

Interactive comment on “A linear method for the retrieval of sun-induced chlorophyll fluorescence from GOME-2 and SCIAMACHY data” by P. Köhler et al.

Anonymous Referee #3

Received and published: 21 January 2015

The authors present an improved retrieval methodology for SIF retrieval. They are continuing previous work by Joiner et al. (2013). The method has now also been applied to a couple of months of Scia data. The topic is interesting but the paper needs to be substantially revised before I can recommend publication. I encourage the authors to do so because I see enough potential for this paper to become publishable. Below I have listed major and minor comments. I would appreciate if these are addressed point by point. Many minor comments however serve as an example. I emphasize here that the authors should use these comments to also critically look at other parts of the manuscript.

C4677

Major comments:

-Overall, I feel that the style of the manuscript needs to be improved. I quite often struggled to understand what the authors are trying to say. Sentences can sometimes be wordy and beside the point. Also, the way of presenting their material is sometimes not very well structured (e.g., introduction and discussion of eq.3 should be moved up to p.12179) or confusing (e.g., discussion of training and test spectra should be more clearly separated). I encourage the authors to reread the manuscript and improve its style. Perhaps a native English speaker can help. I have listed detailed comments until section 3.1 to make my point. Please address these comments. They should also be used, however, to thoroughly improve the remainder of the manuscript, which is necessary... I may be nitpicky but these detailed comments are intended to help the authors make their story more clear.

-In the description of the retrieval methodology, there is quite some overlap with Joiner et al. (2013). I think the description can be made much more concise and to the point. Also, the focus should be on differences with Joiner et al. (2013). These should definitely be more explicitly stated and argued. The unique contribution that the authors make in their present manuscript to the retrieval methodology should stand out more clearly.

-The authors linearize the forward model by estimating T_{up} in a pre-processing step, but this comes at a cost (loss of accuracy). As far as I can see, the only reason for linearizing the forward model is computational speed: for the automated determination of PCs according to the BIC, retrieval for a single target pixel has to be run multiple times. Linearization is not needed for the backward elimination per se (or am I missing something?). The authors need to show whether the supposed accuracy improvement when optimizing the choice of PCs outweighs the loss of accuracy due to the linearization. The appropriate comparison for the retrieval simulations of sect.4.3 would be to compare the backward elimination mode (including the linearization) with the full-PC non(!)-linear forward model fit. Also, is computational speed indeed an issue here?

C4678

Can the authors give an estimate of the increase in computation time when using backward elimination with the non-linear forward model instead of backward elimination with the linear forward model. Other comments and questions that I have about the analysis in sect.4.3 can be found below. I encourage the authors to rethink their analysis and way of presenting it. There are interesting results, but it can be presented more convincingly.

-I have stopped doing a detailed analysis of sect.5. As an overall remark, this section is primarily a description of the SIF data sets. Maps, trend plots and biome-specific analyses are very similar to the previous papers and I am struggling to see a unique contribution here. If the point of this paper is an improvement in retrieval methodology, also the focus of this section should be much more on presenting and discussing differences between data sets to support the conclusions from the retrieval simulations, rather than just trying to convince the reader that the retrieval algorithm is picking up a fluorescence signal, which has been done before. If I think about possible ways to improve this section and make it more worthwhile, the PNAS paper comes to mind. You can ask yourself what is needed to produce a SIF database that can readily be compared with GPP flux tower measurements? To my recollection, in that paper corrections were made, for example, for the area of a grid box actually covered by vegetation. Also, can you make corrections for the presence of clouds? This is just a suggestion to help the authors. I am open to other ways of improving this section.

