This manuscript investigates the joint effects of shortwave solar irradiance (SW) and vapor pressure deficit (VPD) on net ecosystem exchange (NEE) across 23 European ICOS eddy-covariance sites. I recognize the timeliness and importance of evaluating compound atmospheric drought alongside solar-brightening feedbacks. However, mathematical artifacts, ecophysiologically implausible response thresholds, flawed machine-learning validation, and unjustified cross-scale extrapolations fundamentally weaken the manuscript's central conclusions. In particular, the reported increase in VPD sensitivity under high radiation does not establish a novel physiological mechanism. Instead, it largely reflects well-documented emergent behavior in coupled photosynthesis–stomatal conductance formulations, the mathematical endogeneity inherent in the Penman–Monteith inversion, and canopy radiative-heating effects. The work requires substantial reanalysis and extensive structural revisions before it is suitable for publication.

Response: We sincerely thank the reviewer for this thorough and constructive assessment of our manuscript. We agree that the coupled responses of photosynthesis and stomatal conductance to radiation and atmospheric dryness are well established in ecophysiological theory. Our objective is not to propose a novel physiological mechanism, but rather to investigate how these interacting environmental drivers influence the emergence and variability of ecosystem-scale VPD response thresholds under contrasting radiation conditions. To address this concern, we will revise the Introduction and Discussion to clarify the study's objectives, better distinguish our findings from established physiological mechanisms, and acknowledge the methodological limitations and uncertainties associated with our analyses.

We acknowledge the reviewer's concerns regarding potential mathematical artifacts, model validation, and the interpretation of the observed response patterns. To address these concerns, we will conduct additional analyses using multiple complementary approaches, including leave-one-site-out cross-validation, generalized additive mixed models, and evaluations informed by established coupled photosynthesis-stomatal

conductance frameworks.

We have conducted preliminary analyses that suggest our main finding, namely a stronger apparent sensitivity of NEE to VPD under high-radiation conditions, remains qualitatively consistent across alternative analytical approaches. Nevertheless, we agree that the underlying mechanisms require more careful evaluation and interpretation. In the revised manuscript, we will further examine the extent to which the observed patterns may arise from known physiological interactions, canopy radiative-heating effects, or methodological constraints, and we will explicitly discuss these possibilities.

We will also substantially revise the manuscript to clarify the distinction between empirical observations and mechanistic interpretations, moderate any statements that could be construed as implying physiological novelty, and more explicitly acknowledge the limitations and uncertainties of the current analysis. We believe these additions and revisions will substantially strengthen the manuscript.

Major Scientific Concerns:

Methodological Artifacts in Machine Learning, Threshold Detection, and Path Modeling

The empirical thresholds (τ_1 = 3.90 hPa, τ_2 = 20.84 hPa) and the attribution models have serious analytical and ecophysiological limitations: The XGBoost framework introduces substantial temporal data leakage by randomly partitioning half-hourly records into 80% training and 20% validation subsets. Because eddy-covariance observations show pronounced diurnal periodicity and temporal autocorrelation, the elevated validation scores (R^2 = 0.72 for NEE, 0.78 for GPP_{DT}) likely reflect interpolation between chronologically adjacent records rather than genuine out-of-sample predictive skill. The authors should evaluate model generalization using block cross-validation (e.g., leave-one-year-out or spatial leave-one-site-out).

Response: We agree that randomly partitioning half-hourly observations into training and validation subsets may lead to overly optimistic estimates of predictive performance because of temporal autocorrelation and the presence of observations from the same sites in both subsets. We furthermore agree that additional validation strategies are needed to better assess model generalizability. To address this concern, we will implement leave-one-site-out (LOSO) cross-validation, whereby each site will be excluded in turn from model training and used exclusively for independent validation.

In addition, we will evaluate whether our main conclusions regarding the joint effects of SW and VPD remain consistent under this alternative validation framework. Preliminary analyses suggest that the identified response patterns and threshold behavior are qualitatively robust, although predictive performance metrics are expected to be lower under LOSO validation. We will report the LOSO validation results in the revised manuscript.

Fitting continuous piecewise regressions to TreeSHAP surrogate values instead of direct observations identifies the algorithmic splitting boundaries of gradient-boosted trees, not biological tipping points. In addition, the unrealistically narrow 95% bootstrap confidence interval reported for τ_1 (3.87–3.93 hPa; ±0.03 hPa) results from fitting regressions to pre-binned medians from hundreds of thousands of pooled observations. This approach collapses observational flux variance and ignores the spatio-temporal clustering of eddy-covariance data.

Response: We agree that TreeSHAP values represent model-derived feature attributions rather than direct observations of ecosystem responses. Consequently, thresholds inferred from SHAP dependence relationships should not be interpreted as physiological tipping points. To address these concerns, we will take the following three steps.

First, we will clarify the role of TreeSHAP in identifying potential response thresholds. Our original motivation for using TreeSHAP was to isolate the contribution of VPD

from those of other correlated environmental drivers, particularly air temperature and shortwave radiation. By decomposing the predictions of a multivariate XGBoost model, TreeSHAP provided a useful framework for examining how the modeled influence of VPD varies across environmental conditions while accounting for interactions among predictors. In the revised manuscript, we will clarify this distinction and frame the identified thresholds as empirical transition points in the modeled response surface rather than direct evidence of underlying physiological thresholds.

