

After more than three decades of global EC measurements, there is still a large data gap in the tropics. As tropical forests contribute largely to carbon and water fluxes, this data gap is a large source of uncertainty in estimating global carbon and water dynamics. This paper adds to closing this gap with EC data from a wet tropical forest site in Costa Rica, even though the data coverage is only $\pm$ one year. I think the dataset is very valuable, however the data processing is not very thoroughly explained in the paper, e.g. info about u^* filtering, data left after QC, shape and value of storage correction, In addition it seems like several references are not on the right place and/or are not the main reference for the method or results mentioned in the text. This might give the indication that the methods and claims are not well supported. The partitioning methods explained in the text are missing details, which leaves the reader in confusion and limits the understanding of the results and the re-usability of the methods. Some large conclusions about partitioning methods are drawn and in my opinion these conclusions are not yet well supported by the analyses carried out in the study. While the set-up and ideas behind the study are very interesting, such as comparing different partitioning techniques, looking into drivers on different timescales, ... the study itself still raises more questions than answers and needs some more work before publishing.

Overall I would advise to check all references, to more accurately describe the analyses on the NEE as this is the strong basis for all following analyses, to more thoroughly describe the partitioning methods used and more clearly indicate (dis)similarities with more common methods and to strengthen the comparison between partitioning methods.

Comments:

1 Introduction

Line 57 to 69: rewrite this section, this reads very unstructured and contains dubious information.

- ➔ Introduce NEE as the sum of GPP and RECO. The sentence “partition net ecosystem exchange (NEE) into its component fluxes”, is not entirely clear.
- ➔ “the nighttime-based method to extrapolate respiration's temperature response to daytime periods” rephrase this to state that it is the nighttime temperature-respiration relation that is extrapolated to the daytime periods. Now it is not clear which respiration temperature response is meant.
- ➔ “The nighttime method often corresponds to the period of day when ...”.
“Method corresponds to a period” is not clear. Better write that the data used to fit the temperature-respiration relation is nighttime data, and then mention the associated uncertainties with nighttime data. Uncertainties in nighttime data are indeed accumulation, but also heavy data filtering is a source of uncertainty and the

assumptions underlying the method. I would argue that if there is respiration underestimation, then this is an issue of your NEE measurements, and not inherently a downside of the partitioning technique. If this is accurately corrected for, using filtering and storage corrections, this underestimation would not influence the partitioning method too much., so I would definitely include additional other uncertainties next to the underestimation, such as heavily filtered nighttime data, light inhibition of leaf respiration, the uncertainty due to the corrections, ..

- ➔ “Meanwhile, limitations associated with the daytime partitioning method include potential overestimation of morning CO₂ fluxes from buildup throughout the night, as well as strong sensitivity of GPP estimates to variability in incoming radiation”.

The overestimation of morning CO₂ from build up, should already been corrected in the NEE, so should not influence the partitioning method, similarly as the previous comment. The uncertainty or variability due to corrections can play a part of course. I am not sure how the “strong sensitivity of GPP estimates to variability in incoming radiation” introduces a bias. Do you mean that the model does not capture fast changing environmental conditions?

- ➔ “Moreover, the daytime method is subject to biases introduced by seasonal changes in phenology, leaf area index, and water stress (Keenan et al., 2019; Lasslop et al., 2010), which alter both the shape and magnitude of ecosystem light-response curves over the growing season.”

The shape of the light response curve is allowed to change, as parameters are fitted on small windows, so I don't see how seasonal changes can introduce strong biases. Maybe here add the light inhibition of leaf respiration, this is an issue with both the night and daytime partitioning method? Also I don't see how the two references here explain the possible biases.

Line 71-72: More recently, the sundown partitioning approach has been proposed as an alternative to better constrain GPP and Reco estimates (Wehr et al., 2016).

Wehr et al 2016 used simultaneously carbon and isotope EC measurements to partition NEE. This paper is not at all related to the sundown method. Are you referring to van Gorsel et al. 2009?

2 Methods

Line 137: closed path needs to be enclosed path, because the LI7200 is an enclosed analyzer.

Line 140 Webb-Pearman-Leuning (WPL) corrections for density fluctuations (Webb et al., 1980) are not needed for closed and enclosed analyzers, however this will not have influenced your data as eddypro will automatically deactivate the setting.

Line 144: The enclosed-path design of the Li-7200RS minimizes rain contamination of the optical path, and remaining rain-affected periods were removed through EddyPro's standard quality flagging

➔ There are not issues with rain contamination of the optical path in an enclosed path.

Maybe add a diel cycle of non-storage corrected, storage corrected NEE and the storage correction term itself in supplementary to indicate the effect of storage.

