General comment

This is an interesting paper that meets the growing need for flexible and accurate tools able to exploring the complexity of the multiple interactions between plants and air pollutants, with ozone playing a prominent role (Dentener et al., 2020). Furthermore, this paper demonstrates the continuous effort of the research group of the University of York to refine and improve the DO3SE model in order to efficiently tackle emerging research questions. The authors answered more than satisfactorily to the Reviewers' comments, modifying the text and figures accordingly. A critical point is represented by the very limited calibration dataset, but given the main purpose of this research, i.e. explaining and exploring how the model works according to the modifications implemented, this limitation can be considered as acceptable. However, such a calibration could not be considered as robust enough to allow for large scale model application; for this reason the conclusions of this simulation experiment can be considered valid only for the pedo-climatic and management conditions explored and cannot be generalized outside the study area. The model calibration methodology is still a bit confusing and needs to be further improved. I suggest to include in the description section a diagram/table detailing the hierarchy of the calibrated processes, and for each process, the treatments, the varieties, the parameters and the target outputs together with the relative variation ranges.

Thanks for this suggestion. We have now included a Table in the supplementary text (Fig S5) which describes the calibration and evaluation process and clearly details the parameters calibrated/evaluated in each step and the data (by year, cultivar and treatment) that are used for each calibration and evaluation step.

Please find below some additional minor comments that ought to be addressed before the manuscript could be published.

Minor comments

• It is not clear what the time step of the model is. Some processes and modules are represented hourly, some others are daily. There is a bit of confusion. I suggest you better specify these details.

Line 138-141: Added a sentence to explain the different time steps of the key DO$_3$SE output data. 'The DO$_3$SE-Crop model requires hourly input meteorological and O_3 concentration data which is used to produce output on either an hourly (i.e. leaf physiology and short-term O_3 damage variables) or daily (i.e. phenology, soil moisture, long-term O_3 damage, C allocation, biomass and yield variables) time step.'

• LINES 109-112, PAGE 3: as stated in the general comment, the results obtained in this research could not be generalized to a global region, since they are representative only for sites with the same soil, weather and cropping system of the calibration site. The wording "global region, where" should be replaced by "in a site located in a region, where".

Line 111: Modified the sentence to 'This allows us to investigate the ability of the model to simulate O_3 damage for a comparable agro-ecological region where crop productivity is severely threatened by both O_3 pollution and climate change.'

• I understood that the water availability is not limiting; however it is not clear how the soil is represented in the model: mono or multi-layer? Does water availability (e.g. optimum vs severely stressed) have an impact on ozone flux regulation (i.e. via the stomata opening/closure mechanism) and therefore on damage? Please explain.

Line 224-228: We have amended the sentence at Line 235 to explain that the DO_3SE model uses a single soil layer bucket model and the stomatal ozone uptake is only affected once PAW < PAWt.

• Is the nitrogen in the plant-soil system simulated? If yes, please add details about nitrogen routines used. If not, please specify that the simulations were conducted under unlimited conditions for nitrogen availability according to experimental data.

Line 405-407: We do not simulate the plant-soil nitrogen system. To make this clear we have added the sentence as suggested at Line 415 'The model assumes non-limiting conditions for soil nitrogen, in accordance with the experimental data'.

• It is not clear how the integration of LAI from single leaf to canopy is handled when the leaf population (n) is higher than one. That means that cohorts of leaves with different ages are present in the four layers? Is the hypothesis of evenly distribution of LAI among layers still valid? Please add explanation on the manuscript.

We have amended the text to briefly explain how the leaf populations are distributed across the LAI defined canopy layers.

Line 186-187: DO_3SE-Crop allows for any number of representative leaf populations (pop) and canopy layers (n) to be defined over the course of the crop growing season by dividing leaf populations as they emerge evenly across the canopy layers defined by LAI

• It would be really interesting to show some result of the sensitivity analysis in the supplementary material (e.g. the ranking of parameters with related contribution in explaining the output variability share).

Plots showing the results of the sensitivity analysis have been added to the supplementary text (Fig S6) and referred to from section 2.2 DO_3SE-Crop calibration and evaluation.

• LINE 549, PAGE 17: Please clarify what you mean by "in realistic plant response" and add a more precise reference.

Line 565: Modified the sentence to '…that would result in appropriate dry matter partitioning within the plant'.

• Table 2. Long-term effect seems higher on tolerant than sensitive varieties: perhaps column labels were inverted? Please check. Also, i would add to the figure the observed % variation compared to the pre-industrial/baseline scenario.

Table 2: The typo has been corrected so that the correct results are shown for tolerant and sensitivity. We considered showing the observed values (i.e. the % Grain_DM loss between the modelled pre-industrial yield and the observed (i) AA and (ii) EO3 Grain _DM. However, the use of modelled the pre-industrial yield values creates uncertainties when compared with observed data; additionally this comparison gives us no further information about the split between instantaneous and long-term O3 effects. We therefore have decided not to include thi as we felt it could confuse the reader.

• Figure 3 and 9: the units of measurement for RMSE should be reported in the figure panel.

The unit for RMSE has been added in Figures 3 and 9.

• Figure 8: letters a) and b) are missing in the figure panels; i would invert the order of varieties 15 and 16 to make the trend (and comparison between measured and simulated data) clearer.

The order of the varieties and additions of a. and b. have been made to Figure 8 as suggested.

• Figure 9: it makes no sense to keep the axis scale in the range 0-1000 since the lowest value is over 400 gm-2. Please reshape the figure starting from 400 g m-2.

Fig. 9: The plot is now modified so that the y axis scale starts only from 400 g/m2.

• The instantaneous O3 effect is negligible. Could this be partly due to the recovery effect?

Line 771-776: Yes, it could be due to the recovery effect. We have added the following text 'We found a negligible effect of O_3 (0 to 0.2 %) on due to the instantaneous effect of O_3 on photosynthesis, which could perhaps be partly due to the to the crops ability to recover photosynthetic capacity overnight, compared to a highly significant (2.86 to 35.85 %) impact due to the long-term O_3 effect on carbon assimilation *via* the enhancement of senescence on final GrainDM.'

• The explanation of the coefficients of the LAI equations is missing and should be added to the paragraph: e.g. kd' and kb' are not defined.

Line 345-346: All explanations of the coefficients used in eqs 21 and 22 are now included in the text starting.

• The calibration was done manually; given the very high number of possible combinations of parameters and the presence of possible local minima, