Biogeosciences Discussions

Biogeosciences Discuss., author comment AC5
https://doi.org/10.5194/bg-2021-354-AC5, 2022

Reply on RC3

Mark Pickering et al.

Author comment on "Sun-induced fluorescence as a proxy for primary productivity across vegetation types and climates" by Mark Pickering et al., Biogeosciences Discuss., https://doi.org/10.5194/bg-2021-354-AC5, 2022

We thank the reviewer for the insightful comments, and welcome the tougher questions to both improve the manuscript and contribute to the general discussion within the SIF community. Please find below a response to all points raised one by one below.

- Do comparisons among products tell us anything?

  I am skeptical of analyses that compare products and interpret the results as containing empirical information or insights into their relationships. For instance, downscaled SIF is not SIF and FLUXCOM GPP is not GPP. To me it seems their relationships would be very sensitive to or determined by their respective errors. I do commend the authors for describing SIF as downscaled SIF and GPP as FLUXCOM GPP in the text and the figuresn, but does their relationship really tell us anything about SIF and GPP? Many other papers written with these products or similar products do not often make this distinction, and simply interpret downscaled SIF products as SIF and GPP products as GPP. So, thanks to the authors for being more diligent.
  On this note, I think it is a fair question to ask why a downscaled product was used - why not use the raw TROPOMI data. Or use them both and discuss how the results differ. If you repeated this analysis with gridded TROPOMI SIF data, would you get the same results? We have demonstrated ways to use ungridded (Doughty et al. 2019 PNAS) and gridded TROPOMI data (Doughty et al. 2021 JGR) for such analyses.

We agree that the downscaled SIF is not the same as the original SIF retrieval (nor the exact SIF emitted by the surface either), and that FLUXCOM GPP is but an estimation of GPP. Both have errors and shortcomings, as do all measurements and models in science. However, we still think that comparing them and examining their relationships can bring some insights about the underlying processes, or better, on where they are jointly well represented and where they don’t. We do not think these patterns would necessarily solely depend on the error structure. There is also quite some independence between both data streams, as SIF currently does not enter as a feature in FLUXCOM. Seeing where and when these match allows to identify zones of interests where more investigation is warranted. We will try to reinforce this message in the manuscript.

Regarding why we use a downscaled product instead of TROPOMI data, a first response is that since we wanted to investigate the interannual variability, using a downscaled product ensures more years are available. There is a bit of a historical reason for this too, as this work indeed started a couple of years ago when there were much fewer years of TROPOMI. Also, the idea was not necessarily to explore any SIF signal to GPP, but specifically the downscaled one. Including TROPOMI data would definitely be interesting. However, downscaled SIF is currently not available for the later years of TROPOMI because the input GOME2 retrievals are discontinued and they suffer from sensor degradation. While other datasets correcting for this might be available,this would still require a lot of cross-comparison and explanation. We fear all this would considerably lengthen an already long manuscript, and we thus consider it to be beyond the scope of the current study.

- *Downscaled SIF products*
  A couple of comments regarding the downscaled SIF products. This is certainly not intended to be a jab at the SIF-LUE product, but I think there are a couple of issues with most of these products that have not really been addressed yet.
  First is that have shown in my JGR 2021 paper that there is a very weak or often no correlation between VIs and SIF in the tropics, and I have found the same to be true when using TROPOMI SIF and TROPOMI surface reflectances. However, the downscaled products use VIs or surface reflectance, along with machine learning or environmental scalars such as we use in LUE models, to predict SIF.
  How sound is it to predict SIF with surface reflectance or VIs in the tropics when they lack a correlation? SIF is affected by physiological processes that do not affect leaf/canopy optical properties - so is it really safe to assume that we can use reflectances to predict SIF? This question is particularly important for the tropics since they are such a strong driver of annual and intra-annual GPP and XCO2.
  Second, do the downscaled products actually reproduce the SIF signal? The downscaled SIF products were produced before we had a sizable amount of TROPOMI data, but now we have four full years of TROPOMI data. Ideally, platforms with more coarse spatial and/or temporal resolutions (GOSAT, GOME-2, OCO2/3) would capture the seasonality of SIF in the tropics as observed by a near-daily observer like TROPOMI - but do we know that yet? And do their downscaled products reproduce the SIF signal, the VI signal, or something in between?

So here we might need to explain more clearly some particularities of the downscaled SIF approach we use. The downscaling is applied independently at every individual time step at which SIF is available using a relationship that is calibrated regionally. The downscaling is thus more akin to an unmixing process of the coarse signal, where the high spatial resolution explanatory data is trying to allocate where there is a higher likelihood to find higher values and where there might be lower values, and thereby try to best distribute these values around. However the actual signal that needs to be distributed remains that of the original SIF (gridded) observation. This is not the same thing as "predicting SIF with surface reflectance", which is what other machine-learning downscaling methods try to do.