Line 165: closed path sonic anemometer

Using the enclosed path **analyzer**, sonic anemometer and tower profile system...

Line 168– 171 The u^* threshold reflects the potential bias in nighttime NEE due to low turbulence conditions.

- ➔ Mention the u^* threshold that you used to filter out low turbulent conditions, mention also how many data points are removed and how many data points were left after the u^* filtering.
- ➔ Uncertainty due to U^* is reflected in the difference between the 5th and 95th percentile of u^* thresholds, mention these thresholds maybe in supplementary.
- ➔ In supplementary Fig A2 there is stated: “Points below the threshold would be excluded due to insufficient turbulence, while points above the threshold would be retained. Storage and turbulent flux components are shown.” With the “would be” construction it seems like the data points below the threshold were not removed?

Line 188: If gaps are left unfilled, how did you calculate the budgets mentioned in Line 420 later on? Are these gaps equally distributed between day and night measurements? If more nighttime gaps are present, this could result in a underestimation of the ecosystem respiration

Line 193 -194: how was this gap-filling done? Just putting the values of the other meteorological station into the gaps? Did the instruments compared well?

Which software did you use to do the partitioning, the gap-filing, ..?

Line 233: formula (1)

- ➔ The method described in Lasslop et al. 2010 used a rectangular hyperbolic light-response curve, which has a different formula that is used here. Mention that this is

different. Of course the non-rectangular hyperbolic fit can also be used, however then the formula is missing a term, because now the square root of $(\alpha PAR + \beta)^2$ just returns the absolute value of $\alpha PAR + \beta$. Check the formula again.

- ➔ “The y-intercept γ is the base respiration level in the absence of light, which is then used to calculate Reco following the assumptions described in Sect. 2.4”

So respiration is constant during day and night and only changes when a new curve is fitted, i.e, once every week? This is not similar to Lasslop et al. 2010 as he used a temperature dependence for respiration. This method seems closer to the method by Gilmanov et al. 2003

Line 242-243 “Using the base respiration rate from the non-rectangular hyperbolic fit, GPP can be derived using the equation $GPP = NEE - Reco$ To reliably fit this curve, a sufficient density of data points is required. This limited the temporal resolution to a minimum of one week. As a result, two timescales were used: daily and weekly. On the weekly timescale, light-response curves were fitted to obtain stable model coefficients, including the base respiration rate. These weekly respiration estimates were then applied to the corresponding daily NEE observations within each week to calculate daily GPP values.”

- ➔ “using the equation $GPP = NEE - Reco$ ”. I think this statement makes things difficult because now there can be confusion: do you mean NEE(PAR) with the NEE in the formula here, or the original measured NEE values? Likely the first
- ➔ “These weekly respiration estimates were then applied to the corresponding daily NEE observations within each week to calculate daily GPP values.”
So this means that you do not calculate half-hourly GPP values? I am confused here with the daily temporal resolution. You fit the curves on half-hourly data?
- ➔ Is this curve fit on day and nighttime data?

Section 2.4.2 In the Morning and Afternoon Daytime Partitioning Methods: no nighttime data is used to fit the curves?

Section 2.4.3: The method by Reichstein et al. 2005 uses an exponential regression model, not a linear regression. You explain why, but just maybe rephrase.

Line 269: “and GPP_N is calculated from $R_{eco, N}$ ”.

Mention that here you subtract the estimated Reco from the measured NEE to result in GPP.

Line 272: “despite evidence that foliar respiration is often suppressed in high light (Cavaleri et al., 2008).”

I agree that light indeed suppresses foliar respiration, however the reference used here mainly talks about dark respiration and there is no mention of light inhibition of leaf respiration.

Section 2.4.4

You only mention Lasslop et al 2010 here, but this paper references the daytime partitioning method, not the sundown partitioning method. Do the methods somehow rely on some result from Eva van Gorsel 2009? You relate the R_{eco} to temperature, is this with a linear relation?

Line 310: “To assess the relative importance of environmental drivers on GPP, a random forest model (RFM) was trained on observed light use efficiency (LUE) data calculated from GPPA divided by PAR analyzed on half hourly, daily, and weekly timescales ($LUE = GPPA / PAR$) (Breiman, 2001; Jung et al., 2011).”

- ➔ Data users are frequently confused on what EC measures and what we estimate of model, so try to be consistent in wording: LUE is not observed when it is calculated
- ➔ It's a complex sentence
- ➔ I am confused with the reference to Jung et al.. This is a paper about upscaling FLUXNET observations of carbon dioxide, water, and energy fluxes to the global scale. Not related to LUE.

