ESTIMATING NET PHOTOSYNTHESIS OF VEGETATION FROM SOLAR-INDUCED CHLOROPHYLL FLUORESCENCE – A MODELLING STUDY

Jochem Verrelst[1], *Christiaan van der Tol*[2], *Federico Magnani*[3], *Juan Pablo Rivera*[1], *Gina Mohammed*[4] *and Jose Moreno*[1]

[1]Image Processing Lab (IPL), University of Valencia, Spain; [2]Faculty ITC, University of Twente, The Netherlands; [3]Department of Agricultural Sciences, University of Bologna, Italy; [4]P&M Technologies, Canada

Jochem.verrelst@uv.es

ABSTRACT

Progress in imaging spectroscopy technology and data processing can enable derivation of the complete solar-induced chlorophyll fluorescence (SIF) emission spectrum from 640-800 nm. This opens up opportunities to fully utilize the SIF signal as an indicator of photosynthetic activity. A SCOPE modelling exercise was conducted to determine how strongly canopy-leaving SIF can be related to *net photosynthesis of the canopy* (NPC) for various simulated canopy configurations. This was done in two ways. First, a variance-based global sensitivity analysis (GSA) for both SIF and NPC was conducted; second, a regression analysis between SIF and NPC was performed. The GSA identified the SCOPE input variables that drive SIF and NPC. Largely the same key variables were identified, which explains why SIF can be related to NPC. Subsequent regression analysis between SIF retrievals and NPC values indicated that for heterogeneous canopy configurations (i.e. varying biochemistry, leaf, canopy variables) individual SIF bands most sensitive to NPC were located around the first emission peak (SIF_{red}).

Index Terms— Fluorescence, photosynthesis, global sensitivity analysis, retrieval, FLEX

1. INTRODUCTION

In preparation for ESA's Earth Explorer 8 candidate mission FLEX (FLuorescence EXplorer), a Photosynthesis Study (PS) has been completed that aimed to quantitatively link fluorescence to photosynthesis based on model and experimental data [1]. One of the objectives of the PS was to develop a prototype inversion algorithm to retrieve photosynthesis from simulated sun-induced fluorescence (SIF) observations. Another objective was to develop biochemical models into an existing soil-vegetation-atmosphere-transfer (SVAT) model. The model 'Soil-Canopy-Observation of Photosynthesis and the Energy balance' (SCOPE) has been selected as baseline model, because it has the ability to simulate the effects of irradiance, vegetation structure and physiology on SIF and photosynthesis [2]. During the PS, SCOPE was extended with various biochemistry submodels (TB12-D, TB12 and MD12) and was optimized to deliver a wide range of outputs (e.g., fluxes, reflectance, fluorescence).

In this study, the targeted flux is "*Net photosynthesis of the canopy*" (NPC), which is important for carbon cycle and climate change research. In order to enable estimation of NPC from optical data, a hybrid retrieval strategy that feeds SCOPE simulation outputs into a regression algorithm has been pursued. This approach facilitates the use of sun-induced fluorescence (SIF) data in retrieval of NPC for a multitude of theoretical canopy configurations. However, because SCOPE is a complex model that consists of over 30 input variables, a first step is to identify the key variables that drive canopy-leaving SIF. Therefore, we had the following objectives: (1) to apply a global sensitivity analysis (GSA) that quantifies the relative importance of SCOPE input variables to SIF; and (2) to assess the predictive power of SIF wavelengths to estimate NPC, particularly when there is variability in the driving vegetation variables.

2. SCOPE

SCOPE is a vertical (1-D) integrated radiative transfer and energy balance model [2]. It calculates radiation transfer in a multilayer canopy, in order to obtain reflectance and fluorescence in the observation direction as a function of the solar zenith angle and leaf inclination distribution. Irradiance and the distribution of absorbed radiation within the canopy is calculated with the Scattering by Arbitrarily Inclined Leaves (SAIL) model. The distribution of absorbed radiation is further used in a micrometeorological representation of the canopy for the calculation of photosynthesis, fluorescence, latent and sensible heat. The fluorescence and thermal radiation emitted by individual leaves is finally propagated through the canopy, again with the SAIL modelling concept [2].

