Reviewer 2

Rogozin et al.'s manuscript presents the nine-year dynamics of WUE, CUE, and LUE in a hemiboreal forest in Estonia, and determines which climatic variables drive them at daily and interannual scales. They found VPD to be the main driver of daily WUE and LUE, PAR to be the main driver of daily CUE, and temperature to be important for all three. SPEI analysis shows complex effects of drought and wetness on the three RUEs. Overall, this paper provides valuable insights on the dynamics of an understudied ecosystem type and is clearly written. Some of the methods choices should be reconsidered though, and more extensive analysis is needed to have clear takeaways from the SPEI analysis – the authors make claims that should be supported by further data and analysis from their site. The relevance of hemiboreal forests and the previous research that has been done on RUEs should also be expanded upon.

We thank the reviewer for the careful and constructive assessment of our manuscript and for recognizing the value of the nine-year eddy-covariance dataset from an understudied hemiboreal forest ecosystem. We appreciate the reviewer's comments and will address them in detail below in the Specific Comments section.

In the revised manuscript, we will improve the clarity of the methodological choices, especially the definition of growing season, the use of ecosystem-scale WUE, and the interpretation of RF and GAM results. We will also revise the SPEI section to make the interpretation more transparent and better supported by the underlying data. In particular, we will provide additional distributional information for SPEI classes and moderate the interpretation where the current wording is too strong.

We also agree that the Introduction should more clearly explain the relevance of hemiboreal forests, the uniqueness of the study site, and how previous RUE studies frame the knowledge gap addressed by this work. These points are addressed in more detail below.

Specific Comments

Clearer relevance of their research. The introduction is superficial and doesn't give the reader sufficient reason for why the study system is unique, nor explanation of what has been researched to this point and what those studies found.

We thank the reviewer for this comment and agree that the Introduction should state the relevance and novelty of the study more clearly. In the revised manuscript, we will expand the

Introduction to better explain why old hemiboreal coniferous forests are important transition ecosystems under climate change and how our site differs from more commonly studied boreal coniferous EC sites. We will also strengthen the literature context by briefly summarizing previous RUE studies in boreal, temperate, and hemiboreal forests and by clarifying the specific knowledge gap addressed by our nine-year dataset.

L51-53: why does occupying the boreal-temperate ecotone make hemiboreal forests especially sensitive to climate variability? More explanation would be good here – why do these forests respond more strongly to the same amount of climatic variability than another forest type would? Or do you mean that the boreal-temperate ecotone is experiencing more variability?

Thank you for pointing this out. We agree that the manuscript should more clearly distinguish between the broader hemiboreal forest zone and the specific stand investigated in this study. Hemiboreal forests are often described at the regional scale as transitional forest landscapes containing both boreal coniferous and temperate broadleaved elements. However, individual stands within this zone are dominated by coniferous species, especially on dry upland podzolic soils. Our EC footprint represents such an old upland coniferous stand dominated by Scots pine and Norway spruce.

Therefore, we do not intend to present this site as representative of all hemiboreal forest types. Rather, it represents a typical old coniferous forest stand located within the hemiboreal boreal–temperate ecotone. This distinction is important because the site combines boreal-type species composition with a transitional hemiboreal climate regime. We will revise the introduction and site description to make this clearer and avoid implying that the studied stand itself has a mixed coniferous–broadleaved canopy.

L65-77: In this paragraph, please make it clearer where studies in the hemiboreal zone have been conducted: it sounds like only in southern Siberia? Maybe start with this, and mention what they found? And then compare and contrast with what was found in regions close to the hemiboreal zone elsewhere?

Thank you for this constructive suggestion. We agree that the geographical scope of previous RUE studies should be presented more clearly. In the submitted manuscript, studies conducted directly within the hemiboreal zone and studies from adjacent boreal or temperate regions were listed together, which may have made the research gap less clear.

In the revised manuscript, we will reorganize this paragraph to distinguish between: (i) studies conducted within hemiboreal or hemiboreal-like transition zones, such as southern Siberia; (ii) studies from adjacent boreal and temperate coniferous forests in North America and Northern Europe; and (iii) the specific gap addressed by our study, namely the lack of long-term, multi-RUEs eddy-covariance analyses from old coniferous stands in the Nordic-Baltic hemiboreal zone. This revision will clarify that previous work provides valuable context, but that long-term integrated analyses of WUE, CUE, and LUE remain scarce for hemiboreal coniferous forests in this region.

L98-100: here in the methods, it sounds like the site is really only coniferous forest. The intro emphasizes that a hemiboreal forest is characterized by a species composition that combines elements of both northern coniferous and southern deciduous forests. The reader needs more justification/explanation here of what the deciduous characteristics are of your forest site, and what distinguishes it from other coniferous eddy covariance sites that have been studied.

We thank the reviewer for this observation. We will clarify that while the site is biogeographically "hemiboreal" (located in the boreal–temperate transition zone), the local tower footprint is dominated by an old (up to 230 years) coniferous stand of Scots pine and Norway spruce.

The site's distinctiveness relative to other EC stations lies in this stand age, coniferous dominance, and location on sandy upland soils within the Nordic-Baltic region. We will revise the introduction to clearly distinguish between the regional landscape classification and the specific local stand composition

Partitioning method (L153-154): consider also calculating daytime partitioning – if your results agree between the two methods, it would make them more robust. Neither partitioning method is completely reliable, and can lead to very different results depending on site. If you don't consider both, please provide a reason for choosing night time partitioning.

Thank you for this suggestion. We agree that both nighttime and daytime flux-partitioning approaches have limitations and that partitioning uncertainty should be acknowledged. However, for the aims of the present study, we prefer to retain the nighttime partitioning approach rather than additionally applying daytime partitioning. Our main reason is methodological independence. Daytime partitioning method (light-response approach of

Lasslop et al., (2010)), estimate GEP directly as a function of radiation and include VPD-related constraints. Because our study analyses LUE as GEP/PAR and evaluates PAR and VPD as potential drivers of RUEs, using a partitioning method in which GEP is already parameterized by radiation and VPD could introduce a built-in dependence between the derived efficiency metrics and the explanatory variables. This would complicate the interpretation of the RF and GAM driver analyses and could increase the risk of circular reasoning.

For this reason, we used the nighttime partitioning method, in which R_{eco} is estimated from the nighttime temperature response and GEP is then derived from NEP and R_{eco}. This approach is not free of uncertainty, but it better preserves PAR and VPD as external environmental drivers in the subsequent analysis. In the revised manuscript, we will explicitly justify this choice and add a statement that partitioning uncertainty propagates into GEP-derived metrics, especially CUE and LUE.

Growing season determination: based on Fig D1 (L763), in the current GS determination method, SOS and EOS dates rely heavily on the shape of the seasonal curve instead of on when plants were active. For instance, years 2020-2022 have SOS/EOS dates well into when plants were active, while for other years the growing season starts and ends right around when plants begin showing activity. Years 2020-2022 have clearer peaks than the other years. Based on this, I believe you should use a metric that is more robust to GEP growing season shape – for instance fitting SSA and then using a threshold of max GEP for that year would yield more reliable results across years. The current approach does not capture the phase of plant activity equally well in different years, which is what growing season should measure.

Thank you for this important comment. We agree that growing-season definition can affect seasonal flux sums and RUEs estimates. Because no direct phenological observations were available at the site, we had to define the growing season objectively from the EC-derived GEP record. We used the double-sigmoid method of Gonsamo et al. (2013) because it was developed for extracting transition dates from carbon-flux data and provided good fits for our daily GEP series in all years ($R^2 = 0.88–0.93$).

We acknowledge that SOS and EOS derived from this method represent a functional period of sustained ecosystem carbon uptake, not directly estimated plant phenology. This distinction is important for our study because RUEs should be calculated over periods of active ecosystem functioning and meaningful CO_2 uptake.

We also note that, although SOS and EOS shifted in 2020–2022, the timing of the peak of the season (POS) remained relatively stable across years. This suggests that the model captured the main phase of ecosystem carbon uptake consistently, while the shorter LOS in 2020–2022 likely reflects a shorter period of sustained carbon uptake rather than only a fitting artefact.

