Author Comments – Response to Referee #2 (RC2)

We thank Referee #2 for their detailed, thoughtful, and constructive feedback. We are pleased that the reviewer found our manuscript well-written and a significant contribution. Below, we address each of the general, specific, and technical comments.

General Comments

RC2: The reviewer notes that while Questions 1 and 2 are well addressed, the answer to Question 3 lacks clarity due to the absence of stomatal conductance and O_3 uptake flux (FO_3).

AC: In response, we have revised the manuscript to more explicitly and quantitatively address research question 3: how ozone impacts interact with other environmental factors, and how an optimised model can help us understand these mechanisms, particularly on high-ozone days. To improve clarity, we have added modelled stomatal conductance, ozone flux to vegetation (FO_3), and soil moisture as key prognostic variables in our high-ozone day analysis (Section 2.4.3). These outputs allow us to distinguish between: 1) stomatal limitation, where high VPD and/or low soil moisture reduces conductance and FO_3, thus limiting O_3 damage, and 2) direct ozone stress, where elevated FO_3 and maintained stomatal conductance lead to reductions in GPP through biochemical effects. In Section 3.3, we now interpret observed and simulated GPP patterns on high-ozone days using these additional variables to identify the dominant mechanisms at each site. This mechanistic analysis is supported by optimised parameter values (e.g. g_1, p_0, a, and FO_3crit), and improves the attribution of GPP reductions to specific environmental and physiological drivers. We believe these additions now provide a clear and complete response to Research Question 3, and we thank the reviewer for prompting this improvement.

Specific Comments

RC2: Since the partitioned GPP is central to the inference made in this manuscript, it would help if the authors offered a description of how GPP was partitioned from observed net carbon flux. This could be as simple as a brief description with a reference to a citation that details the methods. Lines 130 – 132 claim that GPP and LE were estimated from net-carbon flux. Net carbon flux is used to estimate net ecosystem exchange and GPP. LE is not estimated from net carbon flux. It is typically estimated from H2O flux. The authors should consider correcting or clarifying if they have developed a technique or used an existing technique to estimate LE from net carbon flux.

AC: We thank the reviewer for pointing this out. We have revised Section 2.2 to clarify that GPP was derived from net ecosystem exchange (NEE) using standard partitioning approaches implemented in the ICOS ONEFlux pipeline. We also corrected the erroneous statement about LE and now clarify that LE is derived from water vapour (H_2O) flux measurements, not carbon flux:

“The half-hourly Gross Primary Production (GPP, μmol m^{-2} s^{-1}) and Latent Heat flux (LE, W m^{-2}) were derived from eddy covariance measurements at each site. GPP was estimated from net ecosystem exchange (NEE) using standard partitioning techniques implemented in the ICOS ONEFlux processing pipeline (Warm Winter 2020 Team, ICOS Ecosystem Thematic Centre, 2022). LE was derived from water vapour fluxes measured by the same system. All meteorological, GPP, and LE data are publicly available via the ICOS data portal. The data follow the standard format of ICOS L2 ecosystem products and are fully compatible with FLUXNET2015. Data processing was performed using the ONEFlux pipeline (https://github.com/icos-etc/ONEFlux). Basic site-level statistics and data coverage are reported in Table 1.”

RC2: The JULES damage scheme calculated the O3 damage factor, F, as a function of the stomatal flux of O3 (equation 7). It appears that this is the instantaneous stomatal flux of O3. However, the cumulative flux of O3 through stomata is typically used as the damaging quantity (Lombardozzi et al., 2013, Wittig et al., 2007). Many threshold-based O3 damage indicators are based on cumulative exposure or cumulative stomatal dose (i.e.: AOT40 and POD6). The authors could consider elaborating on this in the discussion section of this manuscript by discussing if it would be worthwhile to use cumulative O3 stomatal flux in future optimization studies. The JULES O3 damage factor, F, as it is formulated in the current study appears to be the same damage factor that is applied to both stomatal conductance (gp) and net photosynthesis (A). However, previous research suggests that net photosynthesis and stomatal conductance are differentially impacted by O3 (Lombardozzi et al., 2012a,b). Both quantities might not exhibit the same sensitivity to O3 or might not change at the same rate as a function of O3 uptake (Lombardozzi et al., 2012b). This suggests the use of separate damage factors, sensitivities, and critical O3 levels for stomatal conductance and net photosynthesis. Are a and FO3crit separately estimated for A and gp? These distinctions are important because they might have implications for modeling transpiration in a land surface model if stomatal conductance is involved. The results report only one value for FO3crit and a which implies that the same damage factor is applied to both A and gp. In the discussion portion of the paper, it would be worth discussing the reasoning behind the JULES modeling choices for the specific formulation of O3 stress on gp and A compared to other methods of incorporating damage factors in land surface models (see Lombardozzi et al. 2012a and b who tried various configurations of an O3 damage factor in the community land model).