Apart from the canopy radiative transfer modules, the following leaf-level modules are relevant:

1. A leaf radiative transfer sub-model (Fluspect) that calculates absorbed photosynthetically active radiation (APAR), reflectance and fluorescence spectra as a function of the irradiance spectrum and the leaf composition. The module calculates excitation-emission probability matrices for both sides of the leaf.
2. Coupled to the leaf sub-model, one mechanistic and two empirical biochemical sub-modules were implemented. These modules describe the relationship between photochemical yield and fluorescence yield empirically for nonstressed, ideal conditions under variable light and CO_2 concentrations (TB12), and for variable drought levels under high light conditions (TB12-D). The third, MD12, module [3] has a more explicit parameterization of fluorescence quenching mechanisms.

3. GLOBAL SENSITIVITY ANALYSIS

SIF-NPC relationships are assessed for various canopy configurations following the identification of variables which are drivers of SIF emission characteristics. To identify these driving variables, a global sensitivity (GSA) analysis was conducted in a related study [4]. Variance-based GSA explores the full input variable space and evaluates the relative importance of each input variable in a model [5]. The method can be used to identify the most influential variables affecting model outputs. In variance-based GSA, the contribution of each input variable to the variation in outputs is averaged over the variation of all input variables, i.e., all input variables are changed together [5]. In our related study [4], Saltelli's method [6] was used to identify the driving variables of canopy-leaving SIF spectrum across its full spectral range. The method has been demonstrated to be effective in identifying both the main sensitivity effects (first-order effects, i.e., the contribution to the variance of the model output by each input variable; S_i) and total sensitivity effects (the first-order effects plus interactions with other input variables; S_{Ti}) of input variables. In this study, the GSA analysis was extended to include the integrated SIF (F_{total}) from 640 to 850 nm and NPC. As such, the relative contribution of each input variable to SIF can be disentangled and quantified. The MD12 biochemical sub-model was used here with the leaf physiological variables qLs and kNPQs fixed, based on earlier performance comparisons to the TB12 modules [1]. Further, the full SCOPE variable space was analyzed for a spherical leaf angle distribution and without varying soil variables. See [4] for details on the boundaries for each variable. The variables were sampled according to Sobol's quasi-random sequence generator. In total, $(N(k + 2))$ model simulations were run, where N is the sample size and equals 1000, and k is the number of input variables and equals 22. This leads to 24000 simulations. Only total order sensitivity effects (S_{Ti}) expressed as percentages were considered.

4. EXPERIMENTAL SETUP - SCOPE SIMULATIONS

Because SCOPE is a multi-scale SVAT model, analysis of input variables to the SIF-NPC relationships can be undertaken by considering varying variables at various scales, e.g., biochemical, leaf, canopy, geometry, and/or micrometeorological. A balance must be achieved between including a sufficient number of ranging variables to achieve good representation of reality but not so many variables that the escalating heterogeneity becomes uninterpretable.

Table 1. SCOPE configurations and ranging variables. See [4] for full name of variables.