Minor comments:

p.12174,l.20: photosynthetically-active -> photosynthetically active p.12175,l.1: Start new paragraph after 'field'? p.12175,l.3: I believe the paper explaining the methodology is Frankenberg et al. (2011a) rather than Frankenberg et al. (2011b), and this reference would be more appropriate here. p.12175,l.6: enabled the possibility to evaluate -> 'enabled evaluation' or 'gave the possibility to evaluate' p.12175,l.9-10: used both the 757 and the 770 nm spectral regions -> I would speak of fit windows instead of spectral regions here? How broad are these fit windows? In view of the discussion

C4679

of GOME2 retrievals below, you could mention that GOSAT only covers the O2 A band. Then, the reader understands why such small (micro) windows are used in the GOSAT retrievals. p.12175,l.10: and is devoid of significant atmospheric absorption -> Do you mean that even for the isolated Fraunhofer line at 770.1 nm there is residual absorption by oxygen? p.12175,l.11-12: Using only the single line in 770 nm simplifies the forward model -> Is this true? The method by Joiner et al. (2011) could equally well be applied to the micro windows of Guanter et al. (2012), I would say. So using an isolated line doesn't really simplify. An important practical problem though is that a measured irradiance should be available. p.12175,l.11-12: single line in 770 nm -> single line at 770 nm (throughout paper!) p.12175,l.14: The method proposed by Frankenberg et al. (2011b)... -> You haven't yet discussed the method by Joiner et al. (2011), which to my recollection relies on having available a measured irradiance spectrum. p.12175,l.21: narrow -> Why narrow? (See your paper.) p.12175,l.28: (2deg x 2deg) which -> What is temporal resolution? Add comma: '2deg), which'. p.12175,l.28-29: high retrieval noise -> Do you have a reference for that? Or, in view of your G2 and SCIA maps, can you argue that GOSAT has relatively larger noise errors than G2 and SCIA? p.12176,l.1: with lower spectral resolutions -> I would insert here 'but better spatial coverage'. p.12176,l.4: instruments which -> instrument, which p.12176,l.8: might be -> is p.12176,l.9-11: For this reason, ...continuous spatial sampling. -> Overly cautious and wordy. Sentence can actually be left out, or changed to 'Thus, the representativeness of spatially mapped SiF is enhanced by using data from instruments which provide a continuous spatial sampling.' p.12176,l.14-15: spectral and radiometric ... enables a significant increase -> Why? Is radiometric performance of G2 superior to SCIA and GOSAT? p.12176,l.18: applied on -> applied to p.12176,l.20-22: Reading this, I ask myself whether there any other ways to get info about SiF? p.12176,l.25-26: consists in -> wordy, replace for example by 'is' p.12176,l.26-27: used singular ... respectively. -> Again wordy: leave out 'used' (throughout paper!). Why make such a formal distinction between PCs (Joiner) and SVs (Guanter) here? It distracts. They are (basically) the same, to my knowledge. For example, simply replace phrase by 'the

C4680

principal components needed for retrieval' (throughout paper!). p.12176,l.28-29: We test ... retrieval method -> A bit awkward, perhaps replace by: 'We apply the proposed retrieval method to data from G2 and SCIA to demonstrate its feasibility' p.12177,l.1: evaluated at 740 nm -> Why do you mention this here? I don't think it is needed at this place, but doing so makes me think about your remark. For example, are you doing something different than previous authors? It doesn't seem so. Is the wavelength at which SiF is retrieved a distinct property of your retrieval method (e.g., depending on the choice of fit window)? No, because 740 nm just is the peak wavelength of the assumed spectral SiF model, which has been determined in laboratory measurements.

p.12177,l.8: Experiment -> Experiment-2 p.12177,l.13: overpass -> equator overpass (also in next section) p.12177,l.15-18: The fourth channel ... Fig. 1. -> Confusing sentence. What point are you trying to make? Perhaps change to 'The fourth channel (590 - 790 nm), which has a spectral resolution of 0.5 nm and a SNR up to 2000, covers the SiF wavelength region (Fig. 1).'

p.12177,l.20: to evaluate ... peak in 740 -> Peak 'at' 740 nm. Also, you are not retrieving SiF at 740 nm per se, but across the entire range between 720 nm and 758 nm. SiF at 720 nm contributes to the reported 740-nm peak value through the assumed spectral model.