We would prefer to retain the double-sigmoid model as the primary method because it provides a standardized and reproducible definition across all nine years. However, as a robustness check, we will follow the reviewer's suggestion and add an alternative threshold-based growing-season estimate using smoothed daily GEP and a fixed percentage of the annual maximum GEP. We will compare SOS, EOS, LOS, and growing-season RUEs between the two approaches and report whether the main conclusions are sensitive to the growing-season definition.

WUE (L176-182): please justify why you're using eWUE instead of another WUE metric (such as inherent or intrinsic WUE). Also, you should exclude at least a day or two after a rainfall event (due to strong soil evaporation signal in ET).

We thank the reviewer for this comment. We agree that the choice of WUE metric should be clarified more explicitly. In this study, WUE is used as ecosystem water-use efficiency, calculated as GEP/ET. This choice follows directly from the objective of the study, which is to assess ecosystem-scale carbon–water coupling from eddy-covariance fluxes, rather than leaf- or canopy-level physiological water-use efficiency.

We did not use intrinsic WUE because canopy conductance and transpiration were not measured directly. Similarly, we did not use inherent WUE because it includes VPD in the metric itself, whereas VPD is one of the main environmental drivers analysed independently in this study. Including VPD in the response variable would partly confound the subsequent driver analysis.

We acknowledge that rainfall can temporarily increase the evaporation component of ET. However, because our WUE metric is explicitly ecosystem-scale, total ET is the relevant denominator and includes both transpiration and evaporation. Removing rainfall days and subsequent days would shift the metric toward a canopy/transpiration-based interpretation, which is not the aim of this study. Moreover, the SPEI-based analysis already evaluates WUE across different hydroclimatic regimes, including wetter conditions. Even with rainfall-influenced days retained, WUE remained comparatively stable across the observed SPEI gradient, supporting the interpretation that ecosystem-scale carbon–water coupling was robust under the hydrometeorological variability captured in this study

LUE (L192): LUE is typically calculated using absorbed PAR, (APAR), not just PAR.

We appreciate the reviewer's comment regarding the use of absorbed PAR (APAR). We acknowledge that while APAR is a more direct measure of the radiation utilized by the canopy for photosynthesis, our research station was not equipped with the specific sensors (such as those required for continuous measurements of canopy reflectance and transmittance) necessary to calculate APAR accurately.

Calculating ecosystem-scale LUE based on incident PAR is a well-established practice in the eddy-covariance community, particularly for long-term stand-level assessments (Lagergren et al., 2005; Launiainen et al., 2022). This methodology allows for a direct comparison of our results with other significant studies in boreal and temperate regions. A similar approach was used in the 17-year analysis of a Scots pine stand in Finland (Launiainen et al., 2022) and in studies of mixed coniferous forests in Sweden (Lagergren et al., 2005).

SPEI (L243): please explain in more detail what is meant. Was 3 month SPEI calculated as a rolling window and compared to each daily RUE? Does SPEI have just 1 value each month, or is a new 3 month SPEI calculated for each day?

Thank you for this comment. The Standardized Precipitation-Evapotranspiration Index (SPEI) was calculated at a monthly resolution using a three-month accumulation timescale (SPEI-3), following the standard methodology of Vicente-Serrano et al. (2010).

To integrate this index with our daily resource-use efficiencies data, the calculated monthly SPEI value was assigned to every day within that specific month. We did not use a daily rolling window; instead, we used the monthly values to group daily RUE observations into the established hydroclimatic categories (from severe drought to severely wet) for the statistical analysis shown in Fig. 7. This approach is consistent with ecosystem-scale studies where monthly SPEI is used to characterize the seasonal moisture state that influences daily physiological responses (Vicente-Serrano et al., 2010; Q. Wang et al., 2021; Y. Wang et al., 2024). We will explicitly clarify this in the revised Materials and Methods section.

L643-666

- **Additional analysis needed to support SPEI claims: In section 4.4, better explanation is needed for why some SPEI bins show significant deviation from normal and others do not. Perhaps it has to do with the number of samples in each bin? This would be good data to provide alongside fig. 7. The range of variation**

for each bin should also be provided for all RUEs. Overall, additional supporting analysis is needed before clear takeaways can be made from the SPEI analysis in my opinion – I find some of the claims the authors make to be over-reaching based on the data provided.

- **In the WUE section (L643-666), authors imply the significant differences found during moderate drought and mildly wet conditions are minor since the actual range of variation spans the normal conditions. Is this not the case for CUE and LUE bins? This data must be provided. If significant differences in WUE can be disregarded, why can't significant differences in CUE and LUE be disregarded?**

We thank the reviewer for this important point. Our intention was not to disregard statistically significant WUE differences, but to distinguish statistical significance from ecological effect size. We agree that this distinction should be stated more clearly and that the distributional information behind the SPEI bins should be made more transparent.

The WUE differences under moderate drought and mildly wet conditions were statistically significant, but their relative magnitude was small: −10% and −6% compared with the Normal SPEI class. We therefore interpret them as minor functional shifts rather than strong changes in ecosystem carbon–water coupling. This interpretation reflects the ecological meaning of ecosystem WUE: variation within approximately 10% suggests that the forest maintained a relatively stable balance between carbon uptake and water loss, likely through conservative stomatal regulation and buffering of ecosystem ET.

In contrast, the relative deviations in CUE and LUE were substantially larger. CUE changed by +52%, +42%, and −31% across significant SPEI classes, indicating much stronger shifts in the balance between carbon uptake and respiratory losses. We interpret these changes as evidence of stronger metabolic uncoupling between photosynthesis and respiration under hydroclimatic anomalies. LUE also showed larger deviations than WUE, ranging from −23% to +21%, suggesting stronger changes in light-use dynamics than in ecosystem carbon–water coupling.

To make this comparison clearer, we will provide the number of daily observations (N) and interquartile ranges (IQR) for each RUE within each SPEI category, either in the Fig. 7 caption or in an separate appendix table. This will allow readers to assess both the sample size and the degree of overlap among SPEI categories. We will also revise the Discussion where needed.

L667 to 687 and L691-692

- **The CUE paragraph (L667 to 687) needs to be supported by another figure showing GEP, RECO, and NEP responses in the different SPEI bins to show that the behavior they are hypothesizing is in fact taking place.**
- **When explaining LUE response to SPEI, the authors appear to contradict previous sections when trying to explain why LUE increases from normal in mild drought/wet conditions, when they say that these conditions 'do not impair ecosystem function, as boreal species are adapted to such fluctuations' (L691-692). This contrasts both WUE and CUE decreasing under mild wet conditions. The explanation for the much higher LUE under severe drought also seems dubious without further evidence to support claims. Even if shaded canopy now has a relatively higher contribution, overall GEP would still be lowered if less of the canopy is photosynthesizing, and since LUE is calculated by GEP/PAR here, LUE should also be lower. Overall, the SPEI results seem difficult to interpret and unclear.**

We thank the reviewer for this critical comment. We agree that the interpretation of the LUE response across SPEI classes, especially under severe drought, was perhaps phrased too strongly in the current version and requires additional support and more cautious wording.

In the revised manuscript, we will clarify that WUE, CUE, and LUE are not expected to respond identically to SPEI classes because they represent different flux ratios: WUE reflects GEP relative to ET, CUE reflects NEP relative to GEP, and LUE reflects GEP relative to PAR. Therefore, an increase in LUE under a given SPEI class does not necessarily imply that WUE or CUE should also increase. However, we agree that the current statement that mild anomalies "do not impair ecosystem function" is too broad and may appear inconsistent with the observed decreases in WUE and CUE under some mild conditions. We will remove or rephrase this statement.

Second, we agree that our explanation for the apparent LUE increase under severe drought is speculative. The shaded-canopy mechanism was intended as a possible hypothesis, not as a demonstrated mechanism. Because we do not have vertical canopy fluxes, leaf-level measurements, or direct information on sunlit versus shaded foliage contributions, we cannot support this explanation with data.

As noted in our response to the previous SPEI-related comment, we will provide additional distributional information. We will also examine the underlying absolute fluxes, especially GEP, NEP, Reco, and PAR, across SPEI bins. This will allow us to determine whether the LUE pattern reflects a meaningful change in GEP relative to PAR or whether it should be treated mainly as an uncertain ratio-based pattern.

If the additional flux comparison does not provide clear support for the severe-drought LUE interpretation, we will explicitly state that the LUE increase under severe drought remains an unexplained observation. In that case, the shaded-canopy explanation will be retained only as a tentative hypothesis requiring further canopy-layer measurements or removed from the main interpretation. We will also add a cautionary statement if the severe-drought bin has a limited sample size.

L125-126: what qualifies as unreliable values?

Thank you for pointing this out. We agree that the term “unreliable values” was too vague. By this we meant records affected by sensor malfunction, maintenance periods, data-logging problems, or physically implausible values outside the expected measurement range of the meteorological sensors. We will revise this sentence to make the quality-control procedure clearer.

L151: need to define NEE

Thank you for this suggestion. To ensure maximum clarity and avoid potential confusion regarding sign conventions, we have decided to remove the term NEE (Net Ecosystem Exchange) from the manuscript entirely and will use NEP (Net Ecosystem Production) consistently throughout the text, figures, and appendices.

As noted in our methodology, we follow the ecological sign convention, where positive NEP values indicate net ecosystem carbon uptake and negative values denote release. Consequently, all diagnostic plots previously referencing NEE, such as the gap-filling performance analysis in Appendix A, have been updated to represent NEP.

L344-346: there’s hardly any difference between which variable dominates each of the RUEs in Table G1, I feel like this sentence is misleading.

Thank you for this comment. We agree that the wording “dominated by” may overstate the contrast between component fluxes, because both components of each ratio showed strong partial correlations with the corresponding RUE. Our intention was to identify the component with the strongest association, not to imply that the other component was unimportant. We will revise the sentence to avoid misleading wording.

L348-356: Lines identical content to L338-346, one of the two passages needs to be deleted.

Thank you for catching this. We agree that this passage is duplicated. The duplication occurred during editing, and we will remove the repeated text.

Fig 5: it's hard to see the black lines, and even harder to see the grey shading.

Thank you for pointing this out. We agree that the visibility of the GAM fits and confidence intervals should be improved. We will revise Fig. 5 to make the black fitted lines and grey confidence shading clearer.

L532-534: Please justify more why IAV of CUE is primarily driven by NEP, when there seem to be very small differences in partial r between NEP and GEP influences on CUE (0.996 vs -0.964). Here GEP seems to drive CUE only very slightly less?

Thank you for pointing this out. We agree that the wording "primarily driven by NEP" may overstate the contrast between the two component fluxes. The partial correlations for CUE were strong for both NEP and GEP, although the association with NEP was slightly stronger. Our intention was to emphasize that interannual CUE variability reflects changes in net carbon retention, but we agree that the current wording could imply that GEP was unimportant. We will rephrase in the revised manuscript.

L576: The structure of section 4.3 is a bit hard for the reader to follow. Perhaps small changes like combining some of the paragraphs that are about the same climatic driver could help.

Thank you for the suggestion. We agree that the readability of Sect. 4.3 can be improved. However, we prefer to retain the current overall structure because it separates the interpretation of the main daily drivers and allows each driver to be discussed in relation to the relevant RUEs. This structure follows the logic of the RF and GAM analyses and avoids mixing distinct mechanisms such as atmospheric dryness, radiation limitation, and thermal effects.

Rather than substantially reorganizing the section, we will improve the transitions between paragraphs and combine the sentences where the same driver is discussed without a clear conceptual break.

