Schackow et al.

Reply to Editor

Thank you for your patience with the review process of this manuscript. We have received comments from two domain experts. Both think the manuscript holds merit, but have several comments.

Both reviewers comment on the issue of generality of results due to the study of a few sites, which should be addressed.
We acknowledge that the number of sites is relatively small. Unfortunately, this is an inherent limitation of in-situ measurements and not something that can be easily overcome. Nevertheless, we now include a figure showing the correlations of the two proposed metrics with hydrometeorological variables for all 15 sites (new Figure 4) and another figure showing the anomalies in normalized slope and area for days with extreme hydrometeorological conditions across all sites (new Figure 6).

Reviewer 1 also raises concerns about the clarity with the discussion of timescales. I also agree with them that the manuscript could go further to links to prior findings about sap flow-VPD hysteresis.
We have now rewritten the introduction and have added explicit mentions to previous studies pointing out hysteresis.

Reviewer 2 points out potentially unclear objectives. On this point, I do wonder if sufficient motivation and/or investigation of alternatives was provided about the chosen slope and area metrics to address the stated questions. Are there other metrics that should be considered? If not, are these sufficient to understand stress?
We have tested different metrics in preliminary stages of this study, namely the afternoon slope, and the slope between peak-SF and peak-VPD or the timing of these two points. We have concluded that AREA and morning SLOPE were sufficient to capture the dynamic responses of SF to hydrometeorological stressors, especially extreme events, and thus decided to focus on those here. We do not consider these preliminary analyses to add meaningful value to the manuscript given its objectives. We have, however, added a note in the methods that other metrics have been evaluated in preliminary analysis:
Line 190-195:
As discussed in the introduction, we hypothesise that the sub-daily dynamics can be used to understand plant stress, and derive two metrics at a daily basis to characterize the diurnal SF–VPD relationship: the magnitude of the morning slope of the SF–VPD curve, referred to as SLOPE, and the area of the hysteresis loop (Zuecco et al., 2016), referred to as AREA. Other metrics were tested in preliminary analysis, such as the magnitude of the afternoon SF–VPD slope, or the temporal lags between peak SF and peak VPD (analogous to e.g., Wan et al. (2023)), but we found the information in these metrics to be strongly correlated, and thus redundant, with the two presented here.

We consider that our results support the hypothesis that these two metrics allow us to distinguish plant responses to hydrometeorological stressors. We hope that our study can motivate further evaluations where more refined information about plant stress (for example plant water potential) might be available. Also, see new details added to RC1C2.

Additionally, one could understand water stress from space at daily and longer timescales with plant functioning-soil moisture relationships, making it easier for remote sensing to capture lower frequency sampling needs. More motivation should be provided about additional information held within diurnal sap flow-VPD relationships.
We do not deny that longer time-scale analyses have value, but it has been pointed also by others that variability in VWC at diurnal time scales carries specific signatures of plant functioning, better targeted to detect vegetation stress, for example Konings et al. 2017. This aspect is now discussed more explicitly, and including more references in the revised version of the introduction (Lines 49-64).

Figure 6 is an interesting contribution and the acronyms should be provided in the caption for clarity.
Thank you, we have now added the information to the figure caption to improve clarity. (now Figure 7)

Reviewer 1

General comments

RC1C1: In their manuscript "Vegetation health monitoring based on sub-daily sap flow variability", Tschackow et al. explore whether high-temporal-resolution sap flow data and its relationship to climatic variables can serve as an early indicator of plant stress. Using SAPFLUXNET data, they link sub-daily sap flow variability to environmental drivers. While the research question is interesting, the presented results only partially support the stated objectives. My main concerns are outlined below.

We thank the reviewer for the constructive feedback.

RC1C2: 1) The rationale for selecting the three focus sites is not sufficiently justified. Beyond identifying patterns absent in the broader dataset, it remains unclear how these sites advance the overarching objective. The connection between the analyses based on the three sites and the global dataset is weak. The authors note (Lines 282–283) the absence of consistent patterns across latitudes, yet proceed to discuss relationships observed only at the three selected sites. This raises concerns about the validity and generalizability of these interpretations—particularly since some sites exhibit near-zero relationships between SLOPE and the key drivers (TSM, PPFD, and Tair). As the study aims to identify patterns that could inform remote sensing–based upscaling of vegetation health monitoring, the lack of consistency across the broader dataset suggests that this objective has not been achieved. To strengthen the manuscript, the authors could consider focusing exclusively on the three detailed sites and clearly state that the conclusions are site-specific. While this would limit the potential for upscaling, it would ensure that the conclusions are well supported by the data.

We show only results for 3 selected sites in Figures 3 and 5 for clarity and conciseness in the description of the metrics and key results of the study, as we cannot analyse all sites in detail in the main manuscript. The three sites were selected to cover a range of background climate conditions and biomes, ranging from humid tropical regions (French Guyana), to semi-arid mid-latitude regions, at a Mediterranean site (Puéchabon) and to a cold permafrost region (Pogorelsky Bor). The rationale for selecting these specific sites was described in detail in Lines 137-143.

In Lines 282-283 (first manuscript) we did not report lack of consistent patterns, but rather that the strength of the relationships varies across sites. This is a very common feature when dealing with still relatively small samples of sites (33 in this case, now reduced to 15, note that we now focus only on sites with both TAir and TSM data) that correspond to different climate zones, tree species, soil types, etc. Such large differences across sites and/or species are also found in the the results of the reference provided by the reviewer, Wan et al., 2023 and Kumar et al. 2023. We note, however, that the relationships in AREA hold beyond the three sites, as demonstrated in the new Figure 4. The differences in SLOPE can be explained by differences in the limiting factors for downregulation across biomes. We do not consider this to be a limitation of these metrics, but rather an indication that they correctly capture the expected responses to hydrometeorological stressors. Furthermore, the manuscript aims to understand whether these two metrics can be reliably derived from coarser sub-daily measurements. The results in Figure 7, shown for all sites in the boxplots, allow to identify 6- or 8-hourly temporal sampling rates as sufficient to capture variability in the AREA and SLOPE metrics across globally distributed sites. We have added an analysis of the AREA and SLOPE for extreme hydrometeorological conditions (as shown in Fig. 5 for the 3 sites) for all 15 sites to evaluate whether the features found apply more generally. This is now included as new Figure 6, reproduced below (Figure R1):

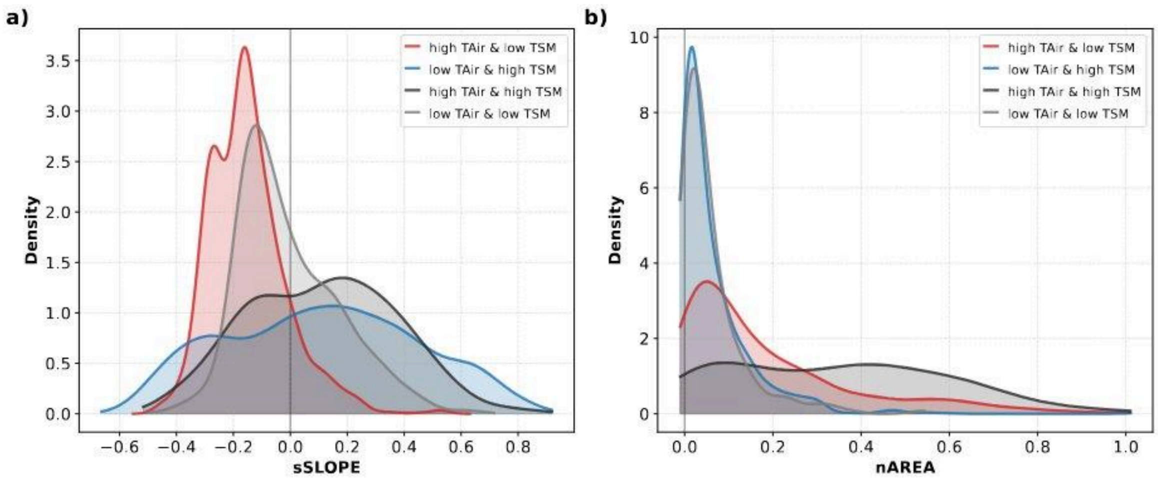

Figure R1: Distribution of sSLOPE (A) and nAREA (B) at extreme anomalies of TAir and TSM for all 15 sites. Red curves correspond to hot and dry days (high Tair and low TSM), blue curves to cool and wet days (low Tair and high TSM), black curves to hot and wet days (high Tair and high TSM) and grey curves to cool and dry days(low Tair and low TSM).

We show that sSLOPE is predominantly negative during dry conditions (red and grey curve) and positive during wet days (blue and black curve), indicating that SLOPE can indeed distinguish between days with high/low soil water availability (the supply side). The distribution curves of nAREA stays close to zero for cold days (blue and grey curve), shifting to higher values under hot and dry conditions (red curve), while hot and wet days (black curve) show the highest values of nAREA. This indicates that AREA is a good proxy for midday regulation in response to high temperatures and atmospheric dryness. Both metrics combined allow to fingerprint vegetation responses to high atmospheric demand and low soil water supply.

RC1C3: 2) The distinction of timescales at which responses are assessed is not very clear: sometimes there is no mention of timescale–which can confuse the reader– and sometimes daily and diurnal are mentioned interchangeably. Often the sub-daily and diel processes are not explained distinctly from daily processes.

We carefully revised the text to ensure that the distinction is made clear. Also see replies to some of the inline comments below. For clarity, we write hourly for the time-steps considered to derive the SF-VPD hysteresis curves. When referring to these curves and the metrics derived from those we refer to “diurnal cycle". The metrics themselves have daily time step, so we explicitly added this note in the methods.

RC1C4: 3) The concepts of "stress" and "vegetation health" are insufficiently defined. It is unclear how the authors determine that exceeding a certain area or threshold constitutes stress. This definition appears to lack physiological justification and is not validated against independent measurements of plant stress (e.g., water potential, leaf turgor, or critical hydraulic thresholds).
We agree with the comment, and reformulate to "vegetation responses" to environmental factors (for the long-term relationships) and stressors (for the extreme days).

RC1C5: 4) Previous work (e.g., Wan et al., 2023, Agricultural and Forest Meteorology) has shown that sap flow–VPD hysteresis is related to sapwood area, yet this factor is neither mentioned nor accounted for. Additionally, the temporal progression of hysteresis (e.g., clockwise vs. counterclockwise loops and their changes over time) is not analyzed. The discussion would benefit from a clearer linkage to prior studies on sap flow hysteresis and environmental drivers. Since hysteresis in sap flow is well documented, the novelty of this finding in the current manuscript remains limited. Without robust evidence that these relationships hold beyond the three case sites, the reliability of these proposed metrics for general vegetation monitoring remains uncertain.
We agree that hysteresis and sap flow is well-documented, and we did not mean to imply that the novelty of this study lies on reporting hysteresis, since we have cited previous works that discuss sub-daily water regulation strategies and that highlight the need for diurnal observations to better characterize forest responses to environmental stressors (e.g. Landsberg and Waring, 2016, Konings et al., 2021, Chakir et al. 2021).
The temporal progression of SF-VPD is given in Figures 3 and 5. In most days across sites the progression is clock-wise. Nevertheless, if the progression is counterclock-wise, this is still reflected in our metrics that focus on the morning slope (regardless of whether it is stronger or weaker than in the afternoon).
We thank the reviewer for the reference to Wan et al. 2023, who analysed how the absolute value of the time-lag between SF and VPD is related with sapwood area. They show a good relationship between time-lag and area of the hysteresis for the SF and incoming solar radiation curve, not to VPD, and only for three selected sites. While it is expected that the time-lag relates to the AREA, it might not fully reflect differences in the afternoon trajectories of SF-VPD that further influence the magnitude of AREA. Furthermore, this factor is likely relevant to explain differences in the absolute magnitude of the AREA values, but not necessarily their relative differences (i.e. the standardized metrics), especially for the evergreen sites, nor the variations in SLOPE (the same time-lag as analysed by Wan et al. can correspond to very different values of SLOPE). Importantly, Fig. 9 in Wan. shows large differences in the magnitude and even the sign of the relationship between the time-lag of SF and VPD and sapwood area across different species. Specifically, for Quercus ilex in FRA (FRA_PUE), there is no relationship between the time-lag and the sapwood area. Therefore, we do not consider that the analysis in Wan et al. invalidates our results in any way.
Sapwood area was not considered here because, as highlighted by Wan et al. (2023), sapwood area explains the variation among trees while we are primarily interested in temporal dynamics and temporal variability. We will however add a reference to the study when discussing the

differences in the absolute magnitude and when comparing across sites where the mean sapwood area, and also the seasonal variation in active sapwood area, may contribute to variability across sites.

Finally, the novelty of our study lies in demonstrating that two simple metrics that can in principle be derived from satellite measurements with few observations per day (the morning slope and area of SF-VPD curves) show characteristic responses to hydrometeorological conditions, and especially extremes, across sites in different climate zones. The relationships in AREA hold beyond the three sites, as demonstrated in Figure 4. The differences in SLOPE can be explained by differences in the limiting factors for downregulation across biomes. We do not consider this to be a limitation of these metrics, but rather an indication that they correctly capture the expected responses to hydrometeorological stressors. We have revised the motivation of this study accordingly (lines 100-109):

> *We then propose a reduced set of descriptors that can be derived from the diurnal hysteresis cycles, and that could be estimated using comparatively sparse sub-daily observations of VWC, or some proxy for it, from satellite remote sensing. We then examine the seasonal and interannual evolution of these descriptors at selected sites to understand how they vary in response to hydrometeorological conditions and particularly extremes. Finally, we compare several observation scenarios to investigate how well these metrics could be determined with sparser data from satellite remote sensing. We test the degree to which it is possible to capture the key characteristics of SF–VPD hysteresis with targeted sub-daily measurements using experiments characterizing hysteresis in SF with plausible data acquisition strategies, which is important for understanding the minimum requirements for observation frequency if observing sub-daily hydraulic dynamics with satellites. This would provide an opportunity for global monitoring of vegetation health and enable the early detection of vegetation stress before structural decline or canopy-level changes become detectable.*

RC1C6: 5) Lastly, the conceptualization (e.g., Figure 1) does not include bimodal sapflow peak responses and thus excludes species that show a midday depression in sapflow due to high VPD (see for example Kumar et al. 2023 Journal of Experimental Botany). To ensure conceptual completeness and global applicability, the authors should revise their framework and discussion to incorporate such species-specific responses.

Kumar et al. (2023) analyse three pioneer and emergent species in a secondary tropical montane forest in Eastern Himalaya, which are not included in our datasets, and only two of those show a mid-day depression, so that it is not necessarily a general feature. In the data we used, generally sapflow is reduced around midday, depending on the time of year and environmental conditions, but not necessarily depressed.

Finally, the scheme is purposely simplified, since it’s an idealised depiction about how to link satellite measurements of VWC, sapflow, and the metrics analysed here.

Specific comments:

Title: vegetation health or vegetation water stress?

Thanks, we have changed the title to *"Detection of tree stress from sub-daily sap flow variability"*, to better match the objectives of the paper regarding detection from other monitoring platforms.