3 Results

Line 447: “Several factors may contribute to this elevated sink strength including methodological choices such as tower configuration and storage flux treatment”

- ➔ How so, references? Most of the time storage flux reduces the sink strength, I would guess

Line 449: Although a variety of GPP partitioning approaches have been used, the majority of previous studies relied solely on the nighttime partitioning method.

- ➔ You are talking about NEE here, partitioning methods can not influence the NEE measurements

Check the values again in Table 1: for example the values of French Guiana are very misleading, as these are values mentioned for different seasons and therefore can not be used as values for sink or source per year. Only use accumulated values of a whole year to make comparisons.

Line 464: “The base respiration parameter γ spans a physiologically unrealistic range, varying from 7.5 to 15.9 $\mu\text{mol m}^{-2} \text{s}^{-1}$ (Cavaleri et al., 2008; Loescher et al., 2003).”

➔ Why is this a physiologically unrealistic range? This is explained further down, so maybe just restructure the two paragraphs

Line 472 – 496: I think this is an interesting analyses, however the two paragraphs can be better structured (see comment above)

Line 474: “This variability is particularly pronounced during morning periods at low PAR ($<500 \mu\text{mol m}^{-2} \text{s}^{-1}$), where high NEE values are observed (Fig. 2d).

➔ could this also be evening hours?

Line 489 – 495: In Fig. 2d, the initial slope is $0.0407 \mu\text{mol C } \mu\text{mol PAR}^{-1}$, whereas in Fig. 2e it is more than double at $0.0927 \mu\text{mol C } \mu\text{mol PAR}^{-1}$. Correspondingly, the estimated base respiration term in Fig. 2d is $8.22 \mu\text{mol m}^{-2} \text{s}^{-1}$, compared to $14.8 \mu\text{mol m}^{-2} \text{s}^{-1}$ in Fig. 2e, which typically averages around $10 \mu\text{mol m}^{-2} \text{s}^{-1}$ based on previous studies from the La Selva, Costa Rica rainforest (Cavaleri et al., 2008; Loescher et al., 2003). In tropical forests, documented responses of GPP and Reco to short-term variations in humidity or VPD are typically on the order of 10–30%, not the $>100\%$ change observed here (Restrepo-Coupe et al., 2013; Saleska et al., 2003).

➔ The two periods compared are a month away, so I don’t completely see how the statement ‘short-term variations in humidity or VPD ...’ is linked to the situation here, because a month is not short term and we have no clear indication of what the difference in meteorological conditions are during these two periods.

Line 521 – 523: The coefficient of determination for the NEE–PAR relationship in the afternoon ($R^2 = 0.62$) is almost double that of the morning ($R^2 = 0.34$), indicating that NEE variability is more strongly explained by PAR in the afternoon than the morning.

➔ Slightly rephrase: indicating that NEE variability in the afternoon is more strongly explained by PAR than the variability in the morning.

Line 527 – 534: Many gap-filling methods rely on fitting light-response curves to daytime NEE data under the assumption that PAR is the dominant driver of variability (Reichstein et al., 2005; Lasslop et al., 2010). However, the elevated scatter and reduced R^2 of PAR during morning periods indicates that including these data can reduce the model’s overall performance. Restricting fits to the afternoon periods when atmospheric mixing is stronger and NEE variability is more tightly coupled to PAR can improve the stability of fitted parameters and reduce bias in gap-filled fluxes.

- ➔ I would leave Reichstein out of this, because the method does not use PAR.
- ➔ I am not convinced that we can make this conclusion yet. Did you test the models' overall performance by comparing the estimate NEE of all three models with the measured NEEs. Because the models performance on all NEE values does not directly depend on the R^2 relation between the subset of NEE and PAR values. Also it seems like these R^2 values are only based on the month of July, looking at the title of the graph, so is this month representative for all R^2 values? You should actively test the overall performance of the different models by calculating R^2 or root mean squared errors between the predictions and all the NEE measurements. Because you can have a very tight fit in the afternoon, but if this fit can not predict morning values, then this fit can not be used to predict overall NEE values.
So I am not saying that the conclusion is wrong, I think you just need more tests to be able to claim this.

562-565 In the context of daytime partitioning, this can bias the fitted intercept upward, potentially overestimating base respiration if the light-response curve interprets these early-morning spikes as baseline flux rather than short-term spikes. Together, these effects suggest that enhanced morning fluxes arise from a combination of physical storage dynamics and ecosystem responses.

- ➔ Will the method not also use all nighttime data to fit this base respiration? will these morning data points have a stronger influence than all evening and night datapoints?

