Specific comments

1. Scientific framing and justification of tool choice.

The introduction would benefit from a clearer articulation of the scientific questions motivating this work. Parts of the current manuscript read as if improving DAISY model performance is the primary objective, rather than a means to address broader bio-ecological questions. Explicitly clarifying what real-world processes or uncertainties this model improvement aims to resolve, and why DAISY is an appropriate tool for this purpose, would strengthen the scientific contribution and better justify the modelling effort.

In addition, the choice of output variables lacks clear justification. The importance and relevance of NEE and LE should be explained more explicitly. These concepts may be straightforward for the authors, but additional explanation would benefit general readers and clarify their scientific relevance.

To account for your remarks as well as the first reviewer's, we suggest updating the last paragraph of the introduction as follows:

"In this case study, five growing seasons of winter wheat, cultivated on a Belgian site, were simulated using Daisy soil–plant–atmosphere model (Hansen et al., 2012). In addition to the stomatal coupling implemented with Farquhar biochemical model (Plauborg et al., 2010), this open-source model explicitly simulates the surface energy budget. Soil water movement is described by Richards' equation while accounting for drainage pipes, which are widely installed in agricultural fields. Together, these features make Daisy suitable for studying crop yield as well as coupled carbon and water flux dynamics under current and future conditions.

In a previous paper, we have identified key parameters controlling crop yield, energy and carbon fluxes and discussed the implications of contrasting water conditions on sensitivity analysis results (Delhez et al., 2025). The objective of this follow-up study was twofold: to improve the simulation of biomass, carbon and water fluxes through calibration and to explore the trade-offs between these outputs. To this end, we applied a swarm-based MOO algorithm called Speed-constrained Multi-objective Particle Swarm Optimisation (SMPSO) targeting (i) dry matter of vegetation organs (DM), (ii) Net CO2 Ecosystem Exchange (NEE) and (iii) latent heat flux (LE). Accurate simulation of these variables is essential for food production assessments and for quantifying crop carbon and water budgets. Furthermore, analysing trade-offs, particularly between carbon (NEE) and water (LE) fluxes, could reveal potential limitations in their coupled representation within the model."

2. Model simplification, dynamic processes, and interpretation

The authors provide a clear and well-reasoned discussion of structural limitations of the current model (e.g. constant SLA). To strengthen this section, it would be helpful to clarify whether introducing additional dynamic processes is expected to significantly improve predictions (e.g. of LE and NEE) for answering the research questions, and what will be the trade-offs by making the model more complex.

You are right to bring that up. Having a dynamic SLA (or specific area of other organs) could help us to better fit biomass, for example, but would worsen NEE. If we increase SLA after anthesis, photosynthesis would be higher as well as dry matter. Therefore, dry matter predictions would more likely be closer to observations, but NEE would be even more negative. This was not an appropriate example. However, we think it's important to discuss the model assumptions, but also the lack of data leading to these assumptions. For instance, dynamic SLA is already implemented in Daisy, but we assumed it was constant as we have only one measurement per season.

To correct this part, we suggest combining this discussion with the previous paragraph about dry matter partitioning:

*“**Wrong modelling assumptions**. Modelling real-world system comes with simplifications. For instance, in this study, the partitioning schemes were derived from S22 since root measurements were only made during that season, as well as early distinction between leaf and stem. These partitioning coefficients were also applied to the other cultivars, but they might be inadequate and partly contribute to this trade-off. If too much assimilated carbon is allocated to the roots at the expense of the stem, the model would underestimate stem DM which is crucial for grain filling during the reproductive stage. Conversely, as leaves are not directly affected and are the major contributor to photosynthesis, the influence on NEE would be attenuated. Besides this example, many assumptions could play a role in these results, such as the organ death rates depending only on development stages or the non-dynamic leaf abscission rate. However, accounting for these within the model would be, in most cases, ineffective as model users hardly have enough data to support this level of complexity.”*

3. Linking model limitations to real-world mechanisms (LE and NEE).

The discussion of limitations in predicting LE and NEE is thoughtful and balanced. It would be particularly insightful to briefly expand on how these diagnosed limitations relate to real-world mechanisms. For example, are the relevant processes already well understood from experimental studies but not yet parameterised in models, or do the results point to areas where targeted experiments or observational designs are needed? Even a short reflection along these lines would enhance the scientific relevance of the work.

We rephrased the section 4.3 to better reflect the advances in experimental studies and advances in modelling:

“Stomatal behaviour plays a central role in plant regulation of water loss and carbon uptake, and has therefore been extensively studied (Damour et al., 2010). While it is well established that mild edaphic drought induces a decrease in carbon assimilation by reducing the stomatal aperture (Beauclaire et al., 2024), plant responses to heatwaves and high atmospheric demand are less evident. Most studies report reduced g_s under high VPD (Grossiord et al., 2020; Bourbia and Brodribb, 2024), aligning with the theoretical predictions of stomatal models (Sabot et al., 2022). On the other hand, high temperatures seem to trigger stomatal opening to promote transpiration cooling (Urban et al., 2017), but reported results vary greatly among species (Moore et al., 2021). Under the combined effect of high temperature and VPD, Marchin et al. (2022) observed a significant increase in g_s for two well-watered species, interpreting this as a strategy to prevent leaf overheating via enhanced LE loss without any increase in carbon assimilation. They later confirmed that this behaviour, referred to as stomatal decoupling, is not restricted to well-watered species (Marchin et al., 2023).

This stomatal decoupling aligns with ecophysiological studies pointing out the role of non-stomatal factors during photosynthesis. Inside the leaf, CO_2 diffuses through air spaces and membranes to reach the sites of carboxylation inside chloroplasts, whereas H_2O moves from the xylem network to the stomata (Flexas et al., 2012; Sack and Holbrook, 2006). These different pathways can be affected independently, where carbon assimilation can be restricted by non-

stomatal factors, often dominated by mesophyll conductance (Flexas et al., 2012). Over the last decade, these non-stomatal limitations have received increased attention, especially under water-stressed conditions (Nadal and Flexas, 2018; Gago et al., 2020). Based on leaf gas exchange measurements, experimental studies proved that these limitations should be considered in leaf models linking photosynthesis and transpiration (Yang et al., 2019; Beauclaire et al., 2024; Chen et al., 2025). Despite that, these complex processes are still not fully understood, and hence barely integrated into crop or terrestrial models (Vidale et al., 2021). As an example, an empirical equation for the mesophyll conductance was implemented in the land surface model JSBACH (Knauer et al., 2019a). The authors extensively discussed the environmental drivers, the modelling implications as well as the effects on the predicted fluxes in a following paper (Knauer et al., 2019b).

These non-stomatal limitations, or stomatal decoupling, are not modelled in Daisy. Based on stomatal coupling (Eq. 1), the increase in VPD reduced the predicted slope between carbon assimilation and g_s. Consequently, the model tended to underestimate g_s and LE fluxes and overestimated sensible heat, leading to a misrepresentation of energy partitioning under such conditions. On the other hand, observations show that plants likely opened their stomata, promoting water loss, while carbon assimilation did not increase accordingly. This stomatal decoupling was particularly visible in April 2019, as the heatwave lasted several days, but also appeared during shorter warm periods, suggesting a recurrent plant strategy. Future work should (i) continue the effort in better understanding the mechanisms of non-stomatal limitations affecting carbon assimilation and (ii) incorporate these mechanisms in process-based models, particularly under such conditions."

4. Clarity and presentation

Overall presentation is good, but several instances of vague wording (e.g. "good accuracy"), long sentences, and **undefined acronyms** reduce clarity. Addressing these issues would improve readability.

We can send it to be editing (English native speakers) for lengthy phrasing and global clarity.

Technical corrections

Abstract

Line 12: "good accuracy" is vague. The term "good" lacks context for readers unfamiliar with the subject. Consider removing it or replacing it with quantitative information.

Is it because readers might be unfamiliar with magnitudes of DM, NEE and LE, and thus the RMSE values are hard to interpret? Or do you think RMSE values are not enough to quantify accuracy? Depending on your answer, we get replace good accuracy by "satisfying results" or add rRMSE in addition to RMSE values.

Line 14: "wrong parameterisation" is a strong claim. The discussion suggests that simplification may contribute to bias, but no quantitative evidence is provided to demonstrate that the parameterisation is incorrect. Model simplification is a design choice; the issue is appropriateness rather than right or wrong. Rephrasing would improve accuracy.

Wrong parameterisation doesn't only refer to the simplifications, but also the underestimation of heterotrophic respiration by Daisy. In section 4.2, we can add this sentence at the end of the corresponding paragraph to make it clearer: "Adjusting soil respiration parameters would require long-term simulations and data from this study site."

Introduction

Line 35: The sentence introducing g_1 is abrupt. It moves directly from physiology to a specific parameter without transition. A brief link from physiological concepts to their model formulation would improve readability.

Indeed, we can add a generic equation linking carbon assimilation (An) and stomatal conductance (gs) as follows:

"In these couplings, carbon and water fluxes between vegetation and atmosphere are assumed to only be regulated by stomatal behaviour. Widely used models of stomatal conductance g_s can be generally expressed as:

$$g_s = g_0 + g_1 \cdot f(E) \cdot A_{net}$$

Where A_{net} is the net leaf CO2 assimilation rate, $f(E)$ is a function of environmental influences such as Vapour Pressure Deficit (VPD) and CO2 concentration at the leaf surface, and g_0 and g_1 are parameters. When water is not limiting, plants open their stomata in order to maximise their carbon uptake. Conversely, they reduce their stomatal aperture under water stressed conditions, compromising between carbon assimilation and water loss (Manzoni et al., 2011). This response is often represented by a decrease of the slope parameter g_1(Zhou et al., 2013; Liu et al., 2009)."

The Equation numbers would be updated for the entire manuscript and the variable description in lines 101-102 will be reduced, as some have already been introduced in this new equation.

Line 51: The term equifinality is introduced without explanation. A brief definition here, similar to the one later in the discussion, would improve accessibility.

Good point, we will add a similar explanation.

Line 57: NEE and LE are introduced without justification. Please briefly clarify why these outputs were selected and what ecological or biophysical significance they represent.

Lines 58–59: The flow may improve if concepts are introduced in a consistent order (e.g. DM, then NEE, then LE), rather than switching between topics.

We took your two comments into account when rephrasing the last part of the introduction (see at the beginning of this document).

Methods

Line 69: TOB, SKY, SAH, and SMA are unclear. Please clarify whether these refer to varieties or growing seasons. If they represent seasons, it may be clearer to use digits.

We can denote them as S11, S15, S17, S19 and S22. All the manuscript will be updated accordingly.

Line 102: Unsure whether g_1 here refers to the same parameter introduced in the introduction.

Yes it is, it should be clearer with the generic equation in the introduction. We can also rephrase lines 99-100 as: "…a stomatal conductance empirical model (Eq. 1). In this version, the stomatal conductance *gs* is based on Leuning (1995) and adjusted to account for hydraulic signalling (Plauborg et al., 2010):"

Line 104: It would be beneficial to clarify why model modifications were required.

This is specified for the diffuse radiation in the appendix A (considering cloudy skies), but not explicitly for the others, we can add it in the appendix as well, or do you think in the main text would be better?

Line 136: The description of "individual exploration" and "social learning" is unclear. Does this refer to local versus global optima?

These terms are defined and commonly used with particle swarm-based techniques. We agree that it might be unclear for unfamiliar users, we can add a reference to local and global optima between brackets, rather than "(its own experience)".

