

Dear Dr. Mukund Palat Rao,

We would like to thank you for taking your time to evaluate our work and foremostly for your interesting and useful comments, questions, remarks and suggestions.

We will try to answer to your questions and we integrated your structural and grammatical suggestions into the manuscript (all answers are in blue color).

Do the RFR style models (RF-F, RF-SIF-R, RF-SIF-R-PFT) outperform the RF-SIF-VI model just because they have more predictor variables (14 vs 4)? Perhaps you can present some analysis on how these models perform on 'validation' data that the models have not been calibrated on?

The RF-R model has 11 inputs, the RF-SIF-R 12 inputs, the RF-SIF-R-FTP 13 inputs and RF-SIF-VI 4 inputs. In the manuscript, we used the adjusted R^2 , which considers the number of samples and predictor variables in its computation to evaluate and compare the performance of our different models (see line 221, page 6). We also compared the models based on the RMSE, which isn't sensitive to the number of explanatory variables.

We also used out-of-bag predictions of RF to calculate the adjusted R^2 and RMSE. We separated our dataset in two datasets: 80% of the data for training and 20% for testing or evaluating the model. Second, on the training dataset we applied a 10-fold cross-validation and 20 iterations to determine the best parameters for each model. Lastly, we evaluated or tested each model on the testing dataset, which were not seen by the model before (see lines 207-213, page 6).

As these points were not clearly stated in the manuscript, the lines 219 to 222 were modified as follow :

“Ultimately, the linear relationships between SIF and GPP were compared based on the coefficient of determination (R^2), Root Mean Squared Error (RMSE), and the p-value metrics. The random forest models were evaluated and compared based on out-of-bag adjusted R^2 and RMSE”.

Page 1, Line 15: “Earliest” studies, based could be “earlier” studies or prior studies?

Line 15: changed to “prior studies”

Page 1, Line 17: “plant functional type” should be plant functional types.

Changed to “plant functional types”.

Page 2, Line 14: “which is the amount of flux carbon taken up by vegetation.” The word ‘amount’ is not needed since the work flux assumes an amount.

L 43 :Changed to “which is the carbon flux taken up by vegetation through photosynthesis”

Page 2, Line 53: “Remote sensing is widely used to upscale canopy GPP to landscape, regional, and global scales and at daily scale using reflected sunlight measured by satellite sensors”. The “and at daily scale” addition seems a bit awkward. Maybe remote sensing is widely used to upscale daily GPP to landscape....

Changed to: “Remote sensing is widely used to upscale daily GPP to landscape, regional, and global scales using reflected sunlight measured by satellite sensors.”

Page 2, Line 60: “and biochemical canopy characteristics (Dechant et al., 2020; PabonMoreno et al., 2022). Although, they suffer.” The way the sentence is framed, I think it would flow better as a continuous sentence.

Changed to: “and biochemical canopy characteristics (Dechant et al., 2020; Pabon-Moreno et al., 2022), but they suffer from contamination by atmosphere and saturation in canopy dense ecosystems and are less sensitive to diurnal and daily variations in photosynthetic status resulting from physiological responses induced by rapid changes of abiotic stresses.”

Page 3, Lime 83: “Early studies relied on ground-based”, should be Earlier studies relying on...

Changed to “Earlier studies relying on”

Page 3, Line 96: “which is on board Sentinel 5-Precursor, represents a novel (???) for understanding?”. A missing word after novel? Tool maybe?

Changed to: “which is on board Sentinel 5-Precursor, represents a novel tool for understanding”

Page 3, Line 97: “it provides a quiet high temporal resolution at daily”. Quite instead of quiet. However, the word quite is not needed either.

Changed to: “it provides a high temporal resolution at daily scale.”

Page 3, Line 104: “comprehensively addressed. Owing to most”. This should be once sentence, or the second sentence should start of as, this is due to the fact or This is because....

Changed to: “However, to the best of our knowledge, an attempt to study the synergy between those variables have not been comprehensively addressed due to the fact that the relationships between structural and functional components are not linear, and have complex interactions over time and space”

Page 4, Line 8: “data products is given in Supplementary Materials in Tab S2.”. the Authors don’t need to do this, but might consider including a column for the spectral band (i.e. visible (R/G/B), NIR, etc. in the table)..

Here the spectral band column was added in tab S2 as you suggested (see supplementary Material in Tab S2).