Regarding the tropics, if we consider a homogenous land cover (e.g. tropical forests), and if we do not have any relationship between SIF and VIs over this area, this relationship will of course not help to disaggregate variation of SIF within that homogenous forest. Changes in LST or NDWI, which might have some relationship, could contribute partially despite a lack of relationship with the main VI (NIRv). But if no relationships are detected, no downscaling is done. The downscaled values retain the same value as the original SIF value at coarse spatial resolution. However, in reality, there will likely be some landscape fragmentation in which different land cover types can be detected at fine spatial resolution but not at coarse, and in these cases the downscaling (unmixing) approach should be able to disaggregate the coarse SIF signal based on differences detectable in NIRv (and LST and NDWI). This might not be sub-variations in SIF within the same landcover type, but it will be valuable to have a better relationships with different land cover types.

Finally, the downscaled (unmixed) product was benchmarked as suggested with respect to OCO2 observations in Duveiller et al 2020, and the agreement was good. Again, we do not pretend it "predicts SIF", but that it tries to "disaggregates the GOME2 signal to a finer spatial resolution".

- *Analysis by land cover type*
  Personally, I am not a fan of grouping land classes to investigate drivers of variables – in this case SIF and GPP. For instance, GPP in EBF in Africa or SE Asia can be driven by a different set of drivers than those in the Amazon. Even within the Amazon basin itself, there is a distinctive gradient in precipitation, temperature, VPD, etc. that is not static in space or time. Drivers of photosynthesis are determined locally by local environmental present and historical factors, disturbance history, species composition, human management, physiological processes, and many other local factors other than just land cover functional type.
  Thus, drivers should be investigated at the pixel level. Why not determine the drivers and their strengths and show it on a map? I am highly skeptical of any results that claim things like 'GPP for this land cover type is driven by x' or 'SIF is driven by x for this vegetation type'.
  Also, the majority threshold used is somewhat subjective and arbitrary. Even at 75% majority land cover type, a sizable portion of the signal (GPP, SIF, or spectra) is driven by a land cover type other than the one you are interested in. Thus, there is an inevitable bias in the results that can't be remedied. For instance, the seasonality in moist EBF of the Amazon is extremely subtle. Thus, even a small area of another vegetation cover type, such as crop or grassland, may dramatically alter the seasonality for a gridcell. Also, setting a 100% land cover threshold is unreasonable as one will end up with very few pixels for analysis, especially at 0.05-degree resolution.
  I have done these analyses myself while writing my 2021 paper published in JGR. I began the analysis by grouping by land cover type, but I obtained very different answers according to the majority % cover threshold that I used. I actually scrapped the entire paper and analysis in favor of showing the SIF-GPP and SIF-VI relationships at the gridcell level as maps, as it was not fair to extrapolate a relationship among all land cover classes globally as being characteristic for that land cover type when in reality the relationships and spatio-temporal relationships were much more complicated. And there are a lot of maps in that paper!

We agree, grouping by PFT has serious shortcoming regarding the simplification of the world and the reduction of local/regional particularities to an abstract grouping. In fact, this is why we are not only working at the PFT level but we also provide maps of continuous relationships between GPP and SIF (see figures 3, 6, 7). However, we also believe that grouping by PFT can still be useful, especially for some in the land surface modelling community who are still currently constrained by the PFT paradigm in order to calibrate their models.
We will take the remark on board and make a more critical point of this in the revised manuscript and discuss more about the caveats of using PFTs.

- *Would we expect the SIF-GPP relationship to be static?*
  The SIF-GPP relationships shown in Figure 5 - wouldn't we expect these relationships to vary over time and according to vegetation stress and other factors? Perhaps there is a seasonality to the relationships? What about a time series of their slopes, R2, or p values?

The relationships here are averaged over the full time period considered (2007-2014) and are growing season mean (i.e. more or less the annual mean). We take the mean GPP/SIF for each pixel over the time period. Then we do a pure spatial linear relationship. As such there is no evolution with time or time series. We do this to break down the SIF-GPP relationship into it’s spatial and temporal components separately. We also try to treat vegetation stress through the lens of fluctuations in meteorological variables in the paper.

It would be an interesting extension to consider the evolution in time of the 8-day spatial SIF-GPP relationship, breaking down the seasonality, however we consider it beyond the scope of the paper. As discussed in a previous comment we will add maps of the slopes of the spatio-temporal relationships in the update, but we are also wary of overloading the, already large, manuscript with figures, for example by including R2 (in addition to the correlation) and p-values.