L599-605: I don't think Fig. 5j really supports the slight decline in response at highest irradiance. It's not clear visually, and if the high VPD effects on WUE and LUE are

discounted as model artifacts in the previous paragraph, it's hard to understand why the much milder signal of CUE decreasing at high PAR (which I honestly can't see in the fig) should be significant and not also a model artifact. Perhaps showing similar figures of GEP, RECO, and NEP vs PAR could help support these claims.
Thank you for this comment. We agree that the slight decline in CUE at the highest PAR values is not visually strong enough in Fig. 5j to support a clear mechanistic interpretation. We also agree that, for consistency, weak edge features of GAM curves should not be over-interpreted. We will therefore remove the statement about a decline in CUE at high irradiance. The revised interpretation will focus on the robust part of the response: CUE increases with PAR at low to intermediate irradiance and then tends toward saturation. This is sufficient for our main conclusion that PAR sets an important envelope for daily CUE, without requiring interpretation of a weak high-PAR edge effect.

L629-631: why is air temp important for CUE in one daily analysis and not the other? Shouldn't that make it a less robust driver of CUE overall, and perhaps not be named the second-most important driver of CUE (L616)? I understand this is because of the results of the random forest model, but in this case you should justify why you trust the random forest over the GAM analysis. Since other prominent drivers of RUEs in the random forest tend to agree more with the GAM analysis than in the case of Tair and CUE, I think your analysis does not necessarily support daily Tair being a main driver of CUE.
Thank you for this important point. We agree that the evidence for air temperature as a main standalone driver of daily CUE is weaker than for the other driver–RUE relationships. The RF and GAM analyses answer different questions: the RF model evaluates predictor importance in a multivariate context, including interactions and covariation among predictors, whereas the single-driver GAM shows the marginal relationship between CUE and air temperature alone. Thus, we do not interpret the weak GAM relationship between CUE and air temperature as evidence for a strong direct daily temperature effect. In the revised manuscript, we will clarify that the RF analysis was used to rank predictor importance in a multivariate setting, whereas the GAM analysis was used to examine the shape of individual response curves. Therefore, in discussion, we will describe air temperature as a secondary, context-dependent predictor of daily CUE rather than as a robust main driver.

References

Lagergren, F., Eklundh, L., Grelle, A., Lundblad, M., Mölder, M., Lankreijer, H., & Lindroth, A. (2005). Net primary production and light use efficiency in a mixed coniferous forest in Sweden. *Plant, Cell & Environment*, *28*(3), 412–423. https://doi.org/10.1111/j.1365-3040.2004.01280.x

Lasslop, G., Reichstein, M., Papale, D., Richardson, A. D., Arneth, A., Barr, A., Stoy, P., & Wohlfahrt, G. (2010). Separation of net ecosystem exchange into assimilation and respiration using a light response curve approach: Critical issues and global evaluation. *Global Change Biology*, *16*(1), 187–208. https://doi.org/10.1111/j.1365-2486.2009.02041.x

Launiainen, S., Katul, G. G., Leppä, K., Kolari, P., Aslan, T., Grönholm, T., Korhonen, L., Mammarella, I., & Vesala, T. (2022). Does growing atmospheric CO2 explain increasing carbon sink in a boreal coniferous forest? *Global Change Biology*, *28*(9), 2910–2929. https://doi.org/10.1111/gcb.16117

Vicente-Serrano, S. M., Beguería, S., & López-Moreno, J. I. (2010). A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index. *Journal of Climate*, *23*(7), 1696–1718. https://doi.org/10.1175/2009JCLI2909.1

Wang, Q., Zeng, J., Qi, J., Zhang, X., Zeng, Y., Shui, W., Xu, Z., Zhang, R., Wu, X., & Cong, J. (2021). A multi-scale daily SPEI dataset for drought characterization at observation stations over mainland China from 1961 to 2018. *Earth System Science Data*, *13*(2), 331–341. https://doi.org/10.5194/essd-13-331-2021

Wang, Y., Peng, L., Chen, T., Yu, P., Zhang, J., & Xia, C. (2024). Driving Forces and Ecological Restoration Revelation in Southwest China Based on the Divergence Characteristics of Ecosystem Compound Use Efficiency. *Forests*, *15*(4), 641. https://doi.org/10.3390/f15040641