Discussion of “Compound soil and atmospheric drought events and CO_2 fluxes of a mixed deciduous forest: Occurrence, impact, and temporal contribution of main drivers”

Author response to Referee 1 comments

Scapucci et al.

May 9, 2024

In the following, *reviewer comments are given in italics*, author comments are given in normal font.

1. General comments

In this manuscript, Scapucci et al. investigate the impact of compound soil and atmospheric droughts on both ecosystem and forest floor carbon fluxes in a montane-mixed deciduous forest. Overall, the manuscript is well-written, and clearly, a lot of fieldwork and data analysis has been done, which is commendable. It has the potential to enrich the literature, and particularly, the idea of examining the responses of both above-canopy and below-canopy carbon fluxes separately to compound droughts is novel. I only have a few concerns that need to be addressed.

Thank you for your positive comments. We have addressed your concerns below.

1) Since SHAP values could also give the overall feature importance, why not just use SHAP for the first driver analysis for daily mean NEP instead of using the conditional variable importance as presented in the manuscript? Are the results based on these two methods consistent?

Thank you for the suggestion. It is true that the SHAP values can also be used to show the overall feature importance by taking the mean of absolute values of all the individual marginal contributions of each feature to model predictions. However, the conditional variable importance (CVI) used in this study is a modified approach for estimating feature importance, specifically designed when there is a collinearity between different features (e.g., Rg and Tair, VPD and SWC), thus it is called ‘conditional’ (Strobl et al., 2008). Therefore, for our study, we chose CVI as a more reliable metric for feature importance. We will clarify our reasoning of using CVI in the section “2.5 Data analysis”.

Nevertheless, based on your suggestions, we also estimated the mean of absolute feature importance for daily mean NEP using SHAP values (Figure R1.1). If we compare this analysis with that presented in Figure R1.2 (Figure 4 of the original manuscript), the CVI and SHAP values from both methods were similar for 2015 and the long-term mean 2005-2022 (Rg was the most important variable followed by VPD, Tair, SWC in both SHAP and CVI analyses), but there were also some discrepancies (i.e., VPD and SWC in 2018 and 2022). These differences were expected as SHAP and CVI are based on two very different mathematical formulations and because CVI does not include positive vs. negative effects of the features (included in the SHAP model) but only absolute effects. As mentioned above, we will highlight the differences of the two methods in the revised manuscript and add the figure R1.1 to the appendix.

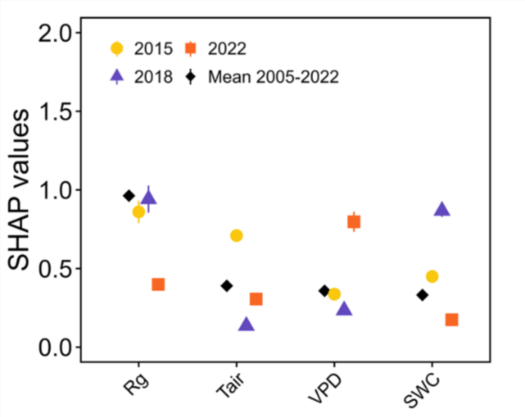

Figure R1.1. Absolute mean SHAP values (±SE) of daily mean NEP obtained with the XGBoost model.

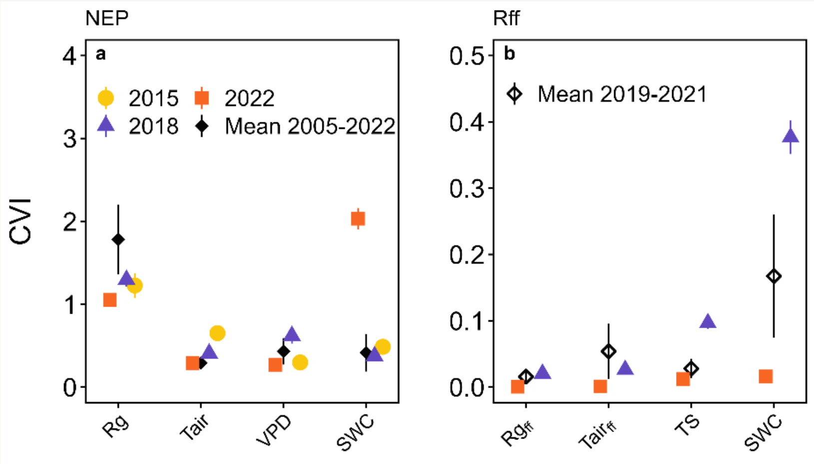

Figure R1.2. Conditional variable importance obtained with the random forest model as presented in Figure 4 of the original manuscript for daily mean NEP (±SE) and Rff (±SE).

