEE Transactions on Geoscience and Remote Sensing, 48, 3358–3368, doi:10.1109/TGRS.2010.2046420, 2010.

Evans, J. R.: Photosynthesis and nitrogen relationships in leaves of C3 plants, Oecologia, 78, 9–19, 1989.

Farquhar, G. D., von Caemmerer, S., and Berry, J. A.: A biochemical model of photosynthetic CO2 assimilation in leaves of C3 species, Planta, 90, 78–90, 1980.

Flexas, J., Escalona, J. M., and Medrano, H.: Water stress induces different levels of photosynthesis and electron transport rate regulation in grapevines, Plant, Cell and Environment, 22, 39–48, doi:10.1046/j.1365-3040.1999.00371.x, 1999.

Flexas, J., Briantais, J.-m., Cerovic, Z., Medrano, H., and Moya, I.: Steady-State and Maximum Chlorophyll Fluorescence Responses to Water Stress in Grapevine Leaves: A New Remote Sensing System, Remote Sensing Environment, 73, 283–297, doi:10.1016/S0034-4257(00)00104-8, 2000.

Flexas, J., Escalona, J. M., Evain, S., Gulías, J., Moya, I., Osmond, C. B., and Medrano, H.: Steady-state chlorophyll fluorescence (Fs) measurements as a tool to follow variations of net CO2 assimilation and stomatal conductance during water-stress in C3 plants., Physiologia plantarum, 114, 231–240, 2002.

Frankenberg, C., Butz, A., and Toon, G. C.: Disentangling chlorophyll fluorescence from atmospheric scattering effects in O 2 A-band spectra of reflected sun-light, Geophysical Research Letters, 38, L03 801, doi:10.1029/2010GL045896, 2011a.

Frankenberg, C., Fisher, J. B., Worden, J., Badgley, G., Saatchi, S. S., Lee, J.-E., Toon, G. C., Butz, A., Jung, M., Kuze, A., and Yokota, T.: New global observations of the terrestrial carbon cycle from GOSAT: Patterns of plant fluorescence with gross primary productivity, Geophysical Research Letters, 38, L17 706, doi:10.1029/2011GL048738, 2011b.

Frankenberg, C., O'Dell, C., Berry, J., Guanter, L., Joiner, J., Köhler, P., Pollock, R., and Taylor, T. E.: Prospects for chlorophyll fluorescence remote sensing from the Orbiting Carbon Observatory-2, Remote Sensing of Environment, 147, 1–12, doi:10.1016/j.rse.2014.02.007, 2014.

Freedman, A., Cavender-Bares, J., Kebabian, P., Bhaskar, R., Scott, H., and Bazzaz, F.: Remote sensing of solar-excited plant fluorescence as a measure of photosynthetic rate, Photosynthetica, 40, 127–132, doi:10.1023/A:1020131332107, 2002.

Friedlingstein, P., Cox, P., Betts, R., Bopp, L., von Bloh, W., Brovkin, V., Cadule, P., Doney, S., Eby, M., Fung, I., Bala, G., John, J., Jones, C., Joos, F., Kato, T., Kawamiya, M., Knorr, W., Lindsay, K., Matthews, H., Raddatz, T., Rayner, P., Reick, C., Roeckner, E., Schnitzler, K., Schnur, R., Strassmann, K., Weaver, A., Yoshikawa, C., and Zeng, N.: Climate–Carbon Cycle Feedback Analysis: Results from the C 4 MIP Model Intercomparison., … of Climate, 19, 3337–3353, doi:http://dx.doi.org/10.1175/JCLI3800.1, 2006.

Guanter, L., Frankenberg, C., Dudhia, A., Lewis, P. E., Gómez-Dans, J., Kuze, A., Suto, H., and Grainger, R. G.: Retrieval and global assessment of terrestrial chlorophyll fluorescence from GOSAT space measurements, Remote Sensing of Environment, 121, 236–251, doi:10.1016/j.rse.2012.02.006, 2012.

Guanter, L., Rossini, M., Colombo, R., Meroni, M., Frankenberg, C., Lee, J.-E., and Joiner, J.: Using field spectroscopy to assess the potential of statistical approaches for the retrieval of sun-induced chlorophyll fluorescence from ground and space, Remote Sensing of Environment, 133, 52–61, doi:10.1016/j.rse.2013.01.017, 2013.