AC: We thank the reviewer for this thoughtful suggestion. In the revised Discussion section, we clarify that the current JULES implementation uses instantaneous stomatal O_3 flux to compute the damage factor F, applied equally to both photosynthesis (A) and stomatal conductance (gp). We acknowledge that alternative approaches—such as those used in the Community Land Model (CLM)—use cumulative O_3 uptake (e.g., POD6) as a more biologically realistic indicator of damage (Lombardozzi et al., 2013; Wittig et al., 2007). We now note that incorporating cumulative dose metrics and distinguishing between photosynthetic and stomatal sensitivities (FO_3crit, a) could improve the representation of O_3 effects in future JULES developments.

“The JULES ozone damage scheme, as applied in this study, uses an instantaneous stomatal flux of ozone to compute a damage factor (F) that is applied equally to net photosynthesis (A) and stomatal conductance (g_p). This approach enables a simple and efficient integration into the model but may not fully capture the temporal dynamics of ozone-induced damage. Many other modeling frameworks use cumulative ozone uptake metrics—such as the phytotoxic ozone dose above a threshold (POD6)—to represent damage accumulation over time (Wittig et al., 2007; Lombardozzi et al., 2013). Moreover, empirical evidence indicates that A and g_p may respond differently to ozone, with distinct sensitivities and temporal responses (Lombardozzi et al., 2012a,b). Future versions of JULES could benefit from decoupling these effects by estimating separate sensitivity parameters (*a*) and critical thresholds ($FO3_{crit}$) for A and g_p, and by transitioning toward cumulative flux-based ozone stress formulations."

RC2: The diurnal cycles of partitioned and JULES simulated GPP are shown in Figure 5. Can the authors clarify whether these diurnal cycles were estimated using data and simulations from all seasons or just the summer?

AC: The diurnal cycles shown in Figure 4 are based on data and simulations from the full year, not limited to the summer season. We will clarify this in the caption of Figure 4:

"Figure 4: Comparison of the observed and simulated GPP diurnal cycles across all sites, averaged over the full year: (a) FI-Hyy, (b) FI-Var, (c) BE-Bra, (d) FR-Fon, (e) IT-BFt and (f) IT-Cp2. Shaded areas encompass plus and minus one standard deviation. The black line represents the observed GPP. The default simulated GPP are the dashed purple line (without O_3) and dashed green line (with O_3), and optimised simulated GPP are the purple line (without O_3) and green line (with O_3)."

RC2: Some statements about the diurnal cycle of GPP need clarification. The authors mention midday depressions in GPP at Mediterranean sites at line 394 and again at lines 465 - 467. Can the authors specify which GPP estimates show these midday depressions (partitioned or simulated)? The partitioned GPP from flux data (black line in diurnal plots) do not show midday depressions at the Italian sites (There does appear to be somewhat of a morning depression in partitioned GPP at IT-BFt). The simulated GPP suggests midday depression and diurnal asymmetry (higher fluxes in the morning) at the IT-BFt.

