

2010). **New rotational-Raman scattering calculation** does not depend upon these parameters, again because there is no significant absorption from  $O_2$  within this dark solar line.

We also computed the filling-in of the K I line due to rotational-Raman scattering (RRS) using the scalar Linearized Discrete Ordinate Radiative Transfer code (LIDORT-RRS) of Spurr et al. (2008). Although the amount of atmospheric scattering is relatively small at these wavelengths, the filling-in effect of rotational-Raman scattering is apparent owing to the large depth of the K I line. The spectral signature is very similar to that of line filling due to fluorescence. The amount of filling-in depends upon solar zenith angle,  $\theta$ , as well as surface albedo. For low to moderate  $\theta$  (up to about  $50^\circ$ ) and typical surface albedos (0.3), this RRS filling-in is about 0.2% at GOSAT spectral resolution. For lower surface albedo, the filling-in increases (approximately 0.4% at albedo of 0.1). Fig. 3 shows the filling-in at  $\theta=70^\circ$ . For typical surface albedo, the value is close to 0.4% that corresponds to a filling-in from fluorescence less than about  $0.5 \text{ mW/m}^2/\text{sr/nm}$ .

In our analysis of GOSAT data, we use both Earth and solar spectra. Due to Doppler shifts and other effects, it is necessary to spectrally align Earth and solar spectra. When shifting one spectrum with respect to another, interpolation is required. If spectra are not adequately sampled, as is the case for most space-based instruments, this interpolation will produce errors (Chance et al., 2005). We have simulated this effect based on the approach of Chance et al. (2005) using spline interpolation. Our results indicate that we should expect a significant spectral effect due to undersampling in GOSAT spectra.

Finally, we simulated the effects of  $O_2$  absorption on GOSAT observations within the spectral region near the K I solar line. Here, we computed spectra convolved with the GOSAT ILSF without  $O_2$  absorption and with absorption at  $20^\circ$  and  $70^\circ$  solar zenith angles. We then examined ratios of these spectra. Our calculations did not show any significant spectral structure due to  $O_2$  absorption within the K I solar line. Spectral structure is shown between stronger  $O_2$  lines (away from the solar K I line) when comparing spectra with and without  $O_2$  absorption.

**New under-sampling discussion**

**$O_2$  absorption effects, new discussion, simulations (we will also fix the misspelling of absorption).**

## 4 Approach

More detailed  
radiative transfer

Figure 4 shows both observed (normalized) solar irradiance and sample Earth-view radiance spectra from the TANSO-FTS band 1 and a zoom in at the wavelengths used here for fluorescence retrieval (bottom panel). Resolved lines within the O<sub>2</sub>-A band are apparent in the Earth spectrum. Solar Fraunhofer lines are seen in both the solar and Earth spectra. Note that there may be small errors in absolute wavelength calibration and that the asymmetric ILSF also produces an apparent wavelength shift between observations and exact positions of solar Fraunhofer lines. There may also be errors in the absolute radiometric calibration as the instrument radiometric sensitivity has decreased by about 10% from pre-flight values in band 1 and vicarious calibration is required for correction.

The potassium (K) I line is one of the deepest solar Fraunhofer lines within GOSAT's band 1 (~755–775 nm). Although there are several other Fraunhofer lines lying closer to the peak of the 740 nm fluorescence peak, the K I line was found to be a good starting point for observing chlorophyll fluorescence owing to its great line depth and lack of terrestrial oxygen absorption. The K I line falls between several weak O<sub>2</sub> lines and can be observed in relative isolation with GOSAT's high spectral resolution.

The satellite-observed Earth spectral radiance,  $I(\lambda)$ , within our spectral region of interest (769.90–770.25 nm), can be approximated using a Lambertian-equivalent reflectivity ( $R$ , assumed constant over our spectral fitting window) surface model and fluorescence,  $F$ , also here assumed to be spectrally constant, i.e.,

$$I(\lambda) = \left( I_o(\lambda) + \frac{RT_d(\lambda)T_u(\lambda)}{1 - RS_b(\lambda)} \right) \frac{E(\lambda) \cos \theta}{\pi} + \frac{FT_u(\lambda)}{1 - RS_b(\lambda)} \quad (1)$$

where  $I_o$  is the radiance (solar-normalized) due to Rayleigh scattered light in the absence of surface reflectance and fluorescence,  $T_d$  is irradiance (flux) transmittance (direct + diffuse),  $S_b$  is the component of the reflected surface radiance that is reflected by the atmosphere back to the surface,  $T_u$  is atmospheric transmittance of the surface

radiance at the top of the atmosphere, and  $E(\lambda)$  is the extraterrestrial solar irradiance. This equation holds for monochromatic radiance.

With GOSAT, we have radiance and irradiance measurements (in two polarizations), i.e., monochromatic radiance and irradiance convolved with the ILSF,  $I^*(\lambda)$  and  $E^*(\lambda)$ , respectively, where  $*$  represents convolution with the ILSF. We now show the effects of different simplifying assumptions. We first assume that the effects of atmospheric scattering and absorption are negligible on our fitting window (i.e.,  $S_b = 0$ ,  $I_o = 0$ ,  $T_u = 1$ , and  $T_d = 1$ ). With these assumptions, the observed total intensity,  $I^*$ , where we now drop the  $(\lambda)$ , can be simplified to

$$I^* = \left( \frac{RE \cos \theta}{\pi} + F \right)^* \quad (2)$$

$$= KE^* + F. \quad (3)$$

Similarly, if we assume that atmospheric scattering and absorption is spectrally constant over the wavelengths of interest, then the observed intensity can be rewritten as

$$I^* = \left( \left[ I_o + \frac{RT_d T_u}{1 - RS_b} \right] \frac{E \cos \theta}{\pi} \right)^* + \frac{FT_u}{1 - RS_b} \quad (4)$$

$$= K' E^* + \epsilon F. \quad (5)$$

Eq. 5 is similar in form to Eq. 3. Note that the value of  $K$  (or  $K'$  or  $R$ ) is of no importance for determining  $F$ . Similarly, a constant calibration offset in the solar irradiance will not affect the derived fluorescence as the error will be absorbed in the retrieved value of  $K$  (or  $K'$ ). Neglect of scattering and absorption will produce a slight scaling error,  $\epsilon$ , in  $F$  (for  $R = 0.3$ , the error is approximately 0.6%).

Since we observe both  $I^*$  and  $E^*$ , there is no explicit use of the ILSF when deriving  $F$  using Eq. 5. Therefore, our approach will not suffer from an imperfect model of the ILSF. Furthermore, if  $\epsilon \simeq 1$ , then radiative transfer calculations are not necessary to derive  $F$ .

More discussion on assumptions made and their impact.

The use of the K I line therefore substantially simplifies a retrieval of  $F$  as compared with the O<sub>2</sub>-A band. The fluorescence signal is not diminished by atmospheric absorption as it is in the O<sub>2</sub>-A band. With GOSAT's spectral resolution, SNR, and large footprint that ensures capture of substantial photon emissions, our calculations, and those conducted independently by Frankenberg et al. (2011), indicate that detection of ecosystem-scale chlorophyll fluorescence can be achieved.

We use a standard weighted least squares (minimum variance) fitting to estimate the state vector,  $\mathbf{x}$  consisting of  $K$  (or  $K'$ ),  $F$ , and the wavelength shift between solar and Earth spectra, i.e.,

$$\mathbf{x} = (\mathbf{H}^T \mathbf{S}_o^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{S}_o^{-1} (\mathbf{y}_{\text{obs}} - \mathbf{y}_{\text{calc}}), \quad (6)$$

where  $\mathbf{H}$  is the Jacobian matrix,  $\mathbf{S}_o$  is the observation error covariance, and  $\mathbf{y}_{\text{obs}}$  and  $\mathbf{y}_{\text{calc}}$  are observed and calculated radiances within our fitting window. We use a diagonal observation (in radiance units) error covariance matrix (i.e., this assumes that observational errors are Gaussian with zero bias and uncorrelated with each other). There are two weak O<sub>2</sub> lines located within the window, as shown in Figs. 2 and 4. We effectively remove wavelengths affected by these lines by assigning large errors or no weight to them. All other wavelengths are weighted equally such that the retrieval does not depend upon the absolute radiance error. The derived spectral residual term is implicitly included in the Jacobian.

For  $E^*$ , we use several GOSAT solar spectra taken closest in time to Earth radiance observations (within a few hours) averaged together to produce a composite spectrum. Before the spectral fit, we first perform a rough adjustment to spectrally align the solar and Earth spectra. We added a wavelength shift term to the fit to further improve the alignment. This term can make the minimization non-linear. However, we found that by providing an initial shift, the problem was sufficiently linear that multiple iterations were not necessary.

Here, we assume a constant wavelength dependence of  $F$  and  $R$  (or  $K$  or  $K'$ ) over this band. This assumption could be relaxed as would be appropriate for a larger spectral fitting window as discussed by e.g., Meroni et al. (2010). We could also account for

More discussion on fitting procedure, reference to Frankenberg et al., 2011 here (and elsewhere)

atmospheric scattering and absorption using a radiative transfer code, though we have not done so here. Note that an error in absolute calibration of  $I(\lambda)$  will be present in the retrieved  $F$ .

Our initial results showed positive values of fluorescence over the Sahara where none was expected. The root cause was found to be unexplained systematic spectral structure in the core of the K I line that we believe is due in part to undersampling and the effects of rotational-Raman scattering.

To account for these affects, we derived averaged spectral signatures from pixels where we expect zero fluorescence. Here we used clear-sky pixels from several days over the Sahara. We derived these residuals separately for several binned values of the derived spectral shifts between Earth and solar spectra. We found that the overall spectral shapes were similar for the various shift values, but differed somewhat in the details. We then incorporate the derived residuals for the appropriate value of the wavelength shift into our spectral fit. Note that this approach does not account for the  $\theta$  dependence of the RRS filling-in. This will result in an overestimate of fluorescence filling-in at high  $\theta$  as the RRS filling-in will not be completely accounted for at these large solar zenith angles.

Samples of the overall fits to normalized reflectance spectra in the P polarization are shown in Fig. 5 for single pixels over active vegetation in the Eastern United States (left) and Sahara (right). Examples of the applied residuals for reflectance spectra are shown in Fig. 6. The fitted filling-in structure due to fluorescence ( $F/KE(\lambda)$  in %) is also shown in Fig. 6 along with the observed minus fitted radiances (in %). For the wavelengths used in the fits, the standard deviation of the fit is generally within about  $\pm 0.3\%$  at the continuum wavelengths, consistent with the expected GOSAT SNR. Residuals are slightly higher ( $\sim 1\%$ ) in the core of the K line where there is higher sensitivity to undersampling and wavelength errors.