Second, we will conduct additional analyses using direct observations to verify the identified response patterns. To evaluate whether the identified patterns are supported independently of the machine-learning framework, we will conduct additional analyses using generalized additive mixed models fitted directly to the observed flux data. These models will explicitly account for the nonlinear and interactive effects of VPD and shortwave radiation while incorporating site-level variability. Preliminary analyses indicate that the first response threshold remains evident in the observation-based framework and occurs within a similar VPD range to that identified from the TreeSHAP analysis. We will systematically compare the response patterns and threshold estimates obtained from the different approaches and report the degree of agreement among them.

Third, we will improve the uncertainty estimates by accounting for site-level clustering. We also acknowledge the reviewer's concern that fitting piecewise regressions to pre-binned median SHAP values may lead to underestimated uncertainty because the binning procedure reduces observational variability and does not fully account for the spatial and temporal dependence structure of eddy-covariance measurements. To address this issue, we will implement a site-level cluster bootstrap procedure and report revised uncertainty estimates that explicitly account for between-site heterogeneity.

Ecophysiologically, an inflection threshold of 3.90 hPa (0.39 kPa) is an implausible trigger for stomatal down-regulation in temperate and Mediterranean forests. In these forests, stomata maintain an operational conductance plateau under low evaporative demand and typically down-regulate only when VPD exceeds 0.8–1.5 kPa.

Under near-saturated air (< 0.4 kPa), latent heat fluxes approach instrument detection limits, and relative measurement errors peak. Indeed, Sections 3.3 and 3.4 expose the fragility of this threshold: the authors contradict their own breakpoint by discarding all data below 5.0 hPa to exclude conditions "without atmospheric dryness stress."

Response: We thank the reviewer for raising this important concern regarding the interpretation of the identified thresholds. We recognize that the manuscript was not sufficiently clear that the identified thresholds are not intended as physiological thresholds, and we will revise all sections to address this. Specifically, we agree that a VPD value of approximately 3.9 hPa (0.39 kPa) is implausible for stomatal down-regulation for emperate and Mediterranean tree species. The thresholds identified in our study were not intended to represent a physiological threshold for stomatal closure; rather, they represent empirical transition points in the ecosystem-scale carbon flux response to atmospheric dryness derived from eddy-covariance observations. Such ecosystem-level thresholds may emerge from the combined effects of canopy physiology, environmental covariation, and ecosystem-scale flux integration, and therefore do not necessarily coincide with leaf-level stomatal response thresholds.

We also note that relatively low VPD thresholds have been reported in studies based on different observational datasets and approaches. For example, our previous work (Zhong et al., 2023) identified VPD thresholds of 3.50 and 3.96 hPa for changes in vegetation productivity sensitivity across northern ecosystems. Li et al. (2026), in which we participated, used tree-ring observations to report a VPD threshold of 0.27 kPa (2.7 hPa) for Picea abies—a species widely distributed across boreal, cool-temperate, and mountainous regions of Europe. Independent work by Kusbach et al. (2025), using field-based observations of Norway spruce forests in Central Europe, reported a nonlinear relationship between VPD and stem radial growth, with growth peaking at approximately 0.4 kPa (4.0 hPa). Although these studies quantify different ecosystem responses and should not be directly equated with stomatal regulation

thresholds, they suggest that nonlinear ecosystem responses to atmospheric dryness may emerge at VPD levels lower than those typically associated with stomatal closure.

We acknowledge that the VPD threshold may vary among vegetation types and environmental conditions. The original 5 hPa cutoff in Sections 3.3 and 3.4 was intended to minimize the influence of nonlinear responses under low-VPD conditions when assessing the sensitivity of ecosystem carbon exchange to atmospheric dryness. This choice was informed by differences among vegetation types, including an estimated transition at approximately 4.53 hPa for evergreen needleleaf forests (ENF). We agree that the rationale for this choice was insufficiently explained and that the selected threshold may influence the subsequent results. To address this issue, we will conduct additional sensitivity analyses using alternative VPD cutoffs, including thresholds near the identified transition point (e.g., 4 hPa) as well as the originally adopted 5 hPa criterion. We will evaluate the extent to which the radiation–VPD interaction patterns depend on this methodological choice and revise our interpretation accordingly.

References

Zhong, Z., He, B., Wang, Y.-P., Chen, H. W., Chen, D., Fu, Y. H., et al. (2023). Disentangling the effects of vapor pressure deficit on northern terrestrial vegetation productivity. Science Advances, 9(32), eadf3166. https://doi.org/10.1126/sciadv.adf3166

Li, T., Chen, H. W., Chen, D., Xiao, J., Guo, L., Liu, L., et al. (2026). Increasing risk of vapor pressure deficit exceeding critical thresholds for tree growth. Proceedings of the National Academy of Sciences, 123(37), e2530363123. https://doi.org/10.1073/pnas.2530363123

Kusbach, A., Jr., Krejza, J., Homolová, L., Fischer, M., Janoutová, R., & Horáček, P. (2025). Productivity of coniferous forests evaluated by remote sensing and field-based

models. Environmental Research Letters, 20(2), 024016. https://doi.org/10.1088/1748-9326/ada2af

Finally, the SW binning strategy used in the covariance-based structural equation modeling (CB-SEM) introduces severe discretization artifacts. Stratifying observations into narrow 50 W m^{-2} bands suppresses radiation variance, which mathematically drives the direct path coefficient (SW⇒NEE) toward zero (from -0.23 to -0.02). Because elevated SW bins coincide with midday conditions characterized by peak temperatures and evaporative demand, this stratification systematically confounds the observed increase in the VPD⇒NEE path coefficient (from -0.01 to 0.64).

Response: Our primary motivation for stratifying observations into 50 W m^{-2} SW intervals was to examine whether the association between VPD and ecosystem carbon exchange varied across different radiation environments while minimizing variation in incident radiation within each subset. We agree, however, that this approach substantially reduces within-bin SW variability and may influence the magnitude and interpretability of standardized path coefficients.

We also agree that higher SW conditions are inherently associated with changes in other environmental factors, including air temperature, evaporative demand, and diurnal timing. Consequently, the increase in the estimated VPD → NEE path coefficient across SW bins could partly reflect these covarying factors rather than an isolated effect of radiation itself. We will therefore revise the manuscript to avoid overinterpreting the SEM results as evidence of an independent physiological mechanism.

To evaluate whether the reported SW–VPD interaction is robust to the discretization imposed by the binning procedure, we will complement the original analysis with continuous interaction models that treat SW and VPD as continuous predictors. These models will explicitly represent nonlinear responses and interactions while accounting for potential confounding effects associated with air temperature, soil moisture, site-

level variability, and diurnal patterns.

Preliminary analyses indicate that the positive SW–VPD interaction remains evident when radiation is modeled continuously rather than through discrete bins. Specifically, the estimated response of NEE to increasing VPD is stronger under high-radiation conditions than under low-radiation conditions, consistent with the pattern inferred from the binned analysis. In addition, a preliminary sensitivity analysis restricted to midday observations (10:00–16:00 local time) yields qualitatively similar interaction estimates, suggesting that the reported interaction is not solely attributable to comparisons across the full diurnal cycle.

Nevertheless, we acknowledge that residual confounding cannot be entirely excluded in an observational study. To further address this concern, we have conducted preliminary complementary SEM analyses using hourly SW anomalies that account for typical diurnal and seasonal variations, as well as daily mean SW anomalies with the seasonal cycle removed. Both analyses show broadly consistent patterns with our original findings, with generally stronger positive VPD–NEE path coefficients under positive SW anomalies than under negative SW anomalies. These preliminary results suggest that the observed SW–VPD interaction is not solely attributable to the typical diurnal or seasonal variability in radiation. In the revised manuscript, we will present these complementary analyses, systematically evaluate the sensitivity of the results to the radiation-binning approach, and place greater emphasis on the observational nature of the identified relationships. We will also clarify that the reported interaction represents a statistical pattern supported across multiple analytical frameworks rather than definitive evidence for a previously unrecognized physiological mechanism.

Mathematical Endogeneity and Conductance Scaling

The reported sensitivity of ecosystem conductance (Gs) to VPD under high irradiance is largely a mathematical and scaling artifact:

First, inverting the Penman–Monteith equation algebraically embeds VPD in the denominator ($G_s \propto LE/VPD$). Because latent heat flux (LE) is substantially higher under high solar irradiance, plotting G_s against VPD mechanically dictates a steeper hyperbolic decline ($y \propto 1/x$). The authors further hardcode this algebraic dependency by fitting the empirical function $G_s = a / (1 + b \cdot VPD)$. Asserting that "this pattern is also consistent with theoretical expectations from the Penman–Monteith framework (Eq. 2), which predicts stronger stomatal regulation" (Section 3.4, L. 366–368) conflates a physical transport inversion with an active ecophysiological feedback.

Response: We thank the reviewer for raising this important concern regarding the potential mathematical dependence of G_s on VPD. We agree that VPD explicitly enters the denominator of the inverted PM equation and that this algebraic dependence may contribute to the apparent negative G_s–VPD relationship. We would like to clarify that our primary aim in analyzing G_s was to provide a complementary assessment of ecosystem-scale water vapor exchange in relation to the observed SW–VPD interactions in NEE, rather than to establish independent evidence of physiological stomatal regulation. We recognize that this distinction was not sufficiently clear in the original manuscript and will revise Section 3.4 accordingly.

While we acknowledge the reviewer's concern regarding the mathematical dependence of PM-derived Gs on VPD, we would like to clarify that elevated SW does not necessarily lead to a stronger inversion-induced Gs–VPD sensitivity, because SW-related changes in both LE and available energy need to be considered. Specifically, the inverted PM equation (Eq. 2) can be expressed as:

$$G_s = \gamma \cdot ga \cdot LE / [\Delta \cdot A + B \cdot VPD - (\Delta + \gamma)\, LE],$$

where $A = R_n - G$ represents available energy (A), and $B = \rho C_p g_a$.

When LE, A, and the other parameters are held constant, the partial derivative of Gs with respect to VPD is indeed negative. However, under otherwise fixed conditions, an

increase in SW can affect both LE and A, which exert opposing effects on the magnitude of this partial derivative. Moreover, both LE and A may covary with VPD under actual environmental conditions. Consequently, the approximation $Gs \propto LE/VPD$ does not fully represent the inverted PM equation, and the net mathematical sensitivity under elevated SW cannot be determined from changes in LE alone. Importantly, this mathematical distinction does not eliminate potential endogeneity in PM-derived G_s.

To further evaluate this concern, we will supplement our analysis with diagnostic tests. Preliminary calculations indicate that, although median LE increases from approximately 58 to 141 W m^{-2} between low- and high-SW conditions, the median magnitude of the fixed-input logarithmic VPD sensitivity of PM-derived G_s decreases from approximately 0.99 to 0.74. These conditional elasticities represent the mathematical dependence of PM inversion rather than observed physiological responses. The results indicate that higher LE under elevated SW is not necessarily accompanied by greater relative mathematical VPD sensitivity, although differences in other meteorological conditions between SW groups may contribute to this contrast.