AC: We will clarify that the midday depression is primarily evident in the simulated GPP at IT-BFt and is only partially observed in the actual data. At IT-Cp2, the model does not show a pronounced midday dip. We will revise our statements in Section 3.3 and the Discussion to reflect this distinction and to better align with Fig. 6. Revised in section 3.3: "Mediterranean sites (IT-BFt and IT-Cp2) experience the highest ozone peaks (>60 ppb). At IT-BFt, the JULES-simulated GPP exhibits a pronounced midday decline, particularly in the optimised configuration with ozone effects, indicating a strong response to midday ozone stress. However, the observed GPP shows only a slight morning dip and continues increasing into the afternoon. At IT-Cp2, no distinct midday depression is observed in either the simulated or partitioned GPP." Revised in the Discussion: "Southern sites like IT-BFt exhibited a pronounced midday decline in simulated GPP, reflecting modelled ozone sensitivity and the interacting influence of high ozone concentrations and elevated VPD. However, the partitioned GPP at this site does not exhibit the same midday depression; instead, it increases gradually into the afternoon. At IT-Cp2, no midday dip is observed in either the simulated or observed GPP."

RC2: The discussion of O3 interactions with environmental factors on high ozone days (in section 3.3 and in the discussion section) needs more clarification and elaboration. It seems that the authors are using LE as a simple proxy for stomatal conductance (LE increases or decreases with changes in stomatal conductance). It could be helpful if the authors plotted the diurnal cycles of JULES simulated stomatal conductance and stomatal flux, FO3, as a third column in Fig. 7. At line 380, the authors mention that the midday peak of VPD and LE facilitates greater O3 uptake through higher stomatal conductance. This appears to be the case at many sites where the reduction in GPP from simulations that did not include O3 (reduction in GPP from purple line to green line) appear to be the highest during the midday period previously defined by the authors

(12 – 16). However, this does not seem to be the case for IT-BFt. The largest reduction in GPP at IT-BFt during high O3 days appears to take place in the morning hours when [O3] is not at peak. It appears that LE and VPD are also high before the 12 – 16 midday period at IT-BFt. Can the authors discuss this interesting exception more? Is there high morning stomatal conductance and morning stomatal O3 flux at this site?

AC: We thank the reviewer for these constructive observations. We agree that interpreting latent heat (LE) as a direct proxy for stomatal conductance can be misleading, as LE is influenced by multiple factors, including VPD and available energy. To better capture stomatal behaviour and ozone uptake, we now include diurnal plots of simulated stomatal conductance and stomatal ozone flux (FO_3) in a third column of Fig. 6, as suggested. This addition provides a more mechanistic view of site-specific O_3 uptake patterns and clarifies why GPP reductions peak at different times across sites. In particular, we now highlight and discuss the case of IT-BFt, where GPP reductions during high O_3 days are most pronounced in the morning, despite ozone concentrations peaking later in the afternoon.

RC2: The results about the boreal sites in section 3.3 can use more elaboration and clarification. Throughout the section, the authors use RMSE reductions to quantify O_3 At line 382, the authors mention that O_3 impacts on the boreal sites (FI-Hyy and FI-Var) are limited. However, the RMSE reductions between optimalizations with and without O3 at FI-Hyy are the largest among the sites (9.97 down to 0.52). This implies the impact of O_3 peaks is the strongest at the boreal site, FI-Hyy, compared to all other sites. Can the authors clarify or limit their statement to FI-Var?

1. **Are the authors referring to the partial correlation analysis when saying that FI-Hyy is less sensitive to O_3 overall (at line 387)? The JULES parameter optimization seems to suggest otherwise: FI-Hyy has higher sensitivity, *a*, and lower $FO3_{crit}$ among the sites (Figure 6). Is FI-Hyy less sensitive to O_3 or does it receive less O_3 exposure outside of select high O_3 days?**

2. **Line 385: Can the authors clarify what they mean by "simulations without O_3 significantly underestimate GPP"? In Fig. 7, it appears that the simulations without O_3 (purple line) estimate much higher GPP compared to the partitioned GPP (black line).**

AC: We thank the reviewer for this important clarification. In response, we have revised Section 3.3 to distinguish between (i) the absolute RMSE reduction at FI-Hyy—which is indeed large due to the model initially overestimating GPP without ozone effects—and (ii) biological sensitivity to O_3, which we interpret based on JULES parameters (*a*, $FO3_{crit}$) and partial correlation analysis. While FI-Hyy shows a strong model performance improvement after including ozone effects, this likely reflects both the correction of structural model bias during high-O_3 episodes and the fact that such episodes are rare at boreal sites (see Table 1). The improvement is therefore event-specific rather than indicative of sustained ecological sensitivity across the growing season. We now explicitly limit our statement about low ozone sensitivity to FI-Var. Additionally, we corrected the misleading phrase at line 385 and clarified that the model without O_3 consistently overestimates GPP at FI-Hyy during high-ozone episodes—even though such events are rare. These rare but impactful events explain the large RMSE reduction when ozone effects are included, despite limited overall ecological sensitivity. This is consistent with the low frequency of elevated ozone concentrations reported in Table 1. In section 3.3, we rephrased the paragraph about boreal sites: "At the two boreal sites (FI-Hyy and FI-Var), ozone peaks reach moderate levels (~46 and 44 ppb, respectively), but their impacts on GPP differ. FI-Var shows minimal response to ozone, with only a 1.3% decrease in RMSE (from 2.34 to 2.31 μmol CO_2

m^{-2} s^{-1}), suggesting low ecological sensitivity. In contrast, FI-Hyy exhibits a large RMSE improvement—from 9.97 to 0.52 µmol CO_2 m^{-2} s^{-1} (a 95% reduction), when ozone effects are included. However, this performance gain does not reflect sustained biological sensitivity. Rather, it stems from a systematic overestimation of GPP by the ozone-free model during high-O_3 episodes. These episodes are rare (see Table 1), but when they do occur, the model without ozone consistently overestimates GPP. The inclusion of ozone damage corrects this bias. The partial correlation analysis and the limited ambient ozone exposure outside these rare events support this interpretation. We therefore distinguish between improved model–data agreement due to structural correction and true ecological ozone sensitivity, the latter being more clearly limited at FI-Var."

RC2: The authors could consider revising the section on Mediterranean sites (starting at like 394). As I mentioned in the previous comment, I am particularly concerned about the claim that compared to other sites, the Italian sites exhibit stronger O3 induced reductions in GPP (line 395). Again, FI-Hyy appears to exhibit the largest reduction in RMSE during high O3 days (a reduction from 9.97 to 0.52). BE-Bra also shows a higher or comparable reduction in RMSE (7.57 down to 3.09) compared to IT-Cp2 and IT-BFt. This needs to be corrected or clarified.

AC: We thank the reviewer for this helpful observation. We have revised the corresponding paragraph in Section 3.3 to clarify that while Mediterranean sites such as IT-Cp2 and IT-BFt experience high ambient ozone concentrations, the magnitude of model improvement (RMSE reduction) is not the highest across all sites. FI-Hyy and BE-Bra show larger or comparable reductions. The revised text reflects this nuance and avoids overstating ozone sensitivity in Mediterranean ecosystems, highlighting instead the complex interplay between ozone concentrations, physiological traits, and model calibration outcomes. In section 3.3, we rephrased the paragraph about Mediterranean sites: "Mediterranean sites (IT-BFt and IT-Cp2) experience the highest ozone peaks (>60 ppb). At IT-BFt, the simulated GPP shows a pronounced midday decline, especially in the optimised configuration with ozone effects, suggesting a strong response to midday ozone stress. However, the observed GPP shows only a slight morning dip and continues increasing into the afternoon. At IT-Cp2, no distinct midday depression is observed in either the simulated or partitioned GPP. While these sites do show reductions in RMSE after including ozone effects—46% at IT-Cp2 (from 5.82 to 3.14 µmol CO_2 m^{-2} s^{-1}) and 0.8% at IT-BFt (from 6.54 to 6.49 µmol CO_2 m^{-2} s^{-1}) these improvements are not the largest among all sites. Indeed, FI-Hyy and BE-Bra show greater RMSE reductions during high ozone days. This suggests that while Mediterranean sites face high ozone concentrations, the degree of ozone-induced GPP reduction may vary depending on the interplay of environmental stressors and model representation. The results highlight the importance of site-specific calibration and caution against generalising Mediterranean sites as the most ozone-sensitive solely based on ozone concentration levels."

RC2: The claim at line 465 needs elaboration: "Southern sites like IT-BFt and IT-Cp2 exhibited pronounced midday declines in GPP, reflecting their heightened sensitivity to ozone and the compounding effects of high VPD and LE." The model simulated midday declines in GPP only appear at IT-BFt in Fig. 7. Please clarify what the authors mean by midday (12 – 16 hour) decline in GPP at IT-Cp2. The authors mention compounding effects of high VPD and LE at the southern sites at line 466 attempting to make a case for multiple stressors exacerbating ozone impacts. At IT-BFt, I can see the authors' claim in the model simulations. The modeling does suggest that GPP declines past the 10th hour when VPD is high and further declines when O3 impacts are added to the modeling. However, the partitioned GPP (black line) does not show this type of compound stress at IT-BFt. Partitioned GPP is showing the opposite. It increased into the afternoon hours (after 10 when VPD is high) which suggest there is not much midday or afternoon

water stress. The authors might want to elaborate on these differences between the partitioned GPP and JULES simulated GPP when discussing the potential of a compound water stress and O3.

AC: We appreciate this detailed observation. In response, we have revised the text in Section 3.3 to clarify that midday GPP declines are primarily present in JULES simulations at IT-BFt, not in the partitioned GPP, and not at all at IT-Cp2. The revised paragraph now distinguishes between modelled and observed responses and emphasises that the simulated declines may reflect model sensitivity to co-occurring VPD and O_3, rather than compound stress seen in observations. We acknowledge that the observed GPP at IT-BFt continues to rise into the afternoon, suggesting that water stress is not as limiting as the model predicts. This discrepancy is now explicitly discussed.

Last paragraph of section 3.3 was revised: "Interestingly, despite the strong midday declines in GPP at Mediterranean sites in the model, Figure 6 shows that this behaviour is not consistently present in the observations. At IT-BFt, the JULES-simulated GPP exhibits a sharp midday reduction, especially when ozone effects are included, suggesting a modelled compound stress due to high VPD and ozone uptake. However, the partitioned GPP at this site increases during the same period (after 10:00), indicating that stomatal closure due to VPD is not occurring to the extent the model assumes. This divergence points to a possible overestimation of midday water limitation in the model configuration. At IT-Cp2, neither the modelled nor observed GPP shows a distinct midday dip, indicating that ozone and VPD effects are less pronounced or not synchronised enough to produce a compound stress response. These site-specific dynamics reinforce the need for more accurate representation of stomatal regulation under co-occurring stresses in Mediterranean systems."

Technical Comments

RC2: Fig. 2a site distinction

AC: We thank the reviewer for this helpful suggestion. We have updated Fig. 2a by increasing colour contrast and line thickness to improve the visual distinction between sites. These changes make the time series more readable, especially when printed or viewed in greyscale.

RC2: The factor 1.6 on line 168 is a factor to convert from conductance to CO2 to conductance to H2O (ratio of CO2 and H2O diffusivities). The conductance to water vapor is gp.

AC: We thank the reviewer for this clarification. We will revise the sentence to explicitly state that the factor 1.6 accounts for the ratio of diffusivities of H_2O and CO_2 through the stomata. This factor is used to convert stomatal conductance from CO_2 to H_2O units, ensuring correct representation of ozone uptake in terms of water vapor conductance (g_p).

RC2: Should FO3 and FO3crit be in different units in equation 7? I am looking at line 201.

AC: Thank you. We will ensure unit consistency in Equation 7 and clarify that both FO_3 and FO_3crit are expressed in nmol m^{-2} s^{-1}.

RC2: Remove second comma after "vegetation" in line 243.

AC: Corrected.

RC2: Figure 7: It might help to double-check the units for VPD on the y-axis. Is it supposed to be displayed in hPa (not kPa)?

AC: We thank the reviewer for bringing this to our attention. We have reviewed the units and confirm that the vapour pressure deficit (VPD) in Figure 6 was plotted in kPa. To avoid confusion, we will explicitly label the axis as "VPD (kPa)" in the figure to ensure clarity of units.

RC2: Consider picking a consistent way to write GPP reductions in section 3.4. The authors make it clear that negatives mean decreases and continue to use negative quantities throughout most of the section. You could consider changing 5.22% to -5.22% at line 424 for consistency.

AC: We agree and have revised Section 3.4 to consistently express GPP reductions as negative percentage values (e.g., -5.22 %) throughout the text. This improves clarity and aligns with the convention used elsewhere in the manuscript when referring to decreases.