The Authors present a new sun-induced chlorophyll fluorescence (SIF) retrieval method based on principal component analyses which they apply to GOME-2 and SCIAMACHY data. Satellite based SIF measurements have the potential to enhance our knowledge about the carbon cycle. The presented method can not only be used for past and current satellite instruments but also be adapted for future satellite missions. Therefore, the paper covers an important topic which fits well to the aims and scopes of AMT and I recommend publication in AMT once my comments have been addressed.

1 Major comments

Orthogonality: You implicitly assume that the SIF spectrum $h_f \cdot T_\uparrow$ is orthogonal to all used PCs. Any correlation will result in the retrieval being not perfectly able to disentangle the amount of SIF (F_s) from the other fitted coefficients. This should at least be discussed. Additionally, you could enhance the confidence in the method by showing that there are no such spectral correlations.

Heterogenity: You emphasise that you need to include soundings above ice in order to construct a training data set that covers enough heterogeneity to be useful for all scenes (including vegetation). What is the justification to assume that vegetated surfaces can be excluded from the training? An indication for this would be if the residuals (measured radiance minus fitted radiance) are not larger for vegetated surfaces than for surfaces included in the training.

Training and test data set: Section 3.2 describes the selection of training and test data sets without making clear what the purpose of theses dataset are. Please add a discussion to the fundamentals.

Eq.1: i) Where is this equation coming from? Please add a reference or derive (Joiner et al. (2013) does not give more information). ii) It should be mentioned that this is already an approximation of the RT and the assumptions made (e.g., plane parallel atmosphere) and their potential influence on the retrieval should be discussed.

Sec.3.3: The terms “apparent reflectance” and “atmospheric path radiance” have not been discussed in respect to Eq.1 and 2. Please use consistent terms and describe how these terms relate to the terms used in Eq.1 and 2.

Eq.3: Eq.3 implies that the total reflectance is a product of transmittance (sum over the PCs) and a polynomial. This is conceptual different from Eq.1 and 2 with the additive term ρ_0 ($\rho_0 + \rho_s \cdot T$). Maybe this clarifies when introducing a proper definition of apparent reflectance to Eq.2. Otherwise the connection to Eq.1 and 2 is unclear.

Description of the retrieval: Sec.3.4,P12184L17: "... for the actual retrieval (Eq.3)..." Eq.3 is the forward model not the retrieval. It includes the to be derived state vector elements. However, a detailed description of how to obtain these parameters is missing. I suppose you perform a least squares fit without any prior constraints (?). Please add a description.

Eq.5: Please explain the intention of introducing γ and the advantage (?) of enhancing the dimensionality of the problem. With $\gamma = \alpha \cdot \beta^T$, the elements of γ cannot be considered independent but within the process of backward elimination they are handled as if they were independent. According to Eq.3 the apparent reflection (polynomial) is multiplied with the transmittance (PCs), i.e., each PC is multiplied with the same polynomial. In contrast to this, Eq.5 allows to have different polynomials for different PCs. Physically and in respect to Eq.3, this seems to be not meaningful.

2 Minor comments

Citations: Check citation style within the entire manuscript.

P12174L14: It is not the first time, see Joiner et al. (2012).

P12176L23-P12177L5: Please references to the corresponding sections of the paper.

Sec.3.2: As you state "the selection of the training set is highly relevant". A map would help to show which soundings have been selected and improve reproducibility.

Sec.3.2: "...the training set is sampled on a daily basis." What does this mean?

Sec.3.2: Please describe the applied cloud/aerosol filtering for the training data set (if any).

Fig.2: Due to the instrumental resolution, the used micro-window at about 722nm may be influenced by H_2O (?) absorption (e.g., dip in the middle of the window). This can result in errors of the derived apparent reflection and hence transmission. How large is the potential influence on the derived SIF?

Sec.3.5: "... previously presented steps..." Which previously presented steps do you mean that reduced the number of coefficients?

Sec.3.5: "Eq.3 contains $i \cdot j$ ..." i and j are only index variables. You could define, e.g., N and M. Additionally, Eq.3 does not contain $N \cdot M + 1$ but $N + M + 1$ coefficients.

P12185L3: Define "a few".

P12185L5: Please explain what you mean by overfitting. Joiner et al. (2013) used more PCs than you do (?). Does this imply they were overfitting?

Eq.6: Why do you not use the noise estimate provided in the L1 files.

P12187L11: Please compare the obtained RSS values with values expected from instrumental noise. This comparison can tell you if you are already fitting noise (too many PCs) or if important PCs are missing.

P12187L16: Define "high average".

P12188L7: Check unit of water vapour columns.

Sec.4.2,P12189L3: Please explain why you decided to use the first eight PCs but not the just introduced backward elimination algorithm.

Fig.4: Fig.4 shows a good correlation between simulated and retrieved SIF. However, from this plot alone, systematic effects cannot be excluded. As an example, high H_2O values could always be hidden in the lower end of the error bar. You could enhance the confidence in the retrieval by showing that retrieved minus simulated SIF does not depend on H_2O, solar zenith angle, AOD, etc.