New fitting procedure  
described

We also show the analogous data for the Enhanced Vegetation Index (EVI) (Huete et al., 2002), a widely used reflectance-based vegetation index, to indicate relative greenness and to infer photosynthetic function, that was obtained over the same time periods from NASA's Aqua MODIS sensor (Fig. 11). The Aqua satellite has an ascending node equator crossing near 13:30 LT, similar to that of GOSAT. EVI data are provided as monthly mean values at  $0.05^\circ$  spatial resolution. We have averaged the data to be compatible with our coarser grid. We found that the standard deviations in EVI for our large grid boxes were fairly high, between about 0.05 and 0.15, for the majority of the vegetated gridboxes. This indicates that there is much spatial heterogeneity in green biomass content within the large grid boxes. Although there are similarities between the scaled  $F$  and EVI values owing to the dependence of both on the amount of green biomass seen by the sensors, some obvious differences are also evident.

Figures 12 and 13 show the full seasonal cycle in scaled  $F$  for the P and S polarizations, respectively, as well as EVI and the normalized difference vegetation index (NDVI) (Tucker, 1979) for a few selected regions. Both standard deviations and error of the mean are shown for the monthly mean  $F$ . Variability that contributes to the standard deviations includes that from retrieval error, the diversity of cover types included within the regions, and natural daily variability.

It should also be noted that the sampling is different for GOSAT and the MODIS EVI. Since the MODIS native spatial resolution is 1 km or better, there is a higher chance of seeing in between clouds. Different cloud screening techniques also contribute to sampling differences.

We use standard error propagation to compute the retrieval error covariance,  $S_r$ , i.e.,

$$S_r = (\mathbf{H}^T \mathbf{S}_o^{-1} \mathbf{H})^{-1}, \quad (8)$$

Using our assumption of a diagonal error covariance with all wavelengths weighted equally (removed wavelengths affected by absorption not included), Eq. 8 simplifies to

$$S_r = \sigma^2 (\mathbf{H}^T \mathbf{H})^{-1}, \quad (9)$$

Precision  
calculation

More discussion on  
sources of scatter

where  $\sigma$  is the radiance error standard deviation that we assume here is approximated by  $I/\text{SNR}$ . With this assumption (note that this is not a good assumption for all types of instrumentation), the error standard deviation in  $F$  is proportional to  $\sigma$  and thus proportional to  $\cos(\theta) \cdot R/\text{SNR}$ . The expected error in scaled  $F$  is then proportional to the  $R/\text{SNR}$ , and the error in terms of % of the continuum radiance goes as  $1/\text{SNR}$ . For our spectral window, we find that the precision of scaled  $F$  ranges from approximately 0.18 for  $\text{SNR}=300$  and  $R=0.2$  to 1.33 for  $R=0.5$  and  $\text{SNR}=100$  for a single measurement in one polarization. For our GOSAT fitting window, the error in terms of % of the continuum is approximately 0.453%(1.35%) for  $\text{SNR}$  of 300(100) for unpolarized measurements.

The actual scaled  $F$  retrieval standard deviations, that encompass both retrieval error and natural variability, are generally between about 0.6 and 1.2 for moderate values of reflectivity (up to 0.5) and higher over the very bright Sahara (between about 1.2 and 2.0). The Sahara results are more consistent with a lower  $\text{SNR}$  close to 100, consistent with our computed spectral residuals in the core of the K line.

As expected, over the Eastern US, a clear increase in scaled  $F$  is seen during the spring months (April and May) with a peak in early summer (June/July). This seasonal cycle is also seen in the EVI that has a slightly greater seasonal contrast (maximum vs. minimum) than scaled  $F$ .

We next examine two Amazonia regions, one near the East coast of Brazil and one more towards the west in the central Amazon rainforest region (Figs. 12-13, panels 2–3). In the westernmost Amazon pixel (panel 2), the dry season (July–November) green up reported by Huete et al. (2006) for 2005 is seen here between September 2009 and January 2010 in both scaled  $F$  and EVI, with scaled  $F$  leading by approximately one month beginning in August. In this region, trees with deep roots may have more continuous access to deep soil moisture (e.g., Nepstad et al., 1994). Here, EVI follows PAR which depends primarily on cloud cover and peaks during the dry season (Huete et al., 2006). However, this temporal pattern is not captured in the NDVI; NDVI values peak in June/July for this region and decline to a minimum in December. In the other

More discussion on errors due to instrument noise.

regions shown here, the NDVI is qualitatively similar (in the seasonal cycle) to the EVI.

In the eastern Amazon area (Figs. 12-13, panel 3), a different pattern is seen with a brown-down in the early part of the dry season in EVI, NDVI, and scaled  $F$ , consistent with analysis over grassland pasture sites in previous years by Huete et al. (2006). In this region, forest conversion has occurred where deep-rooted trees have been replaced with shallow-rooted vegetation that responds more to soil moisture stress. The steepest decline in scaled  $F$  occurs from April to May, while a sharp decline in EVI is seen from June through September. The scaled  $F$  shows an increase starting in September, while the EVI increase begins in October.