Air temperature is not important for daily mean NEP during CASD based on conditional variable importance in Figure 4, while air temperature is still important for daytime mean NEP during CSAD using SHAP in Figure 5. Although daily mean NEP and daytime mean NEP are different, the results using the two methods seem to be inconsistent. Therefore, please also report the overall feature importance of predicting daily mean NEP based on SHAP.

Thanks for highlighting this point. The differences in feature importance/marginal contribution results (driver analysis) in Figures 4 and 5 (in the original manuscript) are due to different data sets and different methods used for two different purposes as described in the section "2.5 Data analysis" (Lines 165-188 of the original manuscript). The exclusion of nighttime NEP values from the SHAP analysis aims to highlight the effects of environmental variables when photosynthetic processes are dominating (during the day), effects that might be hidden if we would include nighttime data. Thus, we expect to see discrepancies between Figures 4 and 5. As you recommended, we calculated the overall feature importance for daily mean NEP with the SHAP analysis and found some similarities but also some differences (see previous comment and Figure R1.1). This discussion will be added to the revised manuscript.

2) The authors used the response curves of SHAP values vs. the abiotic factors to derive the driver thresholds. They observed an increase in Tair_NEPmax and a decrease in SWC_NEPmax from 2015 to 2022, thereby concluding drought acclimation of NEP to higher VPD and lower SWC. However, the thresholds they found were responding to the maximum positive marginal effects of drivers, which means the drivers at these thresholds increase the NEP. As drivers at the 'real' thresholds are expected *to decrease the NEP during droughts, the identified thresholds are likely not relevant to drought. Instead, deriving these thresholds by taking values where SHAP values transition from positive to negative makes more sense. Furthermore, the SHAP values came from a single XGB model trained on the entire period of 2005-2022 (for NEP), which assumed that the NEP-meteorology relationship was stationary over the studied period. But if the drought acclimation of NEP exists, a shifted NEE-meteorology relationship is expected. Therefore, training the XGB models on each year separately (or several years using a moving window) and computing the SHAP values for each model might be a better option although the training dataset may be too small. Overall, the evidence of the acclimation of NEP to droughts is weak given the data availability and analysis, and thereby I suggest removing the associated results and conclusion if more convincing evidence is not found.*

Thank you for your valuable insights. Before we answer your comment, we would like to clarify that we defined *acclimation* as NEP sensitivity to abiotic factors like Tair, VPD, and SWC during each growing season (we have mentioned this in the discussion section (lines 438-447). But we will further clarify this aspect in the revised version of the manuscript, in accordance with what we are stating here.

In the original manuscript, we tried to show "seasonal acclimation" (Grossman, 2023), i.e., that the NEP response to changing environmental conditions differs among different growing season. For example, this means that during a particularly dry growing season, the SWC value corresponding to maximum NEP would be lower compared to a wetter growing season (Fig. 7 in the original manuscript). In our manuscript, we estimated this acclimation (NEP sensitivity) to Tair, VPD, and SWC by estimating the Tair, VPD, and SWC values that indicated maximum marginal contribution to NEP (feature_NEPmax), i.e., maximum SHAP values for the respective driver variable (feature) for different growing seasons. We will explain this better in the revised version of the manuscript.

In our original manuscript, XGB models were indeed trained for each year separately as recommended by the reviewer (we will further clarify this in the method section). Originally, we estimated optimum values of abiotic factors for NEPmax to understand if they would change depending on the environmental conditions among different growing seasons (Figs. 7 and 8). In addition, we now extended our analysis to test if feature_NEPmax values vary with the corresponding mean feature values during the same growing season (Figure R1.3), indicating acclimation of the forest to abiotic factors not only during but also among growing seasons. These (new) results indicated that for example, for the driest year 2022, the forest showed highest feature_NEPmax values at higher soil and air dryness (high VPD and low SWC values) compared to all other years. According to the definition of acclimation stated before, we argue that there was an acclimation of NEP to soil and air dryness during the 2022 growing season. Moreover, this acclimation differed among growing seasons, dependent on the environmental conditions during the respective growing season. Since our focus was on the three CSAD years, we only showed the results of these three years in the original manuscript. In the revised manuscript, we will now show this analysis for all growing seasons and compare the optimum abiotic factors with its seasonal mean values (as shown in Figure R1.3).