Line 6: diurnal or sub-daily? Diurnal is not about sub-daily variations but variation from one day to another
A "diurnal cycle" corresponds to a pattern that recurs every 24 hours (see definition), so we believe this is correct. However the "sub-daily" was redundant, so the sentence has been changed to:

> *Based on SAPFLUXNET measurements of sap flow across tropical, temperate and boreal biomes, we demonstrate how variations in the diurnal cycle of sap flow as a function of vapor pressure deficit (VPD) measurements can elucidate the different levels of plant hydraulic stress.*

How is stress defined? And how is health and stress related? How do the authors define poor or good vegetation health? (at which stress level and why?)
We agree that given the data used, we can only refer to stomatal downregulation as a proxy for hydraulic stress. We carefully revised the manuscript to avoid confusion.

Not clear from the Abstract what species or climate region are the focus of the findings
Added, see reply to Line 6 comment.

Line 10: soil moisture at what time scale? Daily? There is typically a lag between stem refilling at night. How do the authors then relate the instantaneous sapflow response to instantaneous soil moisture?
Thanks for pointing out the imprecision, this refers to the seasonal patterns, the sentence was corrected to:

> *We find that seasonal variations in the morning slope are positively associated with top (0-30cm) soil moisture.*

Line 12: at what time scale?
Same as above. We have now revised the sentence:

> *The area of the diurnal cycle, characterizing the degree of daily hysteresis between sap flow and VPD, increases with sap flow downregulation before peak VPD and is sensitive to temperature and soil moisture variability at seasonal time scales.*

Line 17: sub-daily hysteresis?
Yes, fixed.

Is Figure 1 a graphical abstract? Otherwise, shouldn't come before the Introduction
No, Figure 1 aims to illustrate the concept that underlies the study, specifically how the diurnal cycle of sapflow and vegetation water content can be monitored at key times of the day to detect hydraulic stress. But we agree that the figure is misplaced and have moved it down.

Figure 1 caption: what determines stress with respect to the hysteresis? Panel a does not describe the concept: What differentiates the different circles? The different curves and the satellites? What satellites are these? Each panel should be described clearly (especially there is no explanations of panels c and e). There is a mention of stress on the figure, but is this a specific event (e.g., point in time) or in general? What is then the conditions shown in the other panels where no stress is mentioned?

The description of the stress and relationship with hysteresis is found in the main text as it is quite lengthy (L80-92). However, we agree that the caption was not detailed enough and will add more information about the different elements:

> *Panel a) illustrates generalized diurnal cycles of SF (blue) and VWC (orange) and the critical times of the day at which a satellite could monitor VWC (midnight, 6am, noon, 6pm, from left to right). Panels b)–e) illustrate the hypothesised diurnal relationships of SF–VPD and VWC–VPD for average conditions (b), d)) and for days marked by hydraulic stress (c), e)). The circles in panels b)–e) show the times of the day represented in panel a). The different processes driving these responses are described in Section 1.*

Line 46: stomatal conductance is only one of the mechanisms that determine plant water uptake, the stem capacitance and refilling of water between day and night provides a buffer to whole plant water uptake which needs to be mentioned as well.

Agreed, we have changed to:

> *Diurnal fluctuations in VWC and SF reflect normal operating conditions of xylem water transport and are <u>regulated by vegetation water storage, refilling and stomatal responses,</u> which balance carbon uptake and transpiration-driven water loss (Jarvis and McNaughton, 1986b).*

Line 49 the statement needs a reference. See for example Zweifel at al. 2021 New Phytologist.

We have added the citation for the earlier review by Tyree and Ewers (1991, New Phytologist, 119, 345-360).

Line 56: the caption does not explain that this is what the figure is showing.

Indeed, but the description is very lengthy for a caption. We have added the following sentence in the caption:

> *The different processes driving these responses are described in Section 1.*

Line 57: be specific please which plant hydraulic functioning you mean?

Removed this expression in the revised version of the introduction.

Line 80: SAPFLUXNET is not a running network like FLUXNET or ICOS. It provides a data product with no commitment to extend beyond 2018.

Indeed, but we did not imply that it is. The sentence read:

Currently, processes associated with plant functioning such as plant productivity and evapotranspiration are monitored at sub-daily temporal resolution in networks of sites such as FLUXNET (Pastorello et al., 2020), ICOS (RI, 2022), or SAPFLUXNET (Poyatos et al., 2021).

Line 85: how do the authors link vegetation water content to sapflow where instantaneous water content is a product of sapflow and stem capacitance. There is no mention of the role of stem capacitance in the whole paper which is an essential component of the plant water uptake.
We do not analyse nor model VWC, we only aim to explain that in principle VWC observations from satellite could be used to derive and analyse SF dynamics. We have, however, revised the introduction and removed the sentence.

Figure 2: What is the range of years for different symbol sizes? How should the readers interpret the number of years from the size of symbols?
We agree that it was not clear, we will added a label for the smallest and biggest symbol and the corresponding temporal coverage.

Please explain what is the aim of selecting three focal points; what analysis here is different from the 33 study sites and why a subset of sites was needed? Were there additional info available at these three sites that were not available in other sites?
The reason to select a subset of sites was mainly to ensure clarity and conciseness in the description of the result of study, as we cannot analyse all sites in detail in the main manuscript. Furthermore, the three sites have long temporal coverage and were selected to cover a range of background climate conditions and biomes, ranging from humid tropical regions (French Guyana), to semi-arid mid-latitude regions, at a Mediterranean site (Puéchabon) and to a cold permafrost region (Pogorelsky Bor). The rationale for selecting these specific sites is described in detail in Lines 144-149.
However, we would like to clarify that there was a mistake in our previous manuscript. While 33 sites include Tair, only 15 include both Tair and TSM. Therefore, we have now limited the analysis to the 15 sites and Figures 1 and A1 have been updated.

Lines 150-184 please give consistent level of information for each site: MAT, MAP, elevation, species botanical names.
Thanks for pointing out, we revised the description.

Line 165: how high is “high temperature”?
We meant MAT above 25ºC, this was added in the revision according to the previous comment.

If possible, please report mean site LAI (any information about canopy cover is important for interpreting forest water use and if not directly measured, can be extracted from satellite products). Please also mention the DBH range for the trees used in the analysis for each site as the size of the trees is a determining factor.

We now report the DBH range and site LAI for the trees as given by the SAPFLUXNET metadata.

Case studies should come in the Methods section not before.
However, the section mostly describes the background climate and data available for each site, so we believe it should remain in “Data”. We rename the section to “study sites” for clarity.

Line 177: how deep is “great depths”?
The sentence was not clear, we mean that most water is stored as ground water, i.e., not accessible to plants. We revised the sentence to:

> *Overall, evapotranspiration exceeds precipitation in this region (Urban et al., 2019), with water stored mostly as ground water.*

Line 179: by site age you mean age of the trees?
The sentence has been revised to:

> *The trees at the experimental site in Pogorelsky Bor have an average age of 50 years and feature a mix of six deciduous larch and three evergreen pine trees.*

Line 190: but Figure 1 includes also soil moisture and not only VPD. Please stay consistent with describing the objectives.
Figure 1 does not include soil moisture. But we agree that an additional sentence can be added for clarity:

> *We then analyse how this relationship varies across the growing season and over multiple years, and how it relates to other environmental variables, including temperature, soil moisture and radiation, as described below.*

Line 207: Figure 4, Figure A2
Removed, in reply to next comment.

Line 207: Please remove reference to result, in the Methods section
Done.

Line 209: upper and lower 20 % of the which data? Were the extremes defined per site?
Corrected to:

> *[...] (upper and lower 20% of all days within the growing season at each site)[...]*

Line 210: what is meant by each sample? And why eight?
Each sample corresponds to a given day, which we think is now clearer with the correction in line with the previous comment. The number eight has been chosen to avoid selecting clusters with too few points (thus unrepresentative), but given that we already select the days based on the percentiles of the marginal distributions of the two variables we have removed this threshold in the revised version of the manuscript.

Line 214: typo in see 1 (should be see Figure 1?)
Yes, thanks.

Figure 3 caption should explain what TSM is.
We are not sure what the reviewer means, since the acronym is explained in the caption:

> *Seasonal dynamics of photosynthetically active portion of incoming radiation (PAR, quantified as PPFD), air temperature (TAir) and* ***top soil moisture (TSM),*** *AREA and SLOPE with interquartile ranges showing the range of the hydrometeorological drivers and the derived metrics per day across years.*

The description on what the measurements correspond to are provided in Section 2, which we believe is appropriate.

Figure 3: X-axis label is missing.
Thanks, we did not include it initially to not crowd the figure even more, since the labels correspond to the months in the year, it is added now.

Figure 3: In the lowest panel right, why is the slope and area missing for some days within the growing season?
The data for this site is only available during the growing season, however, even in some of those days there is missing data in the time-series.

Figure 3 Left panel: please explain the caption what the vertical and horizontal dashed lines mark.
Added: *The dashed vertical and horizontal lines indicate the maximum values of the mean seasonal cycles of VPD and SF, and the corresponding average time of the day.*

Figure 3: be explicit please in the caption which variables, instead of writing hydrometeorological drivers.
We do not follow this point, as they are listed in the same sentence:

> *Seasonal dynamics of photosynthetically active portion of incoming radiation (PAR, quantified as PPFD), air temperature (TAir) and top soil moisture (TSM), AREA and SLOPE with interquartile ranges showing the range of the hydrometeorological drivers and the derived metrics per day across years.*

Figure 3: displayed PPFD is also diurnal? Please clarify in the caption by adding "daily" before each variable.
Added *[...] quantified as daily PPFD [...].* The units will be added to the figure.

Lines 224-239: this part reads as a mixture of a repetition of the introduction and what could potentially be the Discussion. Please first report on results.
Thanks for pointing out, we shortened considerably and focused on describing the results.

Lines 301-312: please indicate where the readers can see these results.
We added a reference to Fig. 5.