p.12177,l.22-24: The GOME-2 level 1B ... device. -> Wordy: 'in photons per second' can be left out. It is perhaps important to mention though that the solar irradiance measurement is done daily in principle. For example, 'The GOME-2 level 1b product consists of radiance spectra and a daily solar irradiance measurement.'

p.12177,l.25: available -> 'used' (GOME-2 level 1b data from the entire mission were available for your study, I guess. See also next section.)

p.12178,l.13: mode which -> mode, which p.12178,l.19-21: That means ... limitations. -> Sentence can be completely left out when 'captured' in l.18 is changed to 'downlinked'.

p.12179,l.2: 'basically' can be left out

p.12179,l.6: main challenge to retrieve SiF -> main challenge when retrieving SiF

p.12179,l.9: to data-driven -> to the data-driven

p.12179,l.10: of the TOA signal -> to the TOA signal

p.12179,l.10-12: In simplified ... by SiF. -> I don't quite under-

C4681

stand what you are trying to say here? What exactly is simplified? It seems that you are stating an obvious thing and I would think that this sentence can be left out.

p.12179,l.15: I_{sc} -> Where does the 'c' stand for? I understand that you want to distinguish s(urface) from s(olar), but perhaps you can come up with more understandable subscripts.

p.12179,l.14,eq.1: Why do you use different symbols (L and F) for the radiance?

p.12179,l.208-225: It seems that you are basically repeating and summarising previous papers here, particularly the one by Joiner et al. (2013). Why not refer to that paper and just state final expressions? Alternatively, you can indeed decide to summarise the derivation, but then you need to give more explanation. I think that the current version raises questions with the reader. For example: if you mention the approximation in l.21, state explicitly the reason for doing so (atmospheric scattering small in the near-infrared), relate T_{down_up} to T_{up} , and then explicitly argue why you can write p.12184,l.8,eq.4 (no diffuse contribution). Also, p.12183,eq.3 is not consistent with eq.2 (ρ_o).

p.12179,l.18: amount of -> can be left out

p.12180,l.2-3: apparent reflectance -> please give definition

p.12180,l.3-6: A subset ... SiF retrievals. -> Please leave out descriptions of detector channels. They are already given before and it is not needed to repeat them here. For example, simply say 'In our SiF retrievals we use a fit window extending from 720 nm to 758 nm'. Does the fit window differ from the one used by Joiner et al. (2013)?

p.12180,l.8-9: at which we evaluate the amount of SiF -> Suppose you choose a fit window that does not cover 740 nm, for example a combination of windows from 720 nm to 739 nm and from 741 nm to 758 nm. Does your retrieval then fail? Do you think you cannot retrieve SiF at 740 nm with the current retrieval algorithm then? (See previous remark on this point.)

p.12180,l.14: requires a more complex model -> Why?

p.12180,l.14-16: Simultaneously ... accuracy. -> In view of the previous bullet, why not stop at 721.5 nm then?

p.12180,l.8-19: I think point 3 and 4 repeat point 2 and can be left out.

p.12180,l.22-23: training set and test set -> Please give a definition of training and test set here or at the start of the next section. Also, I find test set a confusing name (what are you testing?). Why not say target set? For example, 'The training set is a set of reflectance spectra over non-vegetated

C4682

areas used to determine PCs and the target set is the set of reflectance spectra over vegetated areas for which a SiF retrieval is attempted.'

I stopped giving comprehensive and detailed comments here.

p.12181,l.20: sun-glint -> Why mention this? You include only land pixels in your training set. p.1218-12181,sect.3.2: You select different spectra for the target set than Joiner et al. (2013). For example, you don't select sea ice or cloudy ocean. Why? Does it make a difference? p.12182,l.1-2: Actually ... test set. -> I don't understand this sentence? p.12182,l.16-22: Why are discussing the spectral dependence of the surface reflectivity for vegetation here? Vegetated pixels are not included in your training set, are they? p.12186,l.9-12: Please add reference. How about modelling noise error for Scia? p.12186,sect.3.6: Why not simply use the noise errors reported in the level-1b products? p.12186,eq.9: Please state explicitly that SD is the standard error of the mean. (Then, formula can perhaps also be left out.) p.12189,sect.4.2: Backward elimination is not used here and the forward model is still non-linear, right? p.12190,l.25: You mention the danger of overfitting and fitting of noise a couple of times, which is the reason why you introduce backward elimination. Can you substantiate and illustrate the claim that you are overfitting for your SIF retrieval methodology? The fact that PCs are thrown out according to the BIC supports this claim, I believe, but it would be nice if you could illustrate the danger of overfitting with some insightful plots for an example case before you actually introduce backward elimination. Alternatively, a short discussion of eigenvalues and noise levels could perhaps also be insightful.