Over Indonesia (Figs. 12-13, panel 6), there is little seasonal change in EVI (and NDVI), although scaled  $F$  shows a significant minimum in June during the dry season when biomass burning typically takes place. El Niño began around this time in 2009 leading to a drier than normal season; widespread haze due to fires was observed in the region from June through September. Logging and forest conversion have occurred here, resulting in more vegetation stress during the dry season, especially during dry years associated with El Niño.

In the north of Australia (Figs. 12-13, panel 7), EVI (and NDVI) generally peaks in the austral autumn owing to the summer monsoonal rains. In contrast, we see a more broad peak in  $F$  starting earlier. Higher spatial scale studies show significant variability in NDVI in this region as well as interannual variability (e.g. Martin et al., 2009). A different pattern is seen in the EVI in the south western region of Australia with a distinct peak in the spring (Fig. 12, panel 9). This pattern is not seen in the derived fluorescence. These regions warrant further study and investigation into potential effects of sub-pixel variability and sources of error in the satellite-derived products.

Similar seasonal cycles are obtained in P and S polarizations. Slightly higher values of scaled  $F$  are obtained in the S polarization that has higher throughput in band 1. The standard deviations are also higher in the S polarization. As the temporal and spatial patterns for P and S polarizations are very similar, all subsequent results will be shown only for the P polarization. Note that small differences in the filling-in of the K I line

More discussion on polarization (new figure showing season cycle in S polarization)

for P and S polarizations can be caused by differences in rotational Raman scattering filling-in that will have different dependences on solar zenith angle.

We next compare and contrast the monthly gridbox averages for the derived and scaled  $F$  with those for EVI, for the eight geographic regions outlined in Fig. 11. The scaling of  $F$  accounts for variations in  $F$  related to incoming PAR while the vegetation indices can serve as a proxy for FPAR. Details are seen in the scatter diagrams of Figs. 14–15.

All of these plots show a large amount of scatter in the relationship between scaled  $F$  and EVI. This scatter results from retrieval error both in the EVI and scaled  $F$ , spatial and temporal sampling differences in the two derived quantities arising in part from sub-gridbox heterogeneity, and natural variability between the two quantities. As shown in Fig. 12, the standard deviation in scaled  $F$  for a relatively small region is about 1 unit. The error of the mean then depends upon the number of observations for a given gridbox. Over most moderately to heavily vegetated areas, the monthly-averaged  $F$  error of the mean is less than about 0.3 for most pixels with a small fraction having errors up to 0.5. The scatter can also be attributed to the fact that scaled  $F$  encompasses a relationship with instantaneous photosynthesis that may be complex rather than a more straight-forward relationship with potential photosynthesis represented by the vegetation indices.

These continental or large regional blocks obviously include a variety of vegetation types such as forests, grasslands, and croplands that have different  $F$  and EVI values. Even though the plots show significant scatter, for most regions there is an evident positive but non-linear relationship between  $F$  and EVI, but with a great deal of scatter.

This relationship is similar in shape and intensity value ranges in July in the Northern Hemisphere (Fig. 14, panels 1, 2, 3, 4). In contrast, these same North American regions exhibit very low  $F$  ( $< \sim 1$ ) across a range of EVI values (up to  $\sim 0.4$ ) in December (Fig. 15). The points with EVI near 0.4 may indicate low photosynthetic activity by standing green vegetation (e.g., coniferous forests) as has been suggested by e.g. Soukupová et al. (2008). Another example where moderate EVI values (0.2–0.4) imply

parameters such as gross primary production (GPP) and light-use efficiency (LUE) (Damm et al., 2010).

We have not taken into account the directionality of the emitted fluorescence, although this is a reasonable approach since  $F$  has traditionally been assumed to be isotropic. There may also be errors in absolute calibration of the Earth radiance data as well as errors associated with our spectral fitting including the neglect of rotational-Raman scattering. Therefore, our reported absolute values must be interpreted with caution. The association of  $F$  with biomes, rather than whole continents or arbitrary large regions, will improve the understanding of this signal.

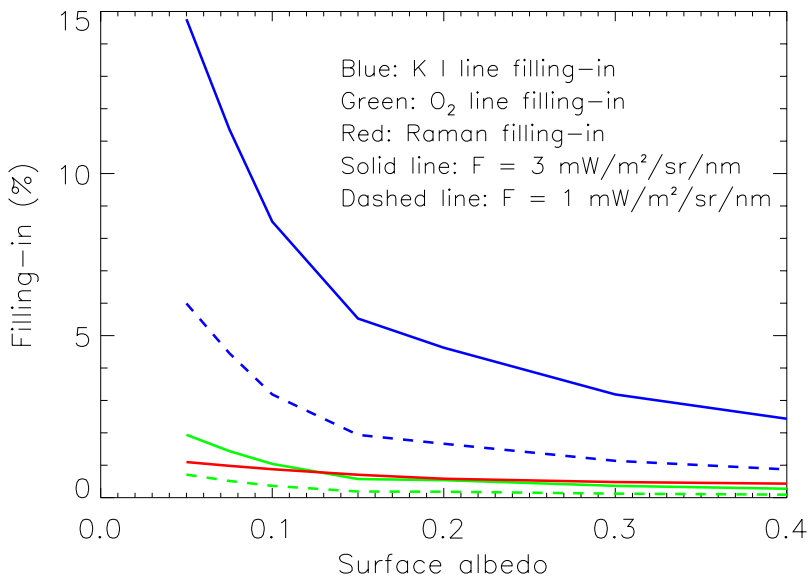
Clearly, future research will need to include developing appropriate validation strategies for these products. Comparisons with point measurements such as net  $\text{CO}_2$  uptake from flux towers present challenges as these measurements may not be representative of the much larger satellite footprints. However, comparisons of temporal variations may be useful. A longer time series of satellite-derived fluorescence will be helpful in conducting such comparisons. In addition, studies are needed to develop a better understanding of the physical linkages between  $F$  and other geophysical and observed (both by remote sensing and in situ) parameters. For example, recent studies have focused on the relationship between satellite-derived scaled Photochemical Reflectance Index (PRI) and light use efficiency (LUE) obtained from an eddy covariance flux tower (e.g., Drolet et al., 2005). We are currently processing more GOSAT data in an effort to map global complete seasonal cycles and to quantify interannual variations in fluorescence owing to events such as extreme drought.

We also plan to improve our retrieval by accounting for instrumental effects such as polarization sensitivity and wavelength calibration as described in Yoshida et al. (2010). A new version of GOSAT level 1B data with improved calibration in band 1 is expected to be released in the near future (A. Kuze, priv. comm., 2011), and we will examine whether these improvements affect our  $F$  retrievals.

Accounting for chlorophyll fluorescence in the K I line could assist in the interpretations made for  $F$  in the  $\text{O}_2$ -A band. Furthermore, this should also help to improve the

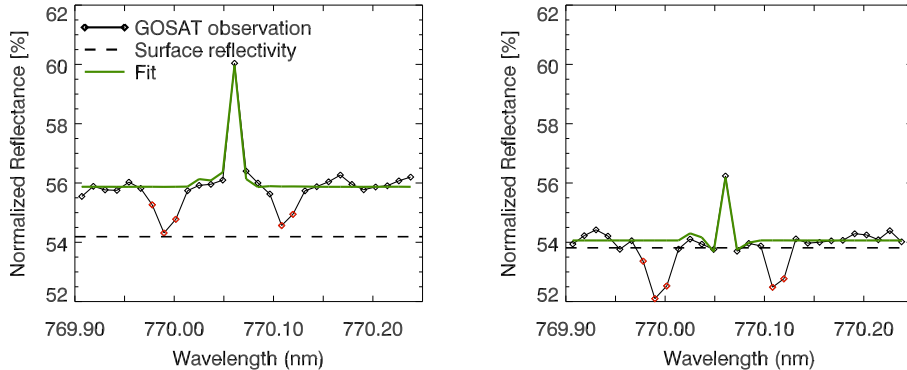
More discussion on validation and future improvements

Removed O<sub>2</sub> absorption, added filling-in due to rotational-Raman scattering



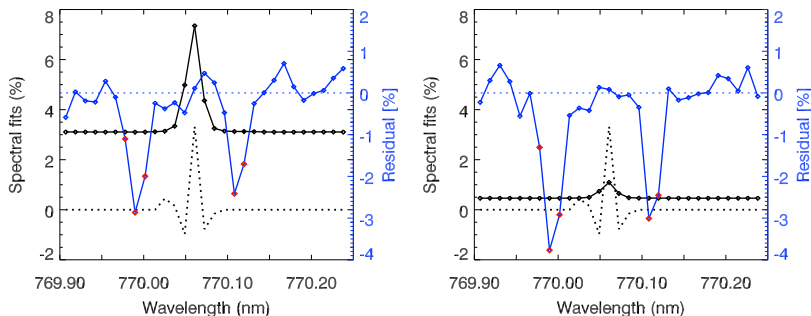
**Fig. 3.** Simulated GOSAT filling-in of the O<sub>2</sub> weak line at 769.9 nm and the K I line as a function of surface albedo. Filling-in of the K I line due to both fluorescence and rotational-Raman scattering is shown.

Now show actual GOSAT wavelengths (not interpolated),  
new fitting procedure used.



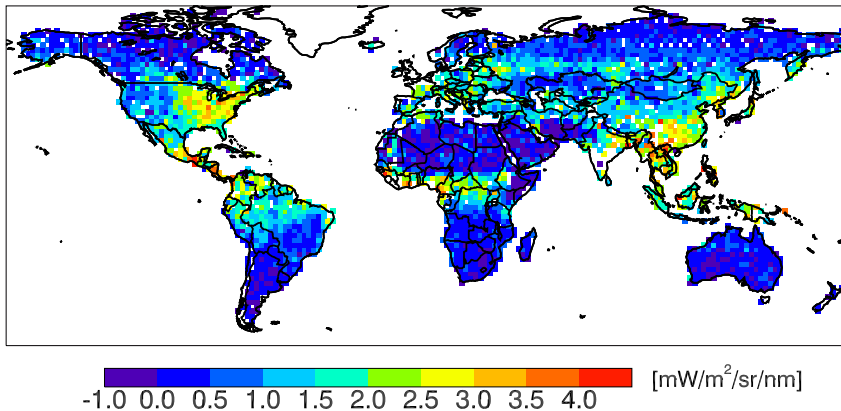
**Fig. 5.** Normalized reflectance (Earth radiance normalized by the solar spectra in the GOSAT P polarization, black and red diamonds; red are not used in spectral fit), spectral fit (green solid line), and derived reflectivity (black dashed line) in the vicinity of the K I Fraunhofer line showing contrast between active vegetation over the Eastern United States (left) and the Sahara (right) on 2 July and 1 July 2009, respectively.