P12191L13: "10 PCs" When applying the algorithm to real data, some PCs may also have to account for instrumental effects which was not the case for the simulations. Are 10 PCs still enough in this case? How often does the back elimination algorithm decide that 10 PCs (i.e., all) are needed.

Sec.5.1,P12193: You use NDVI as proxy for cloud contamination. This seems to be very indirect. What if cloud contamination effects NDVI much less than the amount of vegetation. In this case the NDVI to NDVI graph would still have a slope close to 1. Therefore, your conclusions "it must be concluded that...most likely..." seems to be too strong. Why not using FRESCO?

Fig.7,Fig.9: Please discuss potential reasons for slopes being different from 1 (especially Fig.9). V25 uses more PCs; if one of the additional PCs is not orthogonal to the SIF spectrum $h_f \cdot T_{\uparrow}$, I would expect too low SIF values. Could this explain the observed discrepancy? If so, I would expect that your retrieval also critically depends on the number of PCs (even though simulated data - Fig.5 - does not suggest this). Additionally, please add a colour bar.

Eq.10: This is not the standard error of the mean. Please check the equation. Additionally, why introducing Eq.10 when Eq.8 was introduced for the same purpose?

Fig.10: Please show the selected areas in a map (e.g., together with the selected training data set, see earlier comment).

P12195L18: "At this point..." Could it help to compare with GOSAT? The lack of validation data sets is a general problem for satellite based SIF retrievals. Are there plans to improve the situation.

Sec.5.3: Please mark the approximate region of the SAA in one of the maps.

Sec.5.5: Please show the analysed area in a map.

Sec.5.5,P12200L2: "...is robust against..." This conclusion seems to be too strong. Fig.13 shows that changing the cloud detection threshold from 0.25 to 1. can result in SIF averages being about 30% off. As the sample without cloud filter (threshold 1.) includes also all cloud free scenes, it can be assumed that the influence on individual soundings is much larger than 30%.

P12201L3: ENVISAT was launched in March 2002.

P12201L10: If the proposed method could also be adapted to CarbonSat (which is probably the case), CarbonSat (Bovensmann et al., 2010; Buchwitz et al., 2013) should be cited in this context because it enables carbon cycle analysis based on simultaneous measurements of XCO2 and SIF.

References

Bovensmann, H., Buchwitz, M., Burrows, J. P., Reuter, M., Krings, T., Gerilowski, K., Schneising, O., Heymann, J., Tretner, A., and Erzinger, J.: A remote sensing technique for global monitoring of power plant CO_2 emissions from space and related applications, Atmospheric Measurement Techniques, 3, 781–811, doi:10.5194/amt-3-781-2010, URL `http://www.atmos-meas-tech.net/3/781/2010/`, 2010.

Buchwitz, M., Reuter, M., Schneising, O., Boesch, H., Guerlet, S., Dils, B., Aben, I., Armante, R., Bergamaschi, P., Blumenstock, T., Bovensmann, H., Brunner, D., Buchmann, B., Burrows, J., Butz, A., Chédin, A., Chevallier, F., Crevoisier, C., Deutscher, N., Frankenberg, C., Hase, F., Hasekamp, O., Heymann, J., Kaminski, T., Laeng, A., Lichtenberg, G., Mazìre, M. D., Noël, S., Notholt, J., Orphal, J., Popp, C., Parker, R., Scholze, M., Sussmann, R., Stiller, G., Warneke, T., Zehner, C., Bril, A., Crisp, D., Griffith, D., Kuze, A., O'Dell, C., Oshchepkov, S., Sherlock, V., Suto, H., Wennberg, P., Wunch, D., Yokota, T., and Yoshida, Y.: The Greenhouse Gas Climate Change Initiative (GHG-CCI): Comparison and quality assessment of near-surface-sensitive satellite-derived CO_2 and CH_4 global data sets, Remote Sensing of Environment, doi:http://dx.doi.org/10.1016/j.rse.2013.04.024, URL http://www.sciencedirect.com/science/article/pii/S0034425713003520, 2013.

Joiner, J., Yoshida, Y., Vasilkov, A. P., Middleton, E. M., Campbell, P. K. E., Yoshida, Y., Kuze, A., and Corp, L. A.: Filling-in of near-infrared solar lines by terrestrial fluorescence and other geophysical effects: simulations and space-based observations from SCIAMACHY and GOSAT, Atmospheric Measurement Techniques, 5, 809–829, doi:10.5194/amt-5-809-2012, URL http://www.atmos-meas-tech.net/5/809/2012/, 2012.

Joiner, J., Guanter, L., Lindstrot, R., Voigt, M., Vasilkov, A. P., Middleton, E. M., Huemmrich, K. F., Yoshida, Y., and Frankenberg, C.: Global monitoring of terrestrial chlorophyll fluorescence from moderate-spectral-resolution near-infrared satellite measurements: methodology, simulations, and application to GOME-2, Atmospheric Measurement Techniques, 6, 2803–2823, doi:10.5194/amt-6-2803-2013, URL http://www.atmos-meas-tech.net/6/2803/2013/, 2013.