Nevertheless, we see the value of another definition of "sensitivity" of NEP acclimation to abiotic drivers over the years as suggested by the reviewer (i.e., the value of the driver variable when the SHAP values turn their sign from positive to negative or vice versa). However, this approach can result in significant biases in the results. SHAP values are in fact highly sensitive to the mean NEP value of each growing season (Gou et al., 2024). Thus, we might observe a vertical shift of the SHAP values among different growing seasons with an increase in NEP that in turn would result in a shift of the feature values to lower/higher feature values corresponding to the intersection of the curve with the SHAP zero line (see Figure R1.4). In contrast, the method we used, i.e., determining the maximum SHAP values (i.e., feature_NEPmax), is not affected by such a vertical shift in NEP. Thus, we favour this method over the intersection point method.

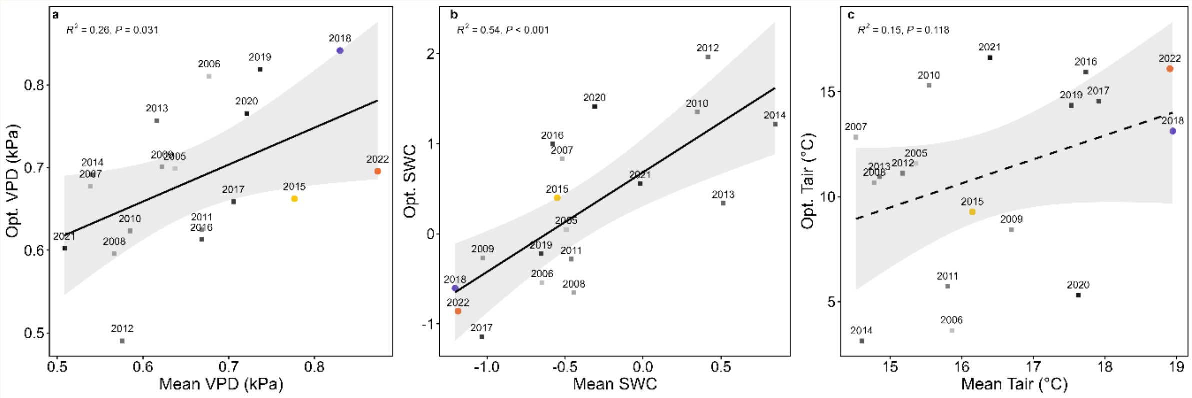

Figure R1.3. Relationships between mean VPD (a), SWC (b), and Tair (c) measured during the respective growing seasons and the VPD, SWC and Tair values of daytime NEPmax. The grey areas around the regression lines indicate the 95% confidence intervals.

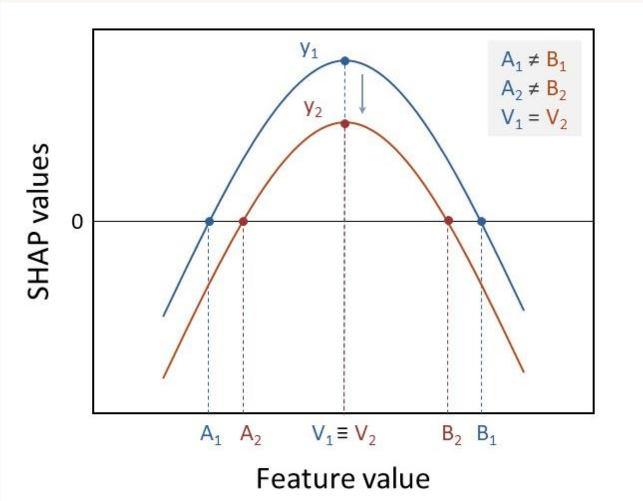

Figure R1.4. Illustration of an eventual vertical shift in the SHAP values (here $y_{1,2}$) that would result in a change of the intersection points of the curves with the SHAP zero line (here $A_{1,2}$ and $B_{1,2}$), but would not change the maximum points (here $V_{1,2}$) of the same curves.