The original range was 7.5 to 15.9 $\mu\text{mol m}^{-2} \text{s}^{-1}$, which was said to be physiologically unrealistic. The new one, from the afternoon methods, is 5.0 $\mu\text{mol m}^{-2} \text{s}^{-1}$ to 11.6 $\mu\text{mol m}^{-2} \text{s}^{-1}$. In absolute terms the range is smaller, but the values themselves are also smaller, so in the original, from the smallest to the largest is a multiplication with 2.1, while in the new values there is a multiplication with 2.3. So I just wonder why the first range is unrealistic, but the second one is not. More the absolute values then?

Line 597-599: While this approach has been noted to face challenges in complex tropical forests with heterogeneous canopy structure (Hutyra et al., 2008; van Gorsel, 2009), our results suggest it performs well at La Selva compared to other partitioning methods.

- ➔ You base this conclusion solely on weekly GPP and RECO, and on comparison with other methods that are themselves not fail-proof. Did you check day-to-day variation, diel cycles, light-inhibition, compared Reco values to soil chambers, ...? How much did the a and b parameter of the nighttime method fluctuate?

Line 601: Under these conditions, u^* is frequently low (Fig. A2), indicating weak turbulent mixing that can lead to CO₂ accumulation near the forest floor, within canopy gaps, and in low-lying areas (Araújo et al., 2008; Goulden et al., 2006). However, the inclusion of profile storage fluxes in our NEE calculations likely mitigates some of the potential underestimation that would otherwise occur when respired CO₂ accumulates below sensor height (Montagnani et al., 2018). This may explain why the nighttime method shows strong agreement with the daytime afternoon and sundown approaches (Fig. 4d), despite the theoretical concerns about low nighttime turbulence.

- ➔ If you performed the u^* filtering, this also likely took out all low turbulent moments, hence minimizing underestimation
- ➔ How many nighttime data points were still present after all quality checks? Because for most tropical sites, the reason why the method performs badly, is because there is too little data to fit the curves with high confidence.

Line 620 – 621: The slightly lower values for GPPN and RecoN were expected, as this method tends to underestimate fluxes (Keenan et al., 2019).

Keenan et al 2019 states “The nighttime method will thus overestimate daytime total ecosystem respiration, and consequently apparent photosynthesis”

However the nighttime tends to underestimate the fluxes IF the nighttime respiration is underestimated, so use other reference, eg Lasslop

Line 621-622 Weekly differences among partitioning methods for the daytime afternoon, nighttime, and sundown partitioning methods were assessed using a one-way ANOVA

- ➔ Mention all statistics in methods section

Line 622-626: For GPP, no statistically significant differences were detected in 79.5% of weeks ($p > 0.05$), while Reco showed no significant differences in 81.8% of weeks, indicating strong agreement among methods at weekly timescales. During our 11-month observational period, none of the three partitioning methods differed significantly for either GPP or Reco ($p > 0.05$).

- ➔ ANOVA with time series might be dangerous as you violate the independence assumption, so I don't think you can easily compare the different methods with anova.

Line 755 – 756 At the daily timescale, predicted LUE declined with increasing temperature, with the strongest temperature sensitivity occurring at $PAR > 700 \mu\text{mol m}^{-2} \text{s}^{-1}$

And Line 763 – 766 The interaction between temperature and light availability on LUE suggests that warming impacts on tropical forest carbon uptake may be most pronounced during lower-light periods, though further work isolating direct thermal effects from covarying VPD would help clarify this mechanism.

- ➔ These two sentences seem contradictory. When the largest difference (sensitivity) is at high PAR, then the warming impacts will be most pronounced at high PAR, no?

Line 758 – 759: This temperature threshold is lower than observed photosynthetic temperature optima of 28 °C for lowland tropical tree species (Slot & Winter, 2017) and temperature thresholds for GPP in tropical forests (Pau et al., 2018).

- ➔ Your temperature optima is an averaged daily temperature, while I think this is more frequently determined on a more instantaneous temperature effect. Likely if you would use maximum daytime temperature or something like that, you will get a higher value of temperature threshold. Because likely the photosynthesis is reduced only for a part of the day with high temperatures, above 28/30 degrees, but then the daily average temperature is likely only 25. So it might be more useful to calculate a temperature threshold on half hourly data.

This narrower daily range likely explains why VPD shows limited apparent influence at the daily scale since the model rarely encounters daily mean VPD values approaching the ~1.5 kPa threshold typically associated with significant stomatal downregulation in tropical wet forests (Grossiord et al., 2020).