Guanter, L., Zhang, Y., Jung, M., Joiner, J., Voigt, M., Berry, J. a., Frankenberg, C., Huete, A. R., Zarco-Tejada, P., Lee, J.-E., Moran, M. S., Ponce-Campos, G., Beer, C., Camps-Valls, G., Buchmann, N., Gianelle, D., Klumpp, K., Cescatti, A., Baker, J. M., and Griffis, T. J.: Global and time-resolved monitoring of crop photosynthesis with chlorophyll fluorescence., Proceedings of the National Academy of Sciences of the United States of America, 111, E1327–33, doi:10.1073/pnas.1320008111, 2014.

Houborg, R., Cescatti, A., Migliavacca, M., and Kustas, W.: Satellite retrievals of leaf chlorophyll and photosynthetic capacity for improved modeling of GPP, Agricultural and Forest Meteorology, 177, 10–23, doi:10.1016/j.agrformet.2013.04.006, 2013.

Hungershoefer, K., Breon, F., Peylin, P., Chevallier, F., Rayner, P., Klonecki, A., Houweling, S., and Marshall, J.: Evaluation of various observing systems for the global monitoring of CO2 surface fluxes, Atmospheric Chemistry and Physics, 10, 10 503–10 520, doi:10.5194/acp-10-10503-2010, 2010.

Kaminski, T., Scholze, M., and Houweling, S.: Quantifying the benefit of A-SCOPE data for reducing uncertainties in terrestrial carbon fluxes in CCDAS, Tellus, 62B, 784–796, doi:10.1111/j.1600-0889.2010.00483.x, 2010.

Kaminski, T., Knorr, W., Scholze, M., Gobron, N., Pinty, B., Giering, R., and Mathieu, P.-P.: Consistent assimilation of MERIS FAPAR and atmospheric CO2 into a terrestrial vegetation model and interactive mission benefit analysis, Biogeosciences, 9, 3173–3184, doi:10.5194/bg-9-3173-2012, 2012.

Kaminski, T., Knorr, W., Schürmann, G., Scholze, M., Rayner, P. J., Zaehle, S., Blessing, S., Dorigo, W., Gayler, V., Giering, R., Gobron, N., Grant, J. P., Heimann, M., Houweling, S., Kato, T., Kattge, J., Kelley, D., Kemp, S., Koffi, E. N., Köstler, C., Mathieu, P., Pinty, B., Reick, C. H., Rödenbeck, C., Schnur, R., Scipal, K., Sebald, C., Stacke, T., Van, A. T., Vossbeck, M., Widmann, H., and Ziehn, T.:

The BETHY/JSBACH Carbon Cycle Data Assimilation System : experiences and challenges, Journal of Geophysical Research: Biogeosciences, 118, 1–13, doi:10.1002/jgrg.20118, 2013.

Kato, S., Loeb, N. G., Rutan, D. A., Rose, F. G., Sun-Mack, S., Miller, W. F., and Chen, Y.: Uncertainty Estimate of Surface Irradiances Computed with MODIS-, CALIPSO-, and CloudSat-Derived Cloud and Aerosol Properties, Surveys in Geophysics, 33, 395–412, doi:10.1007/s10712-012-9179-x, 2012.

Kattge, J., Knorr, W., Raddatz, T., and Wirth, C.: Quantifying photosynthetic capacity and its relationship to leaf nitrogen content for global scale terrestrial biosphere models, Global Change Biology, 15, 976–991, doi:10.1111/j.1365-2486.2008.01744.x, 2009.

Knorr, W., Kaminski, T., Scholze, M., Gobron, N., Pinty, B., Giering, R., and Mathieu, P.-P.: Carbon cycle data assimilation with a generic phenology model, Journal of Geophysical Research, 115, G04 017, doi:10.1029/2009JG001119, 2010.

Koffi, E. N., Rayner, P. J., Scholze, M., Chevallier, F., and Kaminski, T.: Quantifying the constraint of biospheric process parameters by CO2 concentration and flux measurement networks through a carbon cycle data assimilation system, Atmospheric Chemistry and Physics, 13, 10 555–10 572, doi:10.5194/acp-13-10555-2013, 2013.