Figure 4 please indicate on the matrix, if any of these correlations are not significant.
We have updated the figure with this information, as reproduced below:

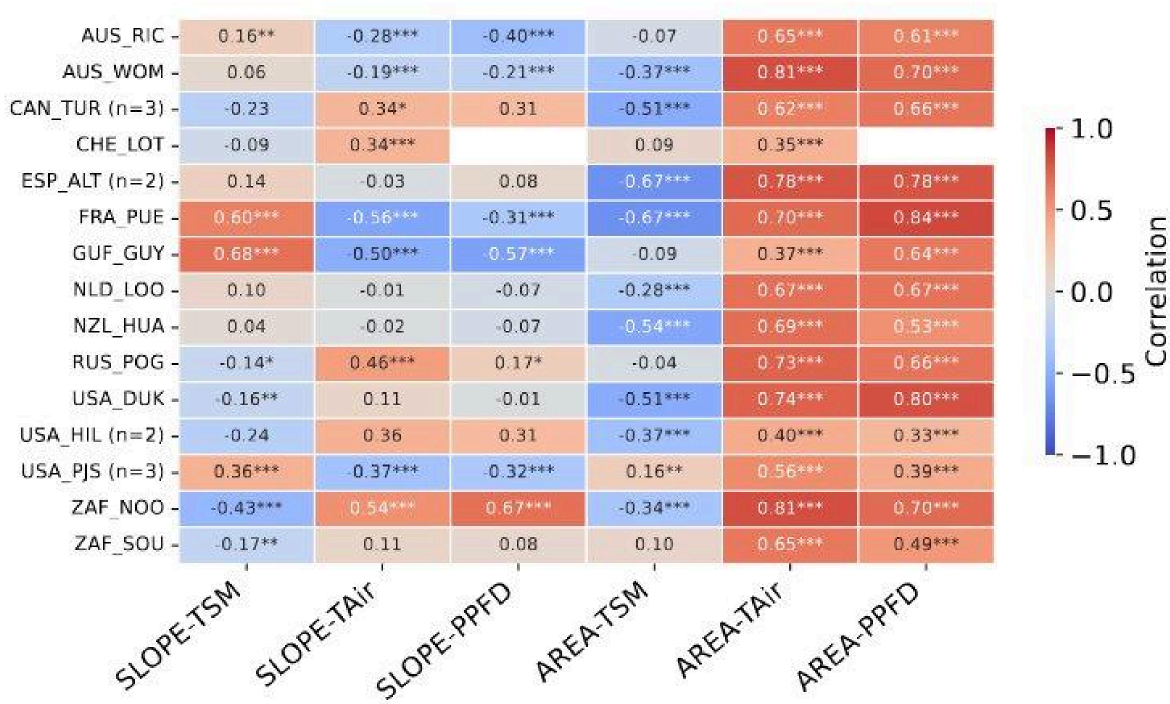

Figure R1.1: Heatmap of correlation between the two metrics and Tair, TSM and PPFD, now with the significance of the correlations indicated.

Overlay of results on the global map, what is the added info to display these on a global map?
Our intention was to facilitate interpretation of differences between sites given background climate, vegetation type, etc. We can present the results in another form, for example a table, if the editor agrees it is more appropriate.

Figure 5: are these slopes all significantly different from zero? Please indicate if they are, and remove those which are not.
We decided to show all points to better represent the full distributions of T vs TSM, in grey, and highlight the selected days by providing Area and Slope information. We will revise the figure to ensure that days with non-significant values of Slope are either excluded or represented with a different marker.

Why is it that the displayed VPD in Figure 5 reaches 3 and 2.5 kPa for the warmer sites, but maximum VPD in Figure 3 doesn't reach that maximum? Please also indicate if VPD and TSM in Figure 5 are mean daily values.
This is because the cycles in Figure 3 correspond to mean values while in Figure 5 we are selecting subsets of days corresponding to more extreme environmental conditions (including

higher VPD). TSM values are daily and VPD is shown for hourly means, it will be added in the caption.

Line 251 and 313: statistically how is the breakpoint determined? The information should be added to the Methods section.
There is no formal breakpoint detection, the purpose here was to describe the sudden change.

Lines 313-319: where do the readers see these results?
In Figure 5, as indicated at the start of the paragraph. We now add letters for clarity.

Please show the breakpoints on the figures.
The breakpoint analysis has been removed.

When deriving the breakpoints, are they estimated based on all slopes, or only the statistically significant ones?
The breakpoints were not derived in a quantitative way, but rather discussed in a qualitative manner as the curve of the mean diurnal cycle departs from the morning slope.

Lines 331-345: text repeats from the Introduction/Methods, and does not belong to Results.
It has been removed.

Line 342: "daily" frequencies
It is actually sub-daily (3-8h), we have added for clarity.

Line 345: of which two variables?
SLOPE and AREA, as mentioned in the second half of the sentence. However, we revised the paragraph for clarity.

Figure 6: caption mentions panels A and B, but these are not marked on the Figure.
We have corrected the figure.

Figure 6: x and y-axis values are missing
Corrected, thanks.

Figure 6: the legend does not include all items that are displayed on the figure (e.g., the solid lines with different colors)
We have removed the legend and indicate all elements in the caption.

Figure 5 and Line 344: why only FRA_PUE? It is not clear.
FRA_PUE was used to illustrate the approach for deriving the R^2 per site. This has been added to the caption.

What is the hypothesis underlying the testing of whether r2 between hourly and daily metrics match?
As mentioned in the caption, and the text, the R^2 reflects how closely the seasonal dynamics of the metrics based on resampled time series track those of the hourly reference across sites.

Line 373: the concept is not new and is well-established, to look at the SF-VPD hysteresis rather than absolute SF. The authors should at least refer to previous studies and discuss their findings in the context of existing knowledge.
We have re-written the paragraph to better focus on the novel aspect of this study - how the proposed metrics can be used as proxies for vegetation responses to changing hydrometeorological conditions.

Lines 375-376: please indicate if this observation holds across all sites or only for a selection of sites, and then adjust the conclusion accordingly.
Clarified.

Line 377: please indicate across which sites, since the study includes both the analysis of a selection of sites, and analysis of all sites.
Thanks, we have clarified it.

Lines 393-397: this repeats text from the Introduction.
The text itself is not repeated, rather we added these sentences to remind the reader about the motivation for deriving the two metrics. We have shortened the sentence.

Line 401-404: Firstly, finding a hysteresis in sapflow data is not new (numerous studies show this) and whether it relates to soil moisture or VPD is not confirmed in this study, apart from on a selection of three sites. So I am not sure how reliable these metrics are for making such conclusions like this statement.
We have removed the sentence since the goal of this paragraph was to focus on how the two metrics can be estimated by coarser temporal data. Furthermore, we now include results from Figure 5 for all sites in Figure 7.

Line 416: which site?
FRA_PUE, which is indicated at the beginning of the paragraph. We have added the acronym for clarity.

Line 419: please provide a reference for this statement.
We have included two references.

Line 420: based on what results do the authors conclude a change in vulnerability over time?
Thanks for pointing out, we meant exposure.

Line 421: what do you mean by "novel climatic stresses"?

Replaced by “more severe and compound climatic stresses”.

Line 424: Figure 4? There seems to be a typo.
Yes, corrected.

Line 426: typo? Figure 5? Same in line 434 and line 435.
Yes, this has been fixed throughout the document.

Line 438: “snow” melting?
Yes, added.

Line 440: what is the reference for the statement that says seasonal drought is often outside the growing season?
It can be seen in the mean seasonal cycle of TSM in Figure 3. We added a reference to the figure.

Line 449: abbreviation (VOD) not explained yet
Added vegetation optical depth.

Reviewer 2

General Comments:
Tschackow et al. explore high temporal sap flow data and its relationship with VPD, temperature, and topsoil moisture across three different sites. In general, the work shows promise but needs more exploration or refining the objectives based on the figures shown.
We thank the reviewer for the feedback and hope that we have addressed the issues raised.

My major concerns are 1) lack of rationale for selecting three sites, and why such different sites. This work would be more impactful if it includes more than one station per ecotone, which exists based on Figure 2.
The reason to select a subset of sites was mainly to ensure clarity and conciseness in the description of the result of study, as we cannot analyse all 15 sites in detail in the main manuscript. Furthermore, the three sites were selected to cover a range of background climate conditions and biomes, ranging from humid tropical regions (French Guyana), to semi-arid mid-latitude regions, at a Mediterranean site (Puéchabon) and to a cold permafrost region (Pogorelsky Bor). The rationale for selecting these specific sites is described in detail in Lines 144-149. Furthermore, we do provide results for the 15 sites in Figures 4, 6 and A2. These cover multiple ecotones. We have extended the analysis of responses to extreme hydrometeorological conditions shown in Figure 5 to all sites, see RC1C2.

2) Objectives are not explicitly clear, nor what is defined as stress or a hot/cold day.
Our objectives / guiding research questions were stated in the Introduction, L95:

"Our analysis is guided by two central research questions: (1) To what extent can plant responses to hydrometeorological stressors be characterized using descriptors of diurnal SF–VPD hysteresis? (2) How often, and at what times, would sparse observations be sufficient to capture these stress signals?"

The definition of hot/cold conditions was defined in the Methods, L208:

"Since we want to evaluate plant functioning and water fluxes under environmental stress, we analyse combined extremes of TAir and TSM (upper and lower 20 % of all days within the growing season at each site). We obtain the following four extreme conditions: cold & wet, hot & wet, hot & dry and cold & dry."

However, we will revise the text to ensure that the objectives and methods are clear to all readers.

3) On line 14 the authors mention that the data is temporally sparse, which might lead to lack of alignment across data sets but do not tackle in the methods how corrections were made to account for that.

This might be a misunderstanding, we do not state that the sapflow data is sparse, but rather that we test approaches to derive similar information from more temporally sparse datasets.

4) The writing would benefit of some revision, particularly on paragraph structure, correct citations, and improving clarity. For example, line 496 is a one-sentence 'paragraph' that does not relate to the rest of the conclusion.

Specific comments:

Definitely improve flow on the abstract, making sure transitions are clear.

§ More concise

We proofread and revised the abstract for clarity.

§ There is also no clear distinction between the abstract and introduction.

We politely disagree. The abstract includes information about the results, which is clearly not included in the introduction. Nevertheless, wel attempted to improve the abstract for clarity.

Unclear how Figure 1 fits in the introduction. Seems more like an abstract

Figure 1 serves as an illustration for the motivation of our study. We propose to keep it, except the Editor considers it should be removed.

On the introduction, there is no clear distinction between how rising CO2 relates to plant water dynamics. I would suggest adding a sentence that links the paragraphs together

The reference to rising CO_2 was meant as the driver of global warming. While indeed elevated CO2 influences plant water dynamics, that is not the focus of this study (rather hydrometeorological conditions). We have revised the sentence to avoid distractions:

"Rising temperatures under global warming intensify the atmospheric demand for water [...]"

Please revise citations

- **Line 48- Foster has no year, is it a one author publication?**
- **Line 50- no year**
- **Line 71- no year**
- **Line 187- requires citation**

Apologies, we realised that the downloaded citations did not contain the full information. This is now corrected. We have added a reference.

Line 76-78: The hypothesis is conceptually interesting but fundamentally circular, since hysteresis itself reflects stress; the authors should clarify which specific aspect of hysteresis they expect to change prior to observable canopy-level decline.

The point in lines 76 and 78 is rather about sub-daily plant hydraulic dynamics providing earlier indicators of responses to stress than optical signals, commonly used in the remote-sensing community: *"hydraulic hysteresis signatures potentially provide early-warning indicators that may significantly precede optically visible signs of vegetation decline"*. Furthermore, this is a hypothesis proposed in the literature, including the references cited in the introduction (e.g. Hammond et al. (2021) or Konings et al. 2021).

We note that our study aims to precisely understand which facets of sub-daily SF allow us to detect responses to hydrometeorological conditions, and particularly extremes. We acknowledge that the lengthy description in the introduction might have made these points less evident, and we have restructured the paragraphs around the motivation and hypothesis of this study for clarity.

Begin by mentioning sap flow acronym at the beginning, not at the middle of intro, and being consistent with the acronyms. If not, there is no need to include it, especially SF for sap flow

Sap flow is introduced in the second paragraph of the introduction and the acronym is introduced in the first instance of the use of "sap flow". We find this appropriate since the first paragraph of the introduction rather introduces the background motivation to understand plant responses to increasing hydrometeorological conditions. We avoided introducing the acronym in the abstract for readability. The acronym is needed in the main text given that it facilitates referring to SF-VPD curves, for example. It is also a common acronym in the literature on the topic. It is not clear what exactly should be changed.

Why contrasting sites? What is the motivation? Why not concurrent mediterranean sites, or concurrent tropical sites? Is three different sites without replication informative enough?

We select three sites along a climate gradient, as shown in Figure A1, to illustrate the seasonal variability of the slope and area metrics and the differences in the diurnal cycles for extreme temperature and TSM conditions. However, we do not limit the analysis only to these sites for the synthesis figures, namely the correlation between the metrics with climate variables (Fig. 4) and the effect of coarser temporal sampling on the reliability of these metrics (Fig. 6).

Paragraph 80, yes and? How does it add to the flow?

The paragraph in line 80 refers to ecosystem monitoring networks. We are not sure exactly what the critique in this point was, but we have carefully revised the introduction to better emphasize the motivation and novelty of this study.

Figure 3- Align the mean seasonal gradient to the right panels, not the Middle. X axis?
We have revised Figure 3.

Would also recommend better ID for the different sites. They are unnecessarily complicated
We use the terminology from SAPFLUXNET.

Line 449- there is no justification to why Vegetation Optical Depth research can serve to motivate your work. It can be informative on its own. I would suggest restructuring towards how it can inform remote sensing initiatives.
Similar comment on line 496. If you want to use your work for VOD, then it is worth mentioning in the conclusion, but it does not add to your argument in any way
The link between VOD and SF is discussed in detail in Section 4.4 and the point is more broadly mentioned in the conclusions as “vegetation water content”. We removed the reference to VOD at the end of Section 4.3 since it was distracting, and have added a reference to Section 4.4 and revised the sentence in former Line 449.

There is no mention of how sap flow is measured (heat thermistor, etc-are the sensor different, and does not cite literature depicting that SF has a clear diurnal cycle- it is not always the case and is very much individual dependent.
The information is provided in the reference to SAPFLUXNET. We consider it to be sufficient for the purpose of this study.

Line 177: great depths?
The sentence was not clear, we mean that most water is stored as ground water, i.e., not accessible to plants. We revised the sentence to:

> *Overall, evapotranspiration exceeds precipitation in this region (Urban et al., 2019), with water stored mostly as ground water.*

Line 179: age of trees? Or age of the forest?
The trees, clarified now.

No division between major groups such as angio/gymno clades, which might inform why there is similar patterns across the different sites.
We provide results for different groups in Supplementary Table 1 now. However, it is not evident that this information explains similarities/differences across sites. We have, though, included more information about soil types and management, and climatic background (Figure A1). We find that the latter two help explaining difference across sites.

There is lack of consistency throughout the writing

We carefully revised the text for consistency, however concrete examples of breaks in the consistency would be appreciated.

What is the broader application of this work? It is not explicit.

The broader application of this work was mentioned both in the introduction, discussion and conclusions: our study proposes two metrics that allow to characterise tree responses to hydrometeorological conditions and extremes, and provides concrete information about the temporal resolution needed in sparse satellite measurements to correctly depict these metrics. We will carefully proof-read the text for clarity.