- ➔ I can not seem to find the threshold of ~1.5 kPa in the mentioned reference

Line 790: Our results demonstrate that instantaneous environmental limitations (particularly VPD and temperature) regulate LUE at the half-hourly timescale,

- ➔ They regulate on the daily timescale, consistent with Fig. 5

Line 795 – 798 This underscores that temporal aggregation itself can obscure sub-daily controls, and scale dependent feature importance should be interpreted alongside the range of each predictor presented to the model at that scale. These findings highlight both the value and limitations of machine learning approaches for understanding ecosystem functioning.

- ➔ “This underscores that temporal aggregation itself can obscure sub-daily controls” I think this sentence is obsolete as temporal aggregation in essence loses sub-daily information. I think the conclusion here is just that looking into drivers on different time-scales is important. I don’t think this is a limitation of machine learning approaches, as this is exactly what was done in the study

4 Conclusions

Line 800: Our results demonstrate that carefully selecting an EC partitioning method is essential for generating realistic and defensible estimates of tropical forest carbon exchange.

- ➔ I completely agree here, but in this study all methods were based on similar assumptions, namely: respiration is same during day and night, and only T dependent or event constant and GPP is fitted using light response. So you could argue that this is a rather small selection of methods from which you chose one, and that choice for the afternoon methods was not very strongly supported.

Line 817 – 819 The growing influence of temperature at longer timescales indicates the cumulative effects of sustained thermal and atmospheric conditions on canopy photosynthetic efficiency.

- ➔ I think this is too strong. The importance of temperature only grows from half hourly to daily, which is still far away from sustained thermal conditions or cumulative effects.

Line 838 – 841: First, the pronounced difference between morning and afternoon light response parameters and substantially higher R² and lower parameter uncertainty of afternoon fits suggests that the daytime partitioning approach may systematically overestimate base respiration at wet tropical sites where morning CO₂ accumulation produces high early morning fluxes.

- ➔ I don't think you can conclude this. It seems like there is no consistent overestimation of the base respiration as in Fig. 3 c) the orange curve is frequently located above or right on the black curve.
- ➔ In addition, you never compared the effect of the differences in parameters, on the actual estimated NEE
- ➔ The fit might be less uncertain, but it is fitted on a smaller amount of data, so the result on the larger data set, might therefore not be better (see previous comment on that specific part)

Small additional corrections:

Line 33: VPD is not yet mentioned, introduce full word.

Line 76 "Improving GPP estimates is essential for quantifying, understanding, and predicting the fate of tropical forests as net carbon sinks or sources.

Lost sentence (has no contribution here)

Line 85: is PAR effect on GPP linear?

Line 96: Like most studies in wet tropical forests, this one relied on the nighttime method (Loescher et al., 2003), but comprehensive evaluations comparing multiple partitioning approaches remain limited.

It is the partitioning methods that relied on the nighttime method

I don't think they use the nighttime method in the paper, because they relate daytime NEE to PPFD

Might include instrumentation list with brand, and measurements height in the supplementary.

Line 311-312 LUE values were computed directly from observed half-hourly GPP and PAR measurements prior to model training.

→ GPP is not observed, is modeled.

Line 359: Environmental conditions over the study period provide context for interpreting both GPP partitioning and the environmental controls on LUE, ~~particularly given the relatively narrow temperature range for this field site.~~

Line 444-446: Compared to other tropical ecosystems, our NEE values fall within the range of NEE estimates found for other wet tropical forest sites (ranging between 95 and -2,534 g C m⁻² yr⁻¹; Table 1), with our estimated NEE falling toward the high end of the range.

Be consistent with minus signs for sink. Also the phrase “our estimated NEE falling toward the high end of the range” is confusing, because the estimate here is very low (negative).

Line 456: which dataset?

Line 646 – 647 remove, this is already mentioned

Line 652 – 654: The variation of GPP as a function of soil moisture and VPD motivates further examination of which environmental drivers most strongly regulate photosynthetic efficiency across timescales, addressed in Sect. 3.4.

→ This variation is not mentioned yet.

Line 656 – 660 remove, already in methods section

Line 681 – 683: Feature performance metrics at each temporal resolution matched the cross-validated R² and RMSE values reported above (Fig. 5a–c), since permutation

importance was computed using the same cross-validation folds as the model performance evaluation.

➔ Somehow strange formulation, especially feature “performance”

Line 699: or light saturation?

Line 700 – 703: This aligns with well-established sensitivities of tropical photosynthesis and ecosystem respiration to sustained warm temperatures, which influence stomatal behavior and multi-day vapor pressure gradients (Lloyd & Farquhar, 2008; Slot & Winter, 2017).

➔ Complex sentence

Line 720 – 730 move to methods

Line 799: conclusions should have different number