Koffi, E. N., Rayner, P. J., Norton, A. J., Frankenberg, C., and Scholze, M.: Investigating the usefulness of satellite-derived fluorescence data in inferring gross primary productivity within the carbon cycle data assimilation system, Biogeosciences, 12, 4067–4084, doi:10.5194/bg-12-4067-2015, 2015.

Kuppel, S., Chevallier, F., and Peylin, P.: Quantifying the model structural error in carbon cycle data assimilation systems, Geoscientific Model Development, 6, 45–55, doi:10.5194/gmd-6-45-2013, 2013.

Lee, J.-e., Frankenberg, C., van der Tol, C., Berry, J. A., Guanter, L., Boyce, C. K., Fisher, J. B., Morrow, E., Worden, J. R., Asefi, S., Badgley, G., and Saatchi, S.: Forest productivity and water stress in Amazonia: observations from GOSAT chlorophyll fluorescence., Proceedings of the Royal Society B, 280, 20130 171, doi:10.1098/rspb.2013.0171, 2013.

Macbean, N., Peylin, P., Chevallier, F., Scholze, M., and Schürmann, G.: Consistent assimilation of multiple data streams in a carbon cycle data assimilation system, Geoscientific Model Development, 9, 3569–3588, doi:10.5194/gmd-9-3569-2016, 2016.

Magney, T. S., Frankenberg, C., Fisher, J. B., Sun, Y., North, G. B., Davis, T. S., Kornfeld, A., and Siebke, K.: Connecting active to passive fluorescence with photosynthesis: a method for evaluating remote sensing measurements of Chl fluorescence, New Phytologist, doi:10.1111/nph.14662, 2017.

Parazoo, N. C., Bowman, K., Frankenberg, C., Lee, J.-E., Fisher, J. B., Worden, J., Jones, D. B. a., Berry, J., Collatz, G. J., Baker, I. T., Jung, M., Liu, J., Osterman, G., O'Dell, C., Sparks, A., Butz, A., Guerlet, S., Yoshida, Y., Chen, H., and Gerbig, C.: Interpreting seasonal changes in the carbon balance of southern Amazonia using measurements of XCO2 and chlorophyll fluorescence from GOSAT, Geophysical Research Letters, 40, 2829–2833, doi:10.1002/grl.50452, 2013.

Parazoo, N. C., Bowman, K., Fisher, J. B., Frankenberg, C., Jones, D. B. A., Cescatti, A., Perez-Priego, O., Wohlfahrt, G., and Montagnani, L.: Terrestrial gross primary production inferred from satellite fluorescence and vegetation models, Global Change Biology, 20, 3103–3121, doi:10.1111/gcb.12652, 2014.

Peylin, P., Bacour, C., Macbean, N., Leonard, S., Rayner, P., Kuppel, S., Koffi, E., Kane, A., Maignan, F., Chevallier, F., Ciais, P., and Prunet, P.: A new stepwise carbon cycle data assimilation system using multiple data streams to constrain the simulated land surface carbon cycle, Geoscientific Model Development, 9, 3321–3346, doi:10.5194/gmd-9-3321-2016, 2016.

Porcar-Castell, A., Tyystjärvi, E., Atherton, J., van der Tol, C., Flexas, J., Pfündel, E. E., Moreno, J., Frankenberg, C., and Berry, J. a.: Linking chlorophyll a fluorescence to photosynthesis for remote sensing applications: mechanisms and challenges., Journal of experimental botany, 65, 4065–95, doi:10.1093/jxb/eru191, 2014.

Rayner, P., Scholze, M., Knorr, W., and Kaminski, T.: Two decades of terrestrial carbon fluxes from a carbon cycle data assimilation system (CCDAS), Biogeochemical Cycles, 19, GB2026, doi:10.1029/2004GB002254, 2005.

Rossini, M., Nedbal, L., Guanter, L., Ač, A., Alonso, L., Burkart, A., Cogliati, S., Colombo, R., Damm, A., Drusch, M., Hanus, J., Janoutova, R., Julitta, T., Kokkalis, P., Moreno, J., Novotny, J., Panigada, C., Pinto, F., Schickling, A., Schüttemeyer, D., Zemek, F., and Rascher, U.: Red and far red Sun-induced chlorophyll fluorescence as a measure of plant photosynthesis, Geophysical Research Letters, 42, 1632–1639, doi:10.1002/2014GL062943, 2015.