	Variables	Justification	# Sim
Biochemistry			
1	V_{cmo}	V_{cmo} is the main biochemical driver of photosynthesis. Hence, this is the theoretical baseline when SIF is not influenced by any other variable.	2000
2	Biochemistry	All biochemical variables (V_{cmo}, m, Rdparam, kV). Represents the most heterogeneous situation at the biochemical scale.	2000
Biochemistry, leaf			
3	V_{cmo}, LCC	Driving biochemical and leaf variables.	2000
4	V_{cmo}, leaf	Driving biochemical variable and all leaf variables (N, Cw, Cdm, Cs, LCC).	2000
5	Biochemistry, leaf	All biochemical and leaf variables. Represents the most heterogeneous situation at biochemical and leaf scales (V_{cmo}, m, Rdparam, kV, N, Cw, Cdm, Cs, LCC).	2000
Biochemistry, leaf , canopy			
6	LCC, LAI	Driving leaf and canopy variables.	2000
7	V_{cmo}, LAI	Driving biochemical variable (V_{cmo}) with driving canopy variable (LAI)	2000
8	V_{cmo}, canopy	Driving biochemical variable (V_{cmo}) with all varying canopy variables (LAI, lw, hc).	2000
9	V_{cmo}, N, Cw, Cdm, Cs, LCC, LAI, hw, hc (spherical LIDF)	Driving biochemical variable (V_{cmo}) with all leaf and all canopy (N, Cw, Cdm, Cs, LCC, LAI, lw, hc).	2000
10	Biochemistry, leaf, canopy	All biochemical, leaf and canopy variables (Vcmo, m, Rdparam, kV, N, Cw, Cdm, Cs, LCC, LAI, lw, hc). Represents the most heterogeneous situation at the canopy scale	2000
All biochemistry, leaf , canopy, geometry, micrometeorology			
11	Key SCOPE variables driving SIF	V_{cmo}, m, Cdm, LCC, LAI, rwc, u, ea, Ca, Ta, Rin. See also Figure 1.	2000
12	All SCOPE variables	All SCOPE variables (V_{cmo}, m, Rdparam, kV, N, Cw, Cdm, Cs, LCC, LAI, lw, hc, VZA, RAA, SZA, rwc, rb, P, u, Oa, ea, Ca, Ta, Rin, Rli). Represents the most heterogeneous configuration.	2000

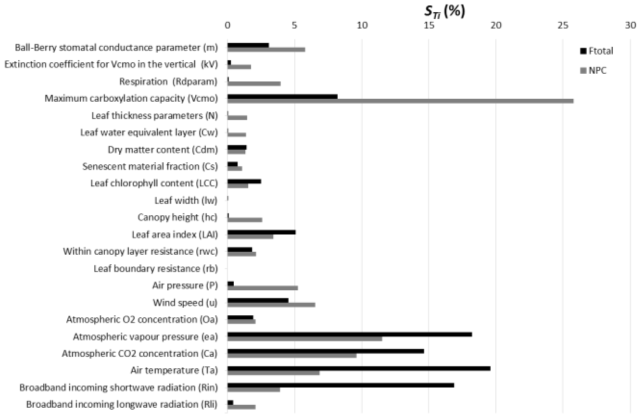
Various canopy configurations with increasing heterogeneity were generated. First, only the main biochemical variables were ranged, then more variables were ranged at biochemical, leaf, and canopy scales, eventually varying all variables at all scales. The key variables taken as a starting point were V_{cmo} at biochemistry scale, LCC at leaf scale, and LAI at canopy scale. The number of ranging variables was then increased to eventually produce a total of 12 canopy configurations, lastly incorporating all SCOPE variables (Table 1). Each variable was randomly sampled 2000 times within their minimum-maximum boundaries according to [4]. From the combined variable space, a uniform random subset of 2000 simulations was selected. If the radiative transfer equations were not successfully resolved, another random simulation was taken. Finally, from all SCOPE output variables the NPC output flux, the SIF spectra, and F_{total} were collected.

6. RESULTS

6.1. GSA results

Figure 1 provides the S_{Ti} results of the SCOPE input variables for F_{total} and NPC. Regarding F_{total}, most input variables exerted a rather small or even negligible effect. The driving variables were: maximum carboxylation capacity (V_{cmo}), Ball-Berry stomatal conductance (m), dry matter content (Cdm), chlorophyll content (LCC), leaf area index (LAI), canopy height (hc), within-canopy-layer resistance (rwc), wind speed (u), atmospheric vapour pressure (ea), atmospheric CO_2 concentration (Ca), air temperature (Ta), and broadband incoming shortwave radiation (Rin). Altogether these 11 input variables explained 96.1% of the total variance (taking interactions into account). These are the F_{total} key variables as shown in Table 1 (case 11).

Figure 1. Driving variables of F_{total} as identified by Global Sensitivity Analysis.