2. Line-by-line comments

Line 22: What is the 30% decrease relative to?

The percentages given in the abstract refer to the decrease in the CO_2 fluxes relative to the long-term means. This will be rephrased in the “Abstract” and in the section “2.5 Data analysis”.

‘largely’ → ‘large’.

Here we meant ‘mostly’. We will replace ‘largely’ with ‘mostly’ to make it clearer.

Line 28: remove the second ‘always’; add ‘has’ after ‘net radiation’.

Thank you for the correction, the sentence will be changed to “Air temperature had negative effects, while net radiation had positive effects on daytime mean NEP during all CSAD events”.

Line 31-32: remove the sentence of acclimation if more convincing evidence is not found.

The sentence will be adjusted accordingly (see our responses above).

Line 61: ‘be it’? ‘particular’ → ‘particularly’

The sentence will be changed to "[…] above the canopy, i.e., net carbon dioxide (CO_2) exchange or remote sensing of vegetation, particularly the latter largely neglecting […]".

Line 107: please add the description of measuring CO2 storage change.
We will clarify it in the manuscript. The sentence will read: "The net ecosystem exchange (NEE) was then calculated as the sum of FC and the CO_2 storage term estimated from concentrations based on 1-point measurements (Greco and Baldocchi, 1996)".

Line 132: please report the depths.
The depths of the soil sensors were mentioned in the footnote of Table A1 "Forest floor meteorological station (profile measurements up to 50 cm depth)". In the revised version, we will mention the exact depths in Table A1 as "[…] (profile measurements at 5, 10, 20, 30, 50 cm depth)" and mention the depth range in the Methods section as well.

Line 134: How to centerly normalzied the SWC data
We used this term "centre-normalized" to not create confusion but were obviously not successful. In the revised version, we will replace this term with the term "normalized" and define it well. The term "centre-normalized" means that we normalized the data with mean = 0 and standard deviation = 1, using the equation $x_{norm} = \frac{x_i - \mu}{\sigma}$, where x_{norm} is the normalized value, x_i the measured value, μ is the mean value, and σ is standard deviation. This normalization is also known as z-score transformation.

Line 171: daytime mean NEP and daily mean NEP are easy to get confused in the many parts of the manuscript. Using 'NEPdaytime' and 'NEPdaily' could help.
Thanks for this excellent suggestion, which will indeed be a good way to make the reading and understanding of the manuscript easier. To better address these variables in the text and in the figures, we will now use the terms "NEP_{DT}" for daytime NEP and "NEP" for daily mean NEP.

Line 177: 'Shapley, 1953' is missing in the reference.
Thanks for the comment. The reference will be added to the reference list.

Line 182: Please clarify why the mean SHAP value instead of the mean absolute SHAP value is used to indicate the overall feature importance.
In the manuscript, we are using the SHAP analysis to understand the time course of the driver effects during the growing seasons of each of the CSAD years and respective CSAD events. Therefore, we calculated the SHAP values separately for each day of the growing seasons included in the general data set (2005-2022 for NEP and 2018-2022 for Rff). As we are interested in the dominating direction of the effect of each feature, we decided to calculate the mean SHAP value during the CSAD events and not the mean of the absolute SHAP values (since this would cancel the direction of the effects). We will add this info to the methods in the revised version of the manuscript.

Line 189-195: please refer to Figure 7.
Thanks for the suggestion, however this sentence is part of the methods section, thus, we cannot refer to a figure which comes much later in the results section.

Line 212-217: All the events' length seem to be 1 day shorter. Same in table 1. Please check.
Thank you. We will correct the lengths in the text and in Table 1 (CSAD 2015 will be 15 and 12 days with a total of 27 days, CSAD 2018 will be 32 days, and CSAD 2022 will be 22 days).

Line 237-204: What are those shade areas around dashed lines?

The shaded areas around the dashed line are the standard error of the mean 2005-2022. The new caption will read "Figure 2. Comparison of 5 day moving averages of daily mean (a-c) Tair, (d-f) VPD, and (g-i) SWC in the years when a CSAD event occurred, against the long-term mean (2005-2022). The area around the dashed line indicates the standard error of the mean 2005-2022. The coloured areas mark the CSAD events, i.e., periods with co-occurring lowest SWC and highest VPD".

Line 241-243: Why Max. or Min. has a standard deviation?
As we are referring to daily means of each of these measurements in Table 2, the standard deviation was derived from the standard deviation of the measurement for the same day. For instance, we found that the maximum daily mean air temperature in 2015 was 26.9°C, with a standard deviation of 3.03°C for the same day.

Line 266-272: 1) What are those shaded areas around dashed lines in the left panels?
The dashed areas around the dashed lines in Fig. 3 are the standard errors of the long-term means. We will add this to the figure caption in the revised version of manuscript.

2) What are those error bars in the right panels?
The error bars in the right panels of Fig. 3 are the sum of the standard errors of the long-term means of the fluxes for the days of CSAD events. For instance, the error bar around GPP in 2018 is the sum of the standard errors of the long-term mean GPP between the days of the year 204-235 (CSAD event 2018).

The standard errors in the plots will be specified in the caption as follow: "Figure 3. [...] against the respective long-term means (a, c, e, g). The grey areas around the dashed lines refer to the standard error of the respective long-term-mean CO_2 fluxes. [...] CSAD event and the respective long-term mean fluxes (2005-2022 for NEP, GPP and Reco; 2019-2021 for Rff). The error bars show the cumulative standard errors of the long-term mean CO_2 fluxes for the same period".

Line 358-359: How to calculate this standard deviation?
It is the residual standard error of the model. The values are an output from the loess object obtained with the fitting of a loess model from the function loess in the *stats* package (version 4.3.3 R Core Team). We will add this info in the methods section of the revised version.

Line 371-373: Since SR vs. TS and SWC during CSAD are not significant, please rephrase 'tend to decrease or increase' as 'non-significant'.
Thanks, the text will be changed as follows: "However, during the CSAD event, SR did not respond to TS (R2=0.19; Figure 9a). On the other hand, independent if a CSAD event was recorded or not, daily mean SR did not respond to variations in SWC (R2 < 0.01 and R2 of 0.3, respectively; Figure 9b)."

Line 385: If still keep 'acclimation', please briefly describe what acclimation is here.
We will add further information about the whole issue of acclimation. See our responses above. Here, in the first paragraph of the discussion, we will rephase the sentence as follows: "In addition, we found first signs of acclimation to such CSAD events within the same and among different growing seasons, based on the shift in the maximum marginal contribution of VPD and SWC to NEP. This also suggests that predictions of site-specific CSADs and their impacts might become more challenging in the future".

"Line 419-421: You found air temperature is not important for daily mean NEP during CASD based on conditional variable importance in Figure 4, while air temperature is still important for daytime mean NEP during CSAD based on SHAP in Figure 5. Although daily mean NEP and daytime mean NEP

are different, the results using the two methods seem to be inconsistent. Therefore, please also report the overall feature importance of predicting daily mean NEP based on SHAP.

Thanks for this comment. As it refers to the issue addressed above (see also the new Figures R1.1 and R1.3), this sentence will change. We will address the new analyses and also use further references (e.g., Granier et al., 2007).

Line 438-447: again, suggest removing if more convincing evidence is not found.

Thanks for the suggestion of an alternative interpretation of our original Figures 7 and 8 (see comments above). As mentioned before, the acclimation or sensitivity of NEP to abiotic drivers based on our analysis is done with the detection of the feature value corresponding to the maximum contribution to NEP (feature_NEPmax) and how these values change among different growing seasons. We indeed found that the VPD and SWC values corresponding to NEPmax were varying depending on the mean feature value of the same growing season, suggesting that there is acclimation to these conditions within the same growing season (see Fig. R1.3 above).

Line 492-493: same as above.

Thank you. We will expand this analysis and its discussion in the revised version of the manuscript (as outlined above). The conclusion will not change, but we will clarify the results and thus better explain this conclusion in the revised manuscript.

References:

Greco, S. and Baldocchi, D. D.: Seasonal variations of CO_2 and water vapour exchange rates over a temperate deciduous forest, Global Change Biology, 2, 183-197, 1996.

Grossman, J. J.: Phenological physiology: seasonal patterns of plant stress tolerance in a changing climate, New Phytologist, 237, 1508-1524, 2023.

Gou, R., Chi, J., Liu, J., Luo, Y., Shekhar, A., Mo, L., and Lin, G.: Atmospheric water demand constrains net ecosystem production in subtropical mangrove forests, Journal of Hydrology, 630, 130651, 2024.