Now show actual GOSAT wavelengths (not interpolated),  
new fitting procedure used.



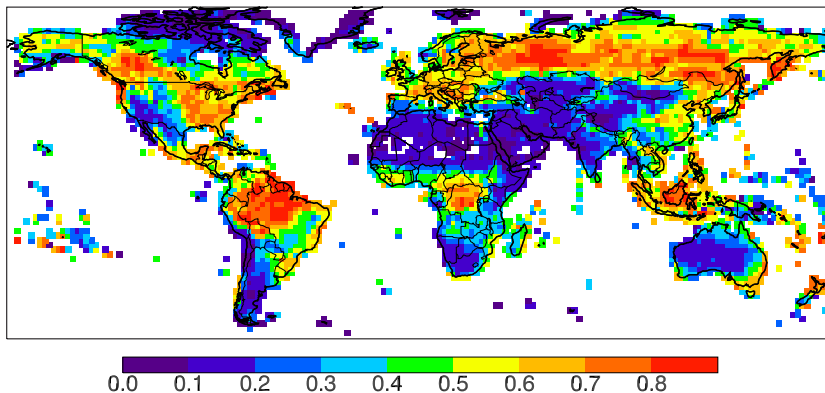
**Fig. 6.** Spectral structure as a percent of the continuum reflectance (left axis): Dotted line is unexplained spectral structure (assumed due to undersampling and rotational-Raman scattering) and black solid line and diamonds represent the spectral fit of the fluorescence. The radiance residuals (observed minus computed) are shown by the blue solid line, where blue diamonds are for fitted wavelengths and red are for those not fitted. These are the same two pixels shown in Fig. 5 (Eastern US on left and Sahara on right). Results are for the P polarization.

Now shown up to highest latitudes, new fitting procedure. Scale now starts from -1.



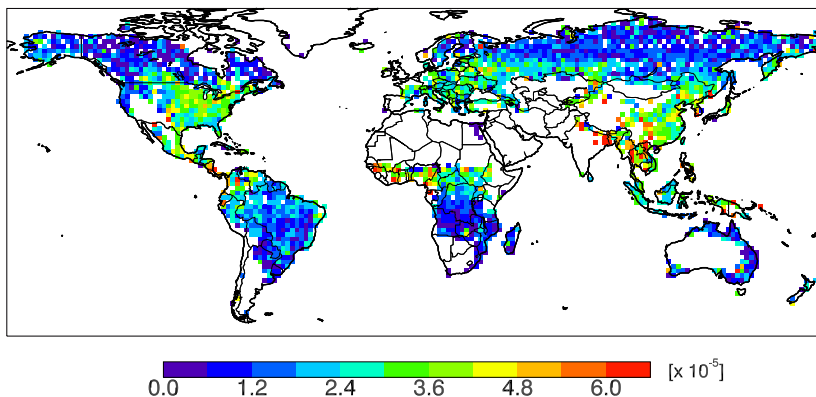
**Fig. 7.** Derived monthly averages for instantaneous fluorescence  $F$  (mW/m<sup>2</sup>/sr/nm) from GOSAT for July 2009 from P polarization.

Now shown up to highest latitudes, new fitting procedure.  
Geolocation errors fixed.



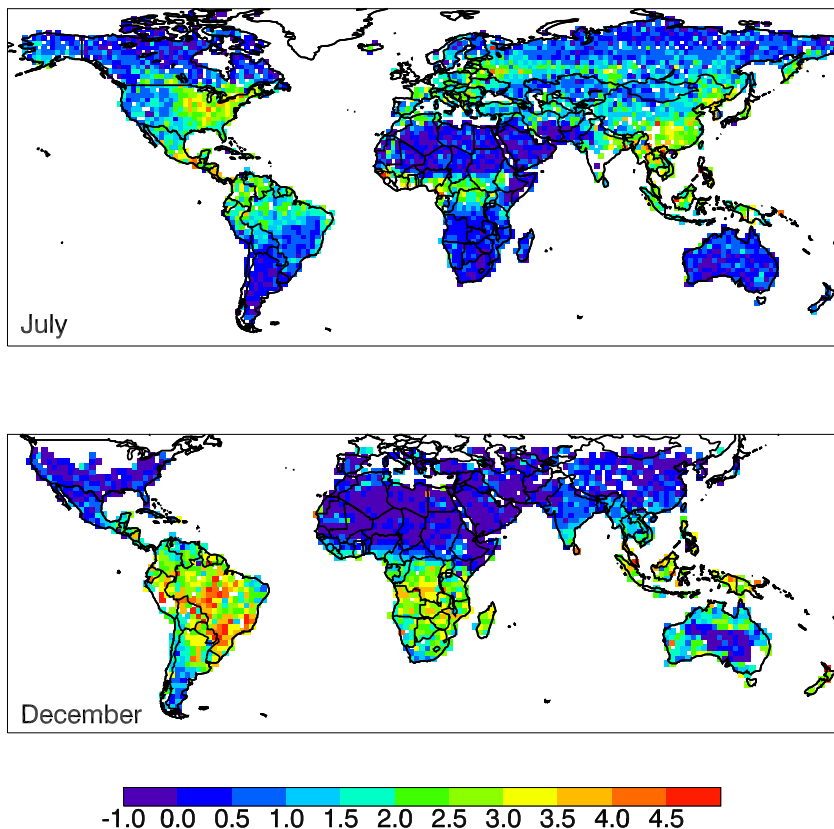
**Fig. 8.** 8-day FPAR product (unitless) from MODIS for 20–27 July 2009.