Acronym	Full Name	Wavelengths (nm)	Band name	Spatial Resolution
B ₁	Surface Reflectance for B ₁	620-670	Red	500 m
B ₂	Surface Reflectance for B ₂	841-876	NIR	
B ₃	Surface Reflectance for B ₃	459-479	Blue	
B ₄	Surface Reflectance for B ₄	545-565	Green	
B ₅	Surface Reflectance for B ₅	1230-1250	SWNIR	
B ₆	Surface Reflectance for B ₆	1628-1652	SWIR	
B ₇	Surface Reflectance for B ₇	2105-2155	SWIR	
B ₈	Surface Reflectance for B ₈	405-420	Blue	1 km
B ₉	Surface Reflectance for B ₉	438-448	Blue	
B ₁₀	Surface Reflectance for B ₁₀	483-493	Blue	
B ₁₁	Surface Reflectance for B ₁₁	526-536	Green	
B ₁₂	Surface Reflectance for B ₁₂	546-556	Green	
B ₁₃	Surface Reflectance for B ₁₃	662-672	Red	
B ₁₄	Surface Reflectance for B ₁₄	673-683	VNIR	
B ₁₅	Surface Reflectance for B ₁₅	743-753	VNIR	

B ₁₆	Surface Reflectance for B ₁₆	862-877	NIR
-----------------	--	---------	-----

Page 6, Line 230-232: “Overall, SIF_d was significantly related with tower-based GPP at the site level and at the daily timescale (as $p < 0.0001$ was statistically highly significant), except for IT-Cp2 site of which GPP and SIF_d relationship was insignificant and weak ($R^2 = 0.001$, $p \leq 0.60$)”. This is of course quite subjective, but despite some of the sites being statistically significant I would not call these relationships as being strong. The reason for the statistical significance and p-value is being driven by the high sample size. In particular, I would add GF-Guy to the list of sites where there is no relationship between GPP-SIF_d. The correlation needed to get an R^2 of 0.2 is around 0.15 which is still quite weak. I would also then add FR-Mej, FR-EM2, and FI-Var to the list of sites with a weak relationship. I know this is mentioned a bit later, but maybe an easier way to frame it would be to not mention the weak relationship at IT-CP2 in the beginning, but then mention all these sites together at the end of the paragraph?

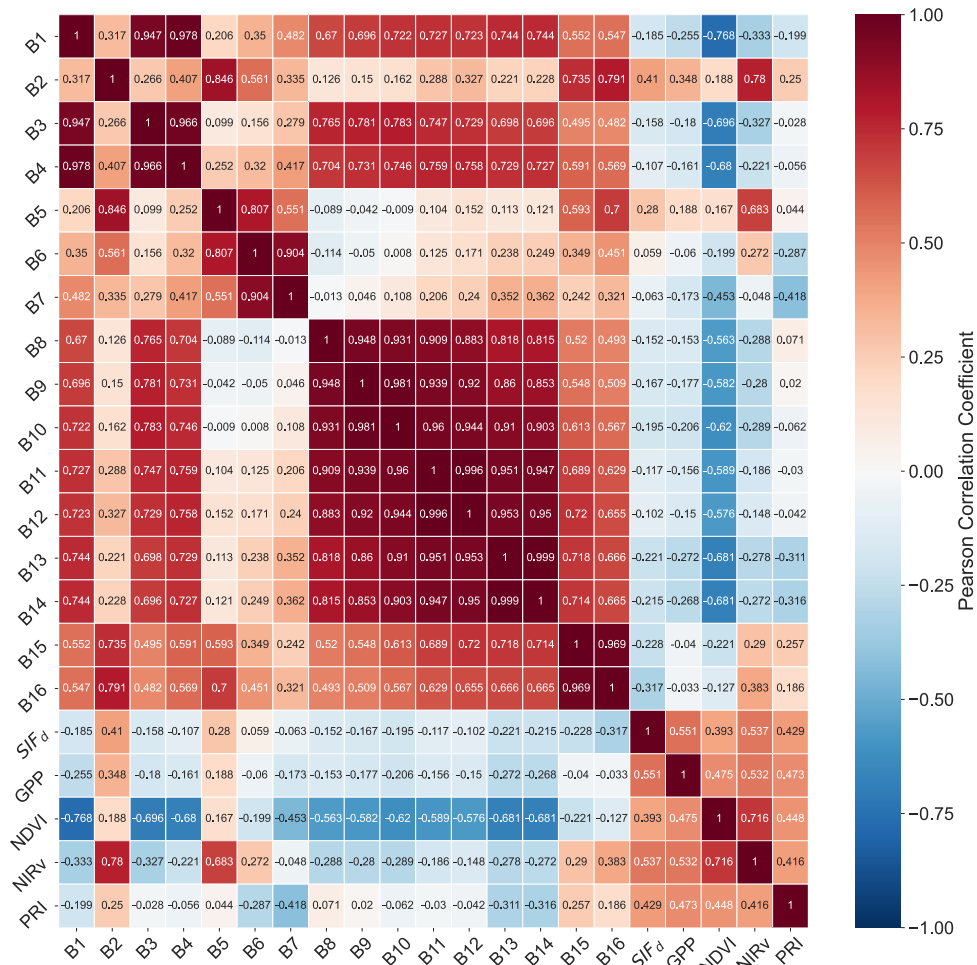
This part was reframed as follows (lines 230-236): “Overall, SIF_d was significantly related with tower-based GPP at the site level and at the daily timescale ($p < 0.0001$). However, Figure 2 indicates that the slopes and intercepts of the linear regression between GPP and SIF_d are site-dependent, suggesting that the difference in plant functional types and spatial heterogeneity across sites may significantly affect the relationships between GPP and SIF_d. The strongest relationships were found at DK-Sor and FR-Fon, which are DBF vegetation type sites, with R^2 values of 0.81 ($p < 0.0001$) and 0.66 ($p < 0.0001$). The weakest linear relationships were recorded at FI-Var, FR-EM2 and FR-Mej sites, and no significant relationship was found at GF-Guy and IT-Cp2.”.