Schimel, D., Pavlick, R., Fisher, J. B., Asner, G. P., Saatchi, S., Townsend, P., Miller, C., Frankenberg, C., Hibbard, K., and Cox, P.: Observing terrestrial ecosystems and the carbon cycle from space, Global Change Biology, 21, 1762–1776, doi:10.1111/gcb.12822, 2015.

Scholze, M., Kaminski, T., Rayner, P., Knorr, W., and Giering, R.: Propagating uncertainty through prognostic carbon cycle data assimilation system simulations, Journal of Geophysical Research, 112, D17 305, doi:10.1029/2007JD008642, 2007.

Scholze, M., Kaminski, T., Knorr, W., Blessing, S., Vossbeck, M., Grant, J. P., and Scipal, K.: Simultaneous assimilation of SMOS soil moisture and atmospheric CO2 in-situ observations to constrain the global terrestrial carbon cycle, Remote Sensing of Environment, 180, 334–345, doi:10.1016/j.rse.2016.02.058, 2016.

Tarantola, A.: Inverse Problem Theory and Methods for Model Parameter Estimation, SIAM, Philadelphia, 2005.

Van der Tol, C., Verhoef, W., Timmermans, J., Verhoef, A., and Su, Z.: An integrated model of soil-canopy spectral radiances, photosynthesis, fluorescence, temperature and energy balance, Biogeosciences, 6, 3109–3129, doi:10.5194/bg-6-3109-2009, 2009.

Van der Tol, C., Berry, J. A., Campbell, P. K. E., and Rascher, U.: Models of fluorescence and photosynthesis for interpreting measurements of solar-induced chlorophyll fluorescence, Journal of Geophysical Research: Biogeosciences, 119, 1–16, doi:10.1002/2014JG002713.We, 2014.

Verrelst, J., Rivera, J. P., Tol, C. V. D., Magnani, F., Mohammed, G., and Moreno, J.: Global sensitivity analysis of the SCOPE model: What drives simulated canopy-leaving sun-induced fluorescence?, Remote Sensing of Environment, 166, 8–21, doi:10.1016/j.rse.2015.06.002, 2015.

Walther, S., Voigt, M., Thum, T., Gonsamo, A., Zhang, Y., Koehler, P., Jung, M., Varlagin, A., and Guanter, L.: Satellite chlorophyll fluorescence measurements reveal large-scale decoupling of photosynthesis and greenness dynamics in boreal evergreen forests, Global Change Biology, 22, 2979–2996, doi:10.1111/gcb.13200, 2016.

Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J., and Viterbo, P.: The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data, Water Resources Research, 50, 7505–7514, doi:10.1002/2014WR015638, 2014.

Wilson, M. and Henderson-Sellers, A.: Global archive of land cover and soils data for use in general-circulation climate models, International Journal of Climatology, 5, 119–143, 1985.

Yang, X., Tang, J., Mustard, J. F., Lee, J.-e., Rossini, M., Joiner, J., Munger, J. W., Kornfeld, A., and Richardson, A. D.: Solar-induced chlorophyll fluorescence that correlates with canopy photosynthesis on diurnal and seasonal scales in a temperate deciduous forest, Geophysical Research Letters, 42, 2977–2987, doi:10.1002/2015GL063201, 2015.

Zhang, Y., Guanter, L., Berry, J. A., Joiner, J., van der Tol, C., Huete, A., and Gitelson, A.: Estimation of vegetation photosynthetic capacity from space-based measurements of chlorophyll fluorescence for terrestrial biosphere models, Global Change Biology, 20, 3727–3742, doi:10.1111/gcb.12664, 2014.

Zhang, Y., Guanter, L., Berry, J. A., van der Tol, C., Yang, X., Tang, J., and Zhang, F.: Model-based analysis of the relationship between sun-induced chlorophyll fluorescence and gross primary production for remote sensing applications, Remote Sensing of Environment, 187, 145–155, doi:10.1016/j.rse.2016.10.016, 2016.

Ziehn, T., Knorr, W., and Scholze, M.: Investigating spatial differentiation of model parameters in a carbon cycle data assimilation system, Global Biogeochemical Cycles, 25, 1–13, doi:10.1029/2010GB003886, 2011.