p.12189,sect.4.3/fig.5: About the bias: I don't understand why the retrieved SIF is always overestimating the input SIF? About correlations: If the backward elimination mode performs better, the correlation should be higher, right? About SD: what is SD here? You say 'the SD of retrieved SIF values' in the caption, but I guess it is the standard deviation of the difference between retrieved and input SIF? (Standard deviation of retrieved SIF values is not meaningful here.) Same for RMSE? So, in this figure $\text{bias}^2 + \text{SD}^2 = \text{RMSE}^2$ should hold? From the figure, this does not seem to be the

C4683

case, but I might be wrong. Furthermore, I doubt whether the comparison between the backward elimination mode and the linear model mode is presented properly. I find it confusing anyway: The number of PCs (x-variable) in the former case is the initial number of PCs that are input to the backward elimination pre-processing step. In the actual retrieval, less PCs are used (how many? are these indeed the PCs with the largest 'explained variance'?). The number of PCs (x-variable) in the latter case however is the number of PCs actually used in the retrieval. In backward elimination mode, the number of PCs selected for the actual retrieval step should saturate, as you suggest in the text, and from that point on it should not matter whether you further increase the initial number of PCs. Thus, from the saturation point onwards, I would expect data points in fig.4 to be basically repetitions (the actual retrieval is perfectly the same). The figure suggests otherwise (independent data points). This should be explained more clearly (why is there still variability in data points for the backward elimination mode for large numbers of initial PCs?). Continuing this point, an alternative and perhaps more appropriate comparison would be a comparison of retrieval outcomes for the two modes when the same number of PCs is used in the actual retrieval. For example, SIF biases when using eight optimally chosen PCs in backward elimination mode and when using the first (in terms of explained variance) eight PCs in (non?)-linear model fit mode. Then, we have the same variable on the x-axis for both modes. To me, the backward elimination step to determine the optimal set of PCs really is a pre-processing step and not part of the actual SIF retrieval.

p.12189,sect.4.3: Can a figure be included showing how many PCs on average are eventually selected by the backward elimination step? Otherwise discuss this. p.12192,l.6-7: should amount at least eight days -> 'should amount to at least eight days'; please give reference p.12193,l.13-14,fig.7: unbiased relationship -> Please provide statistics to support the claim that the NDVI-slope is not significantly different from one, whereas the SIF slope is. The NDVI-slope is larger than one, which agrees with the hypothesis of cloud contamination. p.12192,l.11-p.12193,l.22: Scia SIF values are lower than G2 SIF values. You investigate the effect of clouds using the NDVI, which is

C4684

a rather indirect approach. Can't you use cloud data directly? For example, cloud fractions from FRESCO or OCRA (don't know actually which ones you used). I think you should also discuss that a difference is not expected because the overpass times are so close together. (I am not an expert on this, but I would say that an hour difference is too small to see differences in vegetation activity. True?) p.12194,eq.11: missing $\wedge^2$ in formula. p.12199,l.8: SIF signal is absorbed -> shielding or extinction, cloud droplets do not absorb in this wvl range (throughout paper). p.12208,fig.4: Is SD the same as SD_Fs in eq.9?? I guess not. Why then use an acronym for a plain and ordinary standard deviation here?

sect.5.: I have stopped doing a detailed analysis of sect.5. As an overall remark, I encourage the authors to rethink also this section and ask themselves what point they are trying to convey here.

Interactive comment on Atmos. Meas. Tech. Discuss., 7, 12173, 2014.