Line 154: The description of "once for each growing season" seems to suggest all seasons were used.

We rephrased this to account for your comment and the first reviewer's: *"For each parameter sets (i.e. potential solution) generated by SMPSO, the parameter values were transcribed into Daisy setup files. These include cultivar-specific parameters, with possibly different values for each season, and site-specific parameters that are shared (Table 2). Hence, the model was executed four times, specifically for S15, S17, S19 and S22. The three objective functions (rRMSE) were then aggregated and computed from these four runs and passed back to SMPSO algorithm."*

Line 157: Please clarify how swarm size and total simulation number were chosen. Were these based on empirical experience or hyperparameter tuning? Showing convergence behaviour (e.g. in an appendix) would increase confidence.

The number of simulations was chosen based on the convergence. We can indeed add the convergence behaviour in appendices. Regarding the swarm size, it was more empirical. Common values are around 100, but after looking at the LE-NEE trade-off, which was really narrow, we increased it to have a better spread and be sure that the shape is not due to local optima. This increases calculation time but improves the exploration.

Line 165: Please clarify what is meant by EF being "easily interpreted." Is this relative to other metrics, or related to its physical meaning for the output variables?

This is because EF has no dimension and the threshold for considering a result acceptable is known. We can write the "which is easily interpreted" at the end of line 169.

Results

Line 176: "Most solutions" is vague. Please consider reporting this as a percentage.

With this sentence, we are not talking about the black solutions but the grey ones, saying that grey solutions don't diverge far from the black solutions (contrary to LE-NEE). We understand that it can be confusing, and if we were talking about the black solutions, we could have reported a percentage. We can rephase this as:

"...where the curved Pareto front extends over a large part of the objective space. This 2D front portrays 29% of the solutions from Fig. 1a, and few other solutions diverge far from this front."

Line 202: "Good overall agreement" lacks quantitative support. Please provide numerical values or remove the qualitative descriptor.

Numerical values are provided for hourly NEE, as we calibrate against these data, but not for daily values. We can remove this part.

"As depicted in Fig. 3, simulated temporal variations generally follow observed trends, although predictions tend to be more negative (i.e. more CO2 uptake) during the vegetative stage."

Line 204: Figure 3 also shows underestimation toward the end of the season; this could be mentioned explicitly.

Indeed, it is relevant. When reaching maturity, Daisy stops the plant processes. Therefore, there is no longer any autotrophic respiration until harvest, which could explain this difference. We will add a few lines about this.

Lines 245 and 253: Table numbering appears inconsistent. Table 1a and 1b may need to be Table 3a and 3b?

Indeed, thank you for the comment!

Discussion

Line 301: The statement that constant values may lead to significant bias is plausible, but quantitative evidence is not presented. Please clarify whether incorporating dynamic SLA (e.g. dependent on leaf age or water status) is expected to substantially improve predictions, or whether this remains a hypothesis.

This is answered in detail under one of your specific comments (3).

Lines 303–305: The sentence describing EC data filtering and uncertainty estimation could be rewritten for clarity.

This was rewritten, we hope this is clearer.

"To account for these, EC data are thoroughly filtered, resulting in low data coverage as in Table 1. Moreover, turbulence conditions are also considered in the uncertainty estimation, with an uncertainty related to the friction velocity, an indicator of turbulence strength (Pastorello et al., 2020)."

Line 335: CLM is introduced without prior explanation. Please define it when first mentioned.

As for STICS or JULES, that's a model name commonly accepted so we didn't describe the acronym. "such as the CLM model" might make it clearer?

Line 374: Please rewrite for clarity. It is unclear what is meant by "This decoupling behaviour." Does this refer to increased transpiration while assimilation remains unchanged? Also, please clarify whether "April 19" refers to a date April 19th or the year 2019.

This is taken into account in the updated paragraph, in response to one of your specific comments.