We also acknowledge that the originally applied hyperbolic function imposes a declining response shape. We will therefore replace the hyperbolic function with more flexible spline models as the primary approach for describing the Gs–VPD relationship, while retaining the original function as a supplementary comparison. Preliminary leave-one-year-out comparisons across 20 site–SW groups indicate that the constrained hyperbolic and flexible spline models each achieve lower prediction errors in 10 groups. Although this comparison does not eliminate the mathematical dependence inherent in PM-derived G_s, it indicates that the imposed hyperbolic function is not uniquely required to describe the observed relationship.

Accordingly, we will revise Section 3.4 to avoid interpreting the PM framework itself as direct evidence of stronger physiological stomatal regulation. We will distinguish the empirical Gs–VPD association from its possible physiological interpretation and explicitly discuss the limitations associated with PM inversion and uncertainties in

surface energy fluxes.

Second, analyzing absolute departures ($\Delta G_s = G_s - \bar{G}_{s,bin}$) rather than proportional adjustments conflates intrinsic sensitivity with baseline conductance. Because high irradiance stimulates baseline stomatal opening ($g_{s,ref}$) to facilitate carbon assimilation, the absolute decline in conductance per unit VPD is naturally magnified as established by classic stomatal scaling theory (Oren et al., 1999: $-dg_s/d \ln D \approx m \cdot g_{s,ref}$) and reaffirmed in recent syntheses (e.g., Grossiord et al., 2020; Novick et al., 2016). Once normalized to relative sensitivity ($-\partial \ln G_s/\partial \ln VPD$), stomatal responsiveness often proves largely invariant across light regimes.

Response:

We agree that larger baseline ecosystem conductance under high-radiation conditions can naturally lead to larger absolute declines in conductance with increasing VPD. As noted by the reviewer, this scaling behavior is consistent with established stomatal theory, whereby the magnitude of stomatal response often scales with reference conductance. To address this concern, we will supplement the original analysis with models using the dimensionless relative VPD sensitivity, defined as:

$$S = -\partial \ln (G_s) / \partial \ln (VPD)$$

This formulation evaluates proportional rather than absolute conductance responses and is therefore less influenced by differences in baseline conductance between radiation environments. We will estimate the difference in relative sensitivity between high- and low-SW conditions ($\Delta S = S_{high} - S_{low}$), while accounting for air temperature, soil moisture, seasonal variability, and diurnal patterns. Our preliminary analysis indicates a site-mean difference of 0.094 ± 0.058 (SE). These preliminary results suggest that the tendency toward stronger VPD sensitivity under elevated SW does not disappear completely when conductance changes are expressed on a relative scale.

In the revised manuscript, we will explicitly distinguish between absolute and relative

sensitivity metrics and avoid attributing differences in absolute G_s responses solely to enhanced stomatal regulation. We will further discuss our findings within the context of established conductance-scaling theory.

Biophysical Disconnects: Microclimatic Heating, Respiration, and Gas Exchange

Although the manuscript mentions leaf-to-air temperature differences in the Introduction, its analytical framework treats the forest canopy as an isothermal box, conflating leaf microclimate and respiratory kinetics with stomatal regulation:

Stomata respond directly to the VPD at the leaf surface within the canopy boundary layer ($D_{leaf} = e_s(T_{leaf}) - e_a$), not to ambient atmospheric dryness measured tens of meters above the canopy (D_{air}). Under high irradiance (SW > 500 W m^{-2}), sensible heat loading raises leaf temperatures several degrees above ambient air ($T_{leaf} > T_{air}$). Because saturation vapor pressure (e_s) increases exponentially with temperature via the Clausius–Clapeyron relationship, sunlit canopies experience a substantially greater true evaporative demand than tower instruments record. Without quantifying canopy decoupling (Ω) or estimating D_{leaf}, the reported "amplification" cannot be isolated as stomatal regulation rather than microclimatic thermodynamic heating.

Response: We thank the reviewer for highlighting the important role of canopy microclimate in determining the evaporative conditions experienced by leaves. We agree that atmospheric VPD measured above the canopy does not necessarily represent the actual vapor pressure gradient at the leaf surface, particularly under high-radiation conditions. As the reviewer notes, elevated shortwave radiation can increase leaf temperature relative to ambient air temperature, which may enhance leaf-to-air VPD through the nonlinear temperature dependence of saturation vapor pressure. We agree that radiation is a major driver of leaf temperature and leaf-level evaporative demand. However, the magnitude of leaf warming relative to ambient air temperature is also influenced by multiple interacting factors, including evaporative cooling, aerodynamic coupling, canopy structure, and water availability (Leuzinger and Körner, 2007; Farella

et al., 2022; Lian et al., 2026). Consequently, although elevated radiation can substantially increase leaf-to-air VPD, the magnitude of this increase may vary across sites and environmental conditions.

We acknowledge that radiation-induced changes in leaf temperature and leaf-to-air VPD could partly explain the stronger apparent sensitivity of ecosystem carbon exchange to atmospheric VPD under elevated SW. Our primary finding is an empirical radiation-dependent VPD response in ecosystem carbon exchange derived directly from observed NEE patterns. This finding does not rely on the Penman-Monteith inversion used to estimate ecosystem conductance, although the conductance analysis was intended to aid interpretation of the observed ecosystem-scale response. The reviewer's proposed mechanism therefore represents an important alternative explanation for the observed amplification effect rather than evidence against the existence of the empirical pattern itself.