Now shown up to highest latitudes, new fitting procedure. Fixed errors due to geolocation error in FPAR.



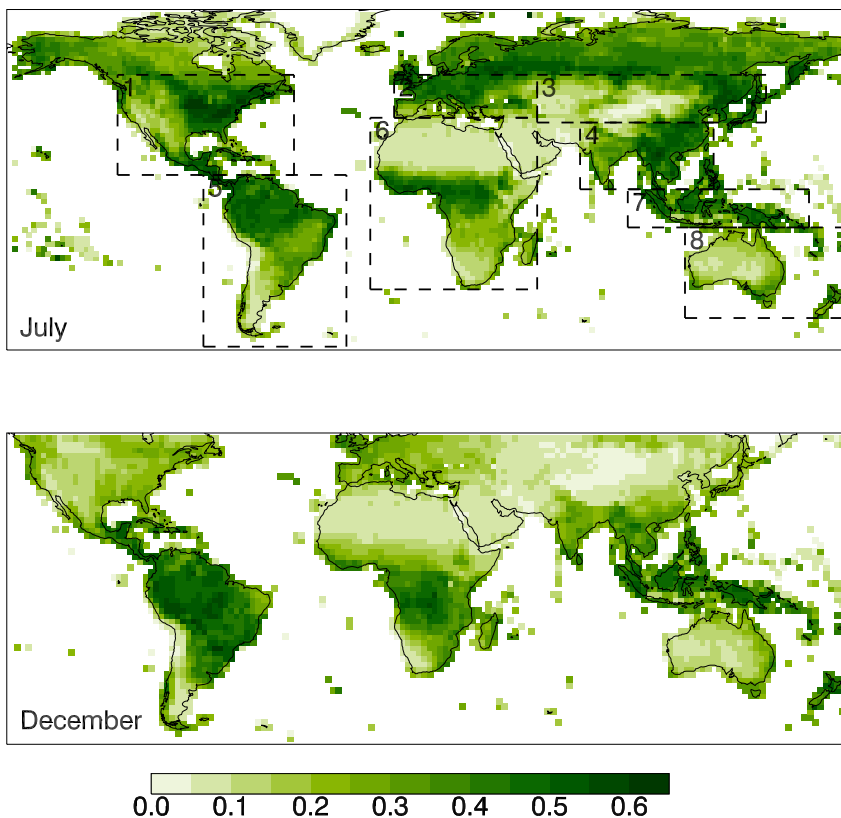
**Fig. 9.** Derived monthly averages for instantaneous fluorescence yield,  $\Phi$ , (unitless) for July 2009 from the GOSAT P polarization.

Now shown up to highest latitudes, new fitting procedure.  
Color scale now starts from -1.



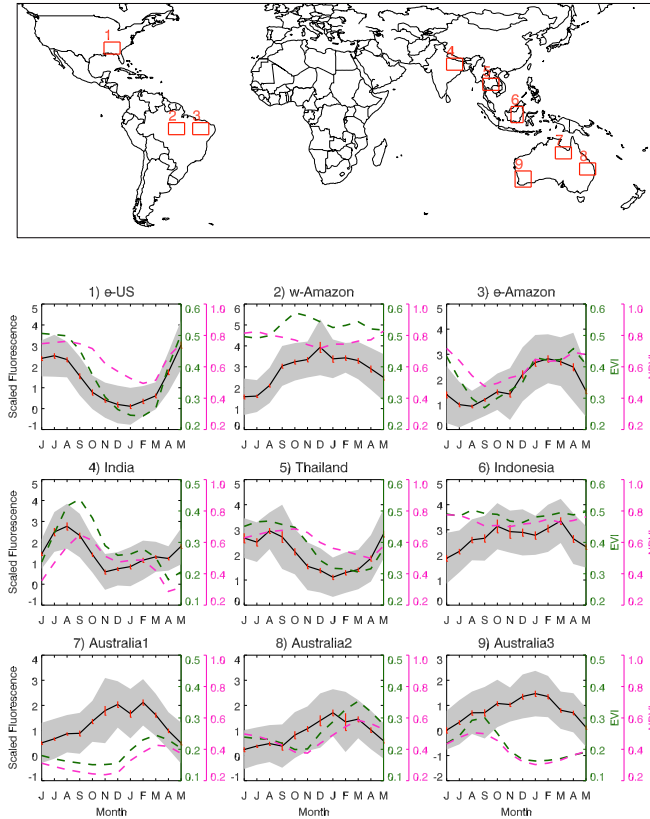
**Fig. 10.** Derived monthly averages for scaled fluorescence (unitless) from GOSAT for July and December 2009 for the P polarization.

Now shown up to highest latitudes.