Page 7, section 4.1.2 I like the progress from the site level (Section 4.1.1) to PFT level SIF_d-GPP relationships. However, the way the PFT level relationships are presented, don’t seem to actually allow us to closely examine within PFT spread in the R^2 ’s, slopes, and intercept. For example, in Table 2, all sites of a PFT are lumped together. If the authors wish to highlight the within PFT spread, one option could be to include boxplots by PFT for the R^2 , slope, and intercept for the SIF_d-GPP relationships.

We did not include the boxplot of R^2 , slopes and intercepts of the GPP vs SIF_d for sites from the same PFT, because, we have low number of sites for some PFT, including OSH (1 site), MF (2 sites), EBF (2 sites), and GRA (3 sites). In addition, the R^2 , slopes and intercepts of the relationships between GPP and SIF_d for each PFT and site are detailed in Tab S3.

Page 9, Section 4.5, Line 284, Supplementary Fig. S1: I would recommend changing the figure to have a diverging colour-bar. The gradient colour-bar from ~ -0.65 to 1 is not intuitive to me and hard to visualize.

In the figure below, the gradient colour-bar for the correlation matrix was changed, as you recommended. The figure was added to the Supplementary Material draft.



Page 12, Table 3: would benefit from a vertical line separating RF-R, RF-SIF-R, and N, and other similar vertical line in the lower panel

The vertical lines separating RF-R, RF-SIF-R and N were added in table 3.

PFT	Sites	N	RF-R				RF-SIF-R			
			Adj. R ²	Slope	Intercept	RMSE	Adj. R ²	Slope	Intercept	RMSE
CRO	9	1171	0.78	1.03±0.03	0.00±0.24	2.67	0.75	1.01±0.03	0.08±0.26	2.89
DBF	6	748	0.92	1.02±0.02	-0.23±0.18	1.41	0.90	1.05±0.02	-0.52±0.21	1.61
EBF	2	188	0.77	0.93±0.07	1.01±0.83	1.23	0.68	0.90±0.09	1.58±0.99	1.45
ENF	13	1385	0.85	1.01±0.02	-0.01±15	1.29	0.78	1.06±0.03	-0.23±0.19	1.54
GRA	3	364	0.81	1.02±0.05	-0.02±32	1.64	0.76	1.07±0.06	-0.17±0.38	1.87
MF	2	117	0.84	1.05±0.08	-0.15±0.76	1.49	0.82	1.12±0.10	-0.62±0.83	1.56
OSH	1	317	0.91	1.02±0.04	-0.09±0.22	0.99	0.88	1.01±0.04	0.01±0.24	1.10
WET	4	599	0.92	0.98±0.02	-0.15±0.10	0.85	0.84	0.98±0.03	-0.37±0.15	1.17
ALL	40	4889	0.86	1.02±0.01	-0.09±0.08	1.72	0.82	1.04±0.01	-0.19±0.10	1.94
			RF-SIF-VI				RF-SIF-R-PFT			

PFT	Sites	N	Adj. R ²	Slope	Intercept	RMSE	Adj. R ²	Slope	Intercept	RMSE
CRO	9	1171	0.70	1.03±0.04	0.01±0.29	3.14	0.75	1.00±0.03	0.12±0.26	2.87
DBF	6	748	0.84	1.05±0.03	-0.58±0.28	2.06	0.91	1.04±0.02	-0.40±0.21	1.56
EBF	2	188	0.51	0.77±0.11	3.42±1.14	1.80	0.72	0.96±0.09	0.74±0.98	1.37
ENF	13	1385	0.66	1.02±0.04	0.10±0.24	1.92	0.79	1.08±0.03	-0.39±0.19	1.5
GRA	3	364	0.69	0.98±0.07	0.02±0.43	2.11	0.77	1.07±0.06	-0.29±0.38	1.84
MF	2	117	0.71	1.04±0.12	0.04±1.07	2.00	0.82	1.12±0.09	-0.73±0.84	1.56
OSH	1	317	0.83	0.98±0.05	0.21±0.29	1.33	0.89	1.02±0.04	-0.06±0.24	1.08
WET	4	599	0.72	0.88±0.04	-0.39±0.21	1.54	0.88	1.05±0.03	-0.29±0.12	0.99
ALL	40	4889	0.75	1.03±0.02	-0.18±0.12	2.28	0.83	1.03±0.01	-0.15±0.09	1.89

Page 15, Line 431: “it can be avoided”, maybe better phrased as “we don’t need to rely on land cover type....and meteorological data”?

Changed to: “we do not need to rely on land cover type and land cover change, and meteorological data.”

Page. 16, Line 465: ERA5 instead of ERAS?

Changed to: “from ERA5 reanalysis.”