In the revised manuscript, we will substantially expand the discussion of potential microclimatic controls, including leaf-to-air temperature differences, canopy-atmosphere coupling, and thermodynamic effects associated with radiative heating. We will also revise our mechanistic interpretation to more clearly distinguish ecosystem-scale observations from the physiological and microclimatic processes that may contribute to them. In particular, we will clarify that the observed radiation-dependent VPD sensitivity of NEE cannot be attributed exclusively to enhanced stomatal regulation and may partly reflect radiation-induced changes in leaf-level evaporative demand.

References

Leuzinger, S., & Körner, C. (2007). Tree species diversity affects canopy leaf temperatures in a mature temperate forest. Agricultural and Forest Meteorology, 146(1–2), 29–37. https://doi.org/10.1016/j.agrformet.2007.05.007

Farella, M. M., Fisher, J. B., Jiao, W., Key, K. B., & Barnes, M. L. (2022). Thermal remote sensing for plant ecology from leaf to globe. Journal of Ecology, 110, 1996–2014. https://doi.org/10.1111/1365-2745.13957

Lian, X., JiJi, J., Fang, J., et al. (2026). Leaf temperature and its departure from ambient air temperature. Nature Plants, 12, 1189–1202. https://doi.org/10.1038/s41477-026-02304-w

Although the authors examine partitioned GPP in Section 3.1 (Fig. 4), standard flux-partitioning algorithms (e.g., Reichstein et al. 2005; Lasslop et al. 2010) rely on prescribed temperature-respiration equations (Q_{10} temperature coefficient and Lloyd–Taylor function) and arbitrary VPD thresholds (e.g., 10 hPa). Thus, training machine-learning models on flux-partitioned GPP risks teaching them the partitioning algorithms' assumptions rather than unconstrained biological relationships.

Response: We agree with the reviewer that flux-partitioned GPP estimates are inherently influenced by the assumptions embedded in partitioning algorithms, particularly those related to temperature-dependent respiration and prescribed VPD responses. As discussed in the original manuscript (Discussion, Lines 428–442), these assumptions may introduce biases into the estimated relationships between GPP and environmental drivers. For this reason, we used directly measured NEE, rather than partitioned GPP, as the primary basis for our main analyses and threshold identification.

The analyses based on GPP derived from both daytime and nighttime partitioning methods (Fig. 4), particularly GPP_{NT}, were included solely as supplementary assessments of robustness and consistency. Their purpose was to examine whether the general response patterns identified from NEE could also be observed in partitioned estimates of ecosystem productivity. Importantly, the two partitioning methods differ in their underlying assumptions. The nighttime method estimates respiration from nighttime observations and extrapolates it to daytime conditions, whereas the daytime

method uses a light-response function that explicitly incorporates a prescribed VPD response. Therefore, comparing results from both methods can partly assess the sensitivity of the observed patterns to the choice of partitioning algorithm. Nevertheless, we acknowledge that these GPP-based analyses are not fully independent of the assumptions underlying flux partitioning and therefore should not be interpreted as independent validation of the identified response thresholds or their physiological interpretation.

In the revised manuscript, we will further clarify the intended role of these analyses and explicitly discuss the limitations associated with the use of flux-partitioned GPP products. We will also reduce the emphasis placed on GPP-derived results and present them as supporting evidence, while maintaining the NEE-based analyses as the primary basis for our conclusions.

In the CB-SEM (Fig. 6 and Section 4.1), the direct path coefficient for TA⇒NEE changes from +0.18 at low SW to −0.44 at high SW. The authors interpret this as high temperatures stimulating net carbon uptake (negative NEE) more than respiration under strong light. However, this claim directly contradicts basic plant thermal optics, midday enzymatic downregulation, and the exponential acceleration of respiration (Reco = Ra + Rh). This counterintuitive shift strongly suggests severe collinearity suppression among SW, TA, and VPD in the SEM, rather than a genuine physiological mechanism.

Response: We agree that the negative TA→NEE coefficient under high SW should not be interpreted as direct evidence that elevated temperatures stimulate photosynthesis more strongly than ecosystem respiration. We acknowledge that elevated temperatures can suppress photosynthesis, particularly above the photosynthetic temperature optimum, while also increasing ecosystem respiration, thereby reducing net carbon uptake. However, the temperature response of ecosystem carbon exchange is nonlinear, and moderate increases in temperature may be associated with enhanced carbon uptake when temperatures remain below the photosynthetic optimum and sufficient radiation

is available (Niu et al., 2012; Crous et al., 2022). Nevertheless, the negative TA→NEE coefficient observed in our SEM does not establish such a physiological mechanism. Instead, it represents a conditional statistical association that may also reflect nonlinear temperature responses, environmental covariation, and differences in the environmental conditions represented by the SW intervals.

To address the concern regarding potential multicollinearity, we conducted preliminary variance inflation factor (VIF) diagnostics. The results indicate that all VIF values in the original SW-stratified models are below 5. Even within the highest SW interval (500–550 W m^{-2}), the VIF values for TA and VPD are approximately 2.57 and 2.51, suggesting that severe linear multicollinearity is unlikely. However, we acknowledge that relatively low VIF values alone cannot exclude statistical suppression or coefficient instability arising from correlated environmental predictors and model specification.

In the revised manuscript, we will explicitly report the VIF diagnostics and further evaluate the robustness of the TA→NEE relationship using alternative model specifications, including continuous SW–VPD interaction models, nonlinear representations of temperature responses, and site-level bootstrap analyses. We will also revise the interpretation presented in Sections 3.3 and 4.1 to avoid attributing the negative TA→NEE coefficient directly to enhanced photosynthetic activity under high-temperature conditions. Instead, we will interpret this coefficient as a conditional association whose underlying physiological mechanisms cannot be determined from the SEM alone.

References

Niu, S., Luo, Y., Fei, S., et al. (2012). Thermal optimality of net ecosystem exchange of carbon dioxide and underlying mechanisms. New Phytologist, 194, 775–783. https://doi.org/10.1111/j.1469-8137.2012.04095.x

Crous, K. Y., et al. (2022). Temperature responses of photosynthesis and respiration in

evergreen trees from boreal to tropical latitudes. New Phytologist. https://doi.org/10.1111/nph.17951

Internal Inconsistency and Unsubstantiated ESM Deficiencies

In the Conclusion (Section 5), the authors caution that if Earth system models (ESMs) fail to capture the amplifying effect of radiation on atmospheric dryness, they risk underestimating climate impacts on carbon uptake. However, this speculation directly contradicts the authors' discussion in Section 4.1, which notes that this interaction is consistent with standard leaf-level formulations, including the Ball–Berry (Ball et al., 1987) and Medlyn (Medlyn et al., 2011) models.

Modern land-surface schemes (e.g., CLM, ORCHIDEE, JULES, and JSBACH) already couple stomatal conductance directly to radiation-driven assimilation through these formulations. Because marginal stomatal sensitivity to vapor pressure deficit ($\partial g_s/\partial D$) naturally scales with assimilation, a greater absolute reduction in conductance under high shortwave radiation emerges from the existing model physics. The authors cannot invoke Ball–Berry/Medlyn schemes to explain their empirical mechanism in Section 4.1 while also hypothesizing an ESM deficiency in Section 5 without benchmarking their observations against actual land-surface model outputs. The speculative claims in Section 5 about model failures and BECCS vulnerabilities should either be supported by explicit model evaluations or removed.

Response: We agree with the reviewer and acknowledge that our original statements regarding potential deficiencies in Earth system models were misleading, particularly in the absence of direct model evaluations. Accordingly, we will remove the speculative statements regarding potential ESM deficiencies and BECCS vulnerabilities from the Conclusion and ensure that the implications of our findings remain within the scope supported by our analyses.

In addition to functional differences, the site network is strongly dominated by

evergreen needleleaf forests (13 of 23 sites), with nearly half located in high-latitude boreal Fennoscandia. This composition may skew aggregate European patterns toward cold, energy-limited systems with narrow hydraulic safety margins, limiting how well the findings generalize to temperate deciduous forests and water-limited Mediterranean forest regimes.

Response: We thank the reviewer for highlighting this important limitation. We agree that the predominance of evergreen needleleaf forests and high-latitude boreal sites may influence the aggregated patterns and limit the extent to which our findings can be generalized to other European forest types and climatic regions. The site selection was primarily constrained by the availability of sites with sufficiently complete observational records and the required environmental variables.

In the revised manuscript, we will explicitly acknowledge this limitation and discuss its implications for the interpretation and generalization of our results. In particular, we will emphasize that the applicability of our findings to temperate deciduous forests and water-limited Mediterranean ecosystems remains uncertain and warrants further investigation using a more balanced representation of forest types and climatic conditions.

Finally, relying solely on uncalibrated surface-layer soil water content overlooks deep root-zone reserves, water stored in fractured bedrock, and access to capillary groundwater, all of which are common in mature European stands. When trees can access subsoil water, high atmospheric VPD can sustain elevated transpiration without triggering stomatal closure, provided whole-plant hydraulic transport capacity (K_{plant}) is not exceeded. Attributing the observed flux dynamics to atmospheric dryness without accounting for root-zone moisture buffering therefore leaves the isolation of atmospheric drought incomplete.

Response: We agree that surface-layer soil water content may not fully represent the water available to mature trees, particularly in ecosystems where deep rooting, fractured

bedrock storage, capillary groundwater access, or other subsurface water sources contribute substantially to plant water supply.

In this study, surface-layer SWC was used because deeper soil moisture observations were not consistently available across the majority of flux tower sites included in our analysis. We acknowledge that this limitation restricts our ability to fully characterize root-zone water availability and its potential buffering effects on plant responses to atmospheric dryness.

We also agree that access to deep water sources may partially decouple transpiration and stomatal behavior from short-term variations in atmospheric VPD. Under such conditions, elevated VPD does not necessarily result in immediate stomatal closure if sufficient water remains available to sustain plant hydraulic functioning. Consequently, observed ecosystem responses to atmospheric dryness are likely influenced by both atmospheric demand and the degree of subsurface water accessibility.

At the same time, we note that our analysis focuses on ecosystem-scale carbon exchange responses to observed atmospheric conditions rather than on isolating purely atmospheric drought effects. Therefore, the absence of comprehensive root-zone moisture information does not invalidate the observed SW–VPD response patterns, but it does introduce uncertainty regarding the mechanisms through which these patterns arise and the extent to which they are moderated by plant water availability.

In the revised manuscript, we will explicitly discuss the limitations associated with using surface-layer SWC as a proxy for ecosystem water status and clarify that root-zone moisture buffering, groundwater access, and plant hydraulic regulation represent important unaccounted sources of variability. We will also temper our interpretation regarding atmospheric drought effects and acknowledge that a more complete separation of atmospheric and soil-water constraints would require deeper soil moisture observations and/or direct indicators of plant water status.

Unjustified Narrative Extrapolation and Acclimation Gaps

In Section 4.3 and the Conclusion, the authors extrapolate sub-daily (30-minute) flux sensitivities across a 5-year window (2019–2024) to multi-decadal climate projections (mid-21st century brightening) and the feasibility of continental BECCS. While Section 4.3 briefly acknowledges temporal scale dependency, the manuscript's broader implications directly contradict this caveat.

Most notably, the authors justify using the 2019–2024 period in Section 2.1 to "minimize potential confounding effects from long-term physiological adjustments of photosynthesis and respiration, such as thermal acclimation". It is logically inconsistent to intentionally isolate instantaneous responses by excluding acclimation and then use those same responses to infer multi-decadal carbon-sink vulnerability, where structural and physiological acclimation, such as shifts in Huber values, redistribution of rooting depth, and thermal acclimation of $V_{c\,max25}$ and J_{max25}, dominate ecosystem trajectories. The speculative statements regarding mid-century forest sinks and BECCS in Sections 4.3 and 5 should either be removed or heavily qualified.

Response: We agree that the short-term flux sensitivities identified in our study cannot be directly extrapolated to multi-decadal changes in forest carbon sequestration, particularly given the potential influence of both physiological and structural acclimation processes.

Nevertheless, we believe that the observed SW–VPD interaction remains relevant for understanding potential ecosystem responses under future climate conditions, especially if high SW and high VPD become more frequent or increasingly co-occur. We acknowledge, however, that the long-term consequences of this interaction remain uncertain and may be substantially modified by acclimation and other ecosystem adjustments.

To address the reviewer's concern, we will explicitly acknowledge the limitations of extrapolating short-term responses to longer timescales in Section 4.3 and will substantially moderate the broader implications presented in the Conclusion. Specifically, we will remove the speculative statements regarding Earth system model deficiencies and BECCS, while retaining a carefully qualified discussion of the potential relevance of SW–VPD interactions under future climate change scenarios.

Specific/minor comments:

Introduction

In Paragraph 3, please clarify whether the cited study by Kašpar et al. (2024) addresses stem radial increment or growth, or instantaneous carbon exchange. This distinction between wood formation kinetics and half-hourly canopy gas exchange is essential for physiological context.

Response: We agree that stem radial growth and instantaneous ecosystem carbon exchange represent distinct processes operating on different temporal scales. Kašpar et al. (2024) examined daily and seasonal stem radial growth using dendrometer observations rather than instantaneous carbon fluxes. We will clarify this distinction in the revised Introduction to prevent potential misunderstanding and ensure that the scope and relevance of the comparison are accurately presented.

In Paragraph 4, it would be better to link the identified knowledge gaps directly to the study's explicit objectives and specific research questions.

Response: We thank the reviewer for this helpful suggestion. We will revise the final paragraph of the Introduction to more explicitly connect the identified knowledge gaps with our research objectives and hypotheses, thereby providing a clearer rationale for the study.

In Section 2.1, pairing 1-km, 4-day MODIS LAI products with half-hourly tower

fluxes creates a pronounced spatial and temporal mismatch with the dynamic eddy-covariance footprint. The authors should clarify how they addressed this discrepancy or assess its influence on the observed relationships.

Response: We thank the reviewer for highlighting the potential spatial and temporal mismatches between satellite-derived LAI and eddy-covariance measurements. In this study, LAI values were extracted at the flux tower locations and matched to half-hourly observations using a nearest-neighbor temporal approach, assuming that LAI changes relatively slowly over time. We acknowledge that this approach cannot fully resolve the spatial and temporal mismatches between satellite-derived LAI and the dynamic eddy-covariance footprint, which may vary with measurement height and meteorological conditions.

To further assess the influence of spatial representativeness, we will conduct a sensitivity analysis using LAI averaged over a 3 × 3 pixel window surrounding each flux tower, in comparison with the original tower-location extraction. We will evaluate whether this alternative spatial sampling approach substantially influences the observed relationships. We will also clarify the LAI extraction and temporal matching procedures in Section 2.1 and explicitly acknowledge the remaining limitations in the revised manuscript.

In Section 2.2, the XGBoost architecture lacks sufficient hyperparameter optimization. Fixing the model at 100 boosting rounds without reporting cross-validated tuning of key parameters, particularly maximum tree depth, learning rate, and regularization (λ, α), creates a substantial risk of depth-dependent overfitting.

Response: We agree that assessing hyperparameter sensitivity is important to reduce the risk of overfitting and to ensure the robustness of the model results. Accordingly, we conducted five-fold site-grouped cross-validation to evaluate the effects of key hyperparameters, including tree depth, learning rate, L1/L2 regularization, and the number of boosting rounds. The results favored a shallower and more strongly

regularized model, while larger tree depths led to greater discrepancies between training and validation performance, indicating an increased risk of overfitting. In the revised manuscript, we will further examine whether the main VPD-SHAP response patterns and the identified threshold values remain consistent across alternative hyperparameter configurations and report the corresponding results to demonstrate the robustness of our findings.

The exclusion of data with VPD ≤ 5.0 hPa (Section 3.3, L. 339–341) is introduced abruptly in the Results rather than defined in the path analysis methodology in Section 2.4. Moreover, discarding observations below 5.0 hPa does not adequately resolve the structural conflict between the demonstrated piecewise nonlinearities and the linear CB-SEM assumptions. Please formally document this screening rule in Section 2.4 and assess whether linear path analysis is robust under non-normal, collinear midday conditions.

Response: We agree that the VPD > 5 hPa screening criterion should have been formally described in the Methods section rather than being introduced in the Results. We will therefore incorporate this screening rule and its rationale into Section 2.4 of the revised manuscript.

We will clarify the rationale for this screening criterion in Section 2.4 and conduct additional robustness analyses to assess the sensitivity of the SEM results to the VPD threshold, predictor collinearity, and distributional assumptions. We will also complement the original SEM with a model including a continuous SW–VPD interaction term, allowing nonlinear responses while accounting for relevant environmental covariates.

In Section 2.5, the inversion of the Penman–Monteith equation assumes complete surface energy balance closure (Rn – G = H + LE). Because eddy-covariance measurements routinely show a 10% to 30% closure gap, inverted canopy conductances may be systematically biased unless an energy balance adjustment is

applied or an uncertainty bound is included.

Response: Thank you for raising this important point. We agree that energy balance non-closure in eddy-covariance measurements may introduce uncertainty into ecosystem conductance estimates derived from the inverted Penman-Monteith equation. To assess the potential influence of this uncertainty, we will conduct an additional sensitivity analysis comparing G_s estimates derived from the original latent heat flux with those obtained after an energy balance closure correction that preserves the Bowen ratio, following established approaches (Wohlfahrt et al., 2009; Pastorello et al., 2020). Where the necessary energy flux measurements are available, we will evaluate whether the G_s–VPD relationships and their sensitivity to SW remain consistent after this correction.

We acknowledge that energy balance corrections themselves rely on assumptions regarding the allocation of the energy imbalance and therefore cannot completely eliminate uncertainty. In the revised manuscript, we will clarify this potential source of uncertainty and its implications for the interpretation of the G_s results, while emphasizing that our primary NEE-based analyses do not depend on PM-derived conductance estimates.

References

Wohlfahrt, G., Haslwanter, A., Hörtnagl, L., Jasoni, R. L., Fenstermaker, L. F., Arnone, J. A., & Hammerle, A. (2009). On the consequences of the energy imbalance for calculating surface conductance to water vapour. Agricultural and Forest Meteorology, 149, 1556–1559. https://doi.org/10.1016/j.agrformet.2009.03.015

Pastorello, G., Trotta, C., Canfora, E., et al. (2020). The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. Scientific Data, 7, 225. https://doi.org/10.1038/s41597-020-0534-3

Throughout Section 3, the text often describes VPD effects as "shifting from negative

to positive" or "strengthening" based solely on SHAP coordinate values. Because SHAP metrics quantify additive model deviations relative to a base value ($E[f(x)]$) conditioned on training sample distributions, please revise the phrasing to clearly distinguish statistical feature contributions from physical flux dynamics.

Response: We will revise the wording throughout Section 3 to distinguish statistical feature contributions from physical flux responses. Specifically, expressions such as "shifting from negative to positive" and "strengthening" will be clarified as changes in SHAP contributions to predicted NEE or GPP, rather than interpreted as direct evidence of changes in physiological responses. We will also clarify this distinction in the interpretation of the corresponding figures.

The reporting of GPP thresholds is ambiguous and internally contradictory. Specifically, the manuscript states that "the first threshold from the nighttime partitioning method has corresponding τ_2 values of 20.38 hPa and 18.76 hPa." It is unclear whether these values correspond to τ_1, τ_2, or distinct thresholds derived from the daytime and nighttime partitioning methods. Please clarify the parameter definitions and ensure consistent notation across both partitioning schemes.

Response: We thank the reviewer for pointing out this ambiguity. We agree that the distinction between the first (τ_1) and second (τ_2) thresholds in the daytime and nighttime GPP partitioning methods was not sufficiently clear. In the revised manuscript, we will revise Section 3.2 to explicitly define and distinguish these two thresholds and ensure that the notation is used consistently throughout the manuscript and the Supplementary Materials.

In Section 3.3 and Figure 6, the abrupt reversal of the TA⇒NEE path coefficient from +0.18 to −0.44 should be explicitly tested for collinearity suppression rather than interpreted as biological stimulation of carbon uptake.

Response: We will assess predictor collinearity and examine the stability of the TA–

NEE path coefficient through additional sensitivity analyses. We will also revise the corresponding interpretation in Sections 3.3 and 4.1 to avoid attributing this reversal directly to enhanced carbon uptake without sufficient supporting evidence.

Discussion

- In Section 4.3, although the discussion evaluates atmospheric dryness, it does not adequately address multi-year soil moisture drought legacies. Please explain how cumulative depletion in the root zone and subsoil, documented during recurrent European compound drought events, may compound or decouple the observed sensitivities to solar radiation and atmospheric evaporative demand.

Response: We thank the reviewer for this important suggestion. We agree that antecedent hydrological conditions may modulate ecosystem responses to atmospheric water demand. Persistent root-zone water depletion can exacerbate VPD-induced constraints on carbon uptake, whereas access to deeper soil water reserves may buffer vegetation against short-term surface moisture deficits. Because our analysis primarily relies on surface-layer soil moisture observations, these longer-term and deeper-profile hydrological influences cannot be fully captured. In the revised manuscript, we will expand Section 4.3 to discuss these mechanisms as an important source of uncertainty and to clarify how the absence of root-zone soil moisture information may affect the interpretation of the observed SW–VPD–NEE relationships.

Abstract and Conclusions

- The Abstract and Conclusion might need to be revised based on the Major Scientific Concerns.

Response: We agree and will revise the Abstract and Conclusion accordingly to reflect the responses to the major scientific concerns and ensure consistency with the revised analyses.