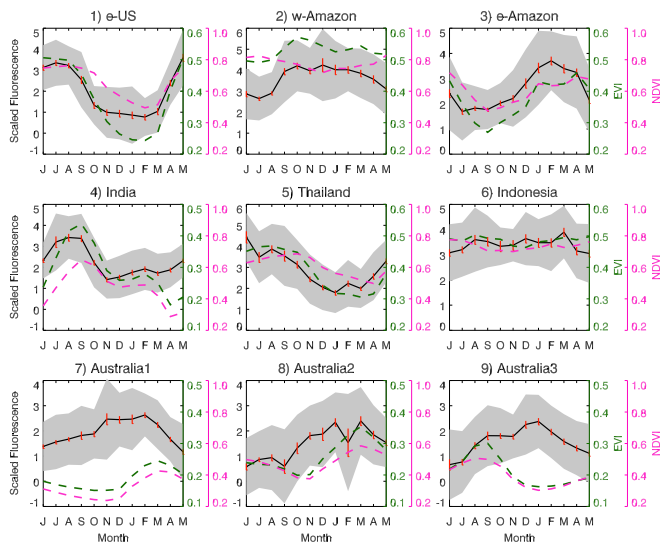
**Fig. 11.** Enhanced Vegetation Index (EVI) (unitless) from Aqua MODIS for July and December 2009. Boxes in top panel indicate regions of interest that will be further explored in Figs. 14–15.

NDVI included. New fitting.



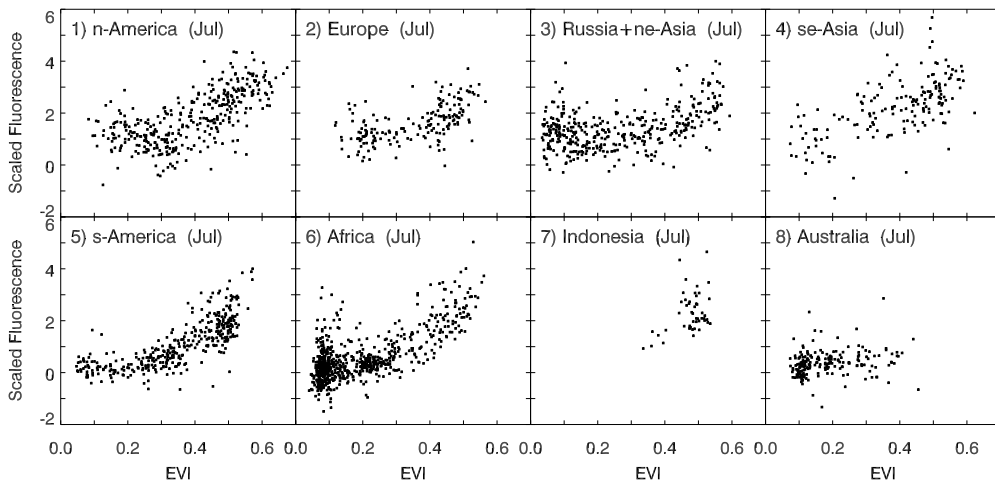
**Fig. 12.** Bottom panel: Full seasonal cycle in scaled  $F$  from the P polarization for June 2009–May 2010 (left axis, unitless) for the regions indicated in the top panel; Black line: Monthly means; Red line: Errors of the mean; Shaded gray: standard deviation; Dashed green(pink) line: MODIS EVI(NDVI) (right axes, unitless).

New figure showing results from S polarization.



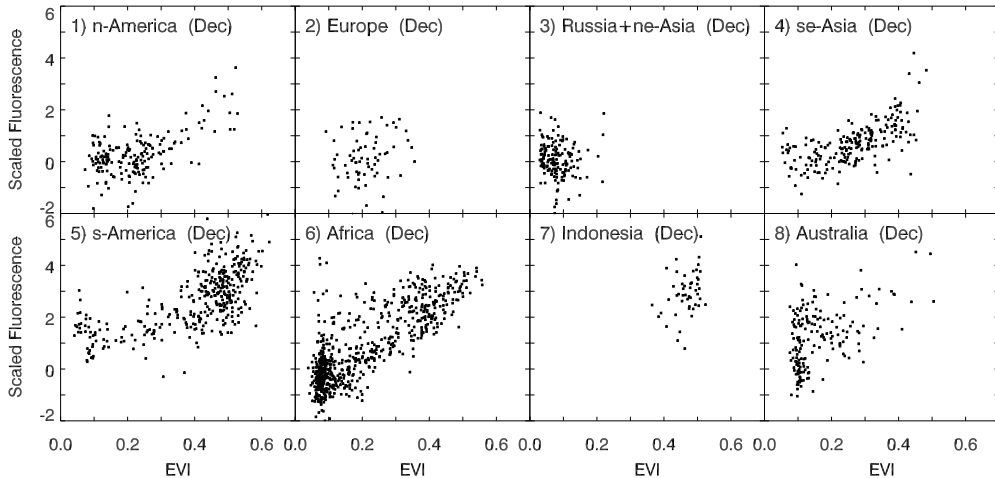
**Fig. 13.** Similar to bottom panel of Fig. 12 but derived from the GOSAT S polarization data.

With new fitting. Low F results now more centered about zero.



**Fig. 14.** Scatter plots of the GOSAT scaled  $F$  from the P polarization versus the MODIS enhanced vegetation index (EVI) (both unitless) for July 2009 for different regions shown in Fig. 11. Each point represents a monthly mean from a single  $2^\circ \times 2^\circ$  grid box.

With new fitting. Low  $F$  results now more centered about zero.



**Fig. 15.** Similar to Fig. 14 but showing scaled  $F$  from the P polarization versus the MODIS enhanced vegetation index (EVI) (both unitless) for December 2009.