Regarding NPC, the most important variable is V_{cmo}, which is understandable as that is the variable that controls photosynthetic capacity. In addition, many other variables play a role in driving NPC. These include the same atmospheric variables as F_{total} (ea, Ca, Ta, Rin, u), but additionally air pressure (P), and to a small extent broadband longwave radiation (Rli) and atmospheric O_2 concentration. Contrary to F_{total}, in general all leaf and canopy variables influence NPC, but especially Ball-Berry stomatal conductance (m), respiration (Rdparam) and canopy height (hc). The overlap in driving variables between SIF and NPC explains why it is possible to establish relationships. Consequently, relationships towards NPC may not occur directly through the variable regulating photosynthetic activity (V_{cmo}), but rather through secondary relationships, e.g. through detected spatial variations in leaf and canopy surface variables. On the other hand, the GSA analysis of both F_{total} and NPC also demonstrates that driving variables are not fully identical. There are thus no reasons to assume that detection of SIF is directly related to capture of NPC in every circumstance. While relationships can be constructed, spatially-explicit retrieval of key variables is required to disentangle the information content related to photosynthetic activity.

6.2. Single band analysis SIF-NPC relationships

The possibility of single SIF bands to correlate with NPC was subsequently assessed. A linear regression was looped over the SIF broadband range (640-800 nm) and the R^2 of validation data is plotted in Figure 2. Results are organized according to ranging variables at the scales of biochemical, leaf, canopy, and at all SCOPE scales. The following trends are observed:

When ranging only variables at *biochemical scale (*Fig. 2a), V_{cmo} was the main variable that drives NPC [7] and resulted in a very strong relationship (R^2 of 0.99). Relationships with NPC weakened when the other biochemical variables (*m*, *Rparam*, *kV*) were varied but impacts were spectrally invariant.

When SCOPE variables were ranged for the *leaf scale* (Fig. 2b), varying the driving variables leaf chlorophyll content (LCC) and V_{cmo} produced strong relationships with NPC. However, these relationships degraded at wavelengths beyond the red emission peak. The relationship degraded further when additionally varying other leaf variables, and a clear distinction between the first and second peak can be observed. When also ranging all biochemical variables, correlations weakened further but the difference between both peaks is less prominent.

When SCOPE variables were ranged for the *canopy scale* (Fig. 2c), strong spectrally invariant relationships were obtained only in the case of V_{cmo} plus LAI and other canopy variables. The relationship broke down for the NIR emission peak when leaf variables were ranged. Also strong relationships can be obtained when varying only the driving leaf and canopy variables LCC and LAI. But this relationship also broke down for the second peak. When simulating fully heterogeneous canopies, and also ranging other biochemical variables then correlations degraded to such an extent that meaningful relationships could not be derived.

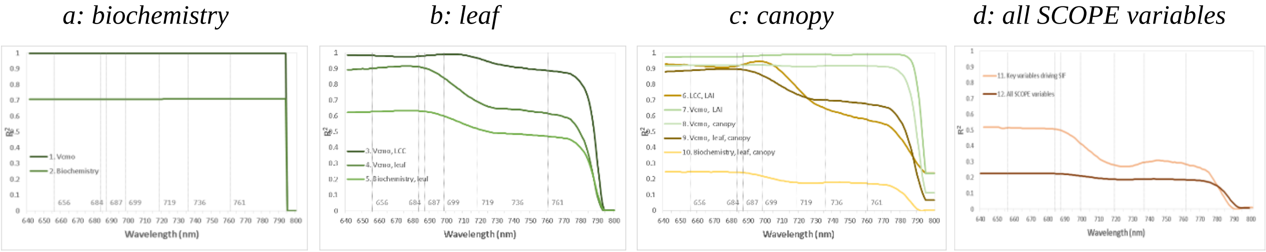

a: biochemistry
b: leaf
c: canopy
d: all SCOPE variables

Fig. 2: Regression analysis results (R^2) of single wavelengths across the 650-800 nm spectral range for different canopy configurations. The gray lines represent important wavelengths (absorption lines, SIF peaks, mid-valley).

Combining input variables at *all SCOPE scales* (Fig. 2d), i.e., including also micrometeorological variables, caused relationships to deteriorate, especially for the second peak. When considering the driving SIF variables as identified by the GSA exercise, the first peak achieves a maximal R^2 of about 0.5, whereas for the second peak it does not exceed about 0.3. Given that this dataset is mainly generated from varying micrometeorological variables, results suggest that these variables tend to weaken SIF-NPC relationships. Therefore, varying all SCOPE variables no longer produced meaningful relationships between SIF and NPC. It underlines the importance of applying a GSA to constrain the number of input variables in an optimized way.

8. CONCLUSIONS

A SCOPE modelling simulation was conducted to examine how successfully canopy-leaving SIF can estimate *net photosynthesis of the canopy* (NPC). To analyze the influence of biochemical, leaf and canopy mechanisms impacting the SIF signal, a global sensitivity analysis (GSA) was conducted. Based on the identified key variables subsequently multiple canopy configurations were simulated. Regression analyses between SIF retrievals and NPC values led to the following general finding: (1) the most sensitive SIF bands to NPC were located around the first (i.e. red) emission peak for heterogeneous canopy configurations. Further, although results not shown here, also band combinations were evaluated, as well the use of a nonlinear regression algorithm. That analysis led to the following general findings: (2) combining two SIF retrieval bands (e.g., O_2-B and O_2-A) led to stronger correlations than using only one SIF band; (3) using the O_2-B and O_2-A bands produced similar or superior performances than using the two emission peaks, while using the peak ratio produced poorer relationships than when both bands were individually entered into the regression model; (4) even stronger correlations were achieved using four main SIF retrieval bands (Hα, O_2-B, water vapour, O_2-A); and (5) nonlinear regression produced stronger relationships than did linear approaches.

The theoretical results generated here can serve as a reference for future SIF-based photosynthesis studies. Overall, it is recommended that the SIF signal in at least the O_2-B and O_2-A bands be sampled to enable robust quantification of canopy photosynthetic activity.

11. REFERENCES

[1] G.H. Mohammed, Y. Goulas, F. Magnani, J. Moreno, J. Olejníčková, U. Rascher, C. van der Tol, W. Verhoef, A. Ač, F. Daumard, A. Gallé, Z. Malenovský, D. Pernokis, J.P. Rivera, J. Verrelst, and M. Drusch, 2012 "FLEX/Sentinel-3 Tandem Mission Photosynthesis Study. Final report," ESTEC Contract No. 4000106396/12/NL/AF. 2014.

[2] C. van der Tol, J.A. Berry, P.K.E. Campbell, and U. Rascher, "Models of fluorescence and photosynthesis for interpreting measurements of solar induced chlorophyll fluorescence," *Journal of Geophysical Research*, 10.1002/2014JG002713. 2014.

[3] F. Magnani, A. Olioso, J. Demarty, V. Germain, W. Verhoef, I. Moya, Y. Goulas, G. Cecchi, G. Agati, P. Zarco-Tejada, G. Mohammed, C. van der Tol, "Assessment of Vegetation Photosynthesis Through Observation of Solar Induced Fluorescence from Space," ESTEC Contract No. 20678/07/NL/HE, Final Report. 2009.

[4] J. Verrelst, J.P. Rivera, C. van der Tol, F. Magnani, G. Mohammed, and J. Moreno, "Global sensitivity analysis of SCOPE v1.53 model: what drives canopy-leaving sun-induced fluorescence?" Remote Sensing of Environment. Submitted. 2015.

[5] A. Saltelli, S. Tarantola, K.P.S. Chan, "A quantitative model-independent method for global sensitivity analysis of model output," Technometrics 41, pp. 39–56. 1999.

[6] A. Saltelli, P. Annoni, I. Azzini, F. Campolongo, M. Ratto, S. Tarantola, "Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index," Computer Physics Communications 181, pp. 259–270, 2010.

[7] Y. Zhang, L. Guanter, J.A. Berry, J. Joiner, C. van der Tol, A. Huete, A. Gitelson, M. Voigt, P. Köhler, 2014. Estimation of vegetation photosynthetic capacity from space-based measurements of chlorophyll fluorescence for terrestrial biosphere models. Global Change Biology. 20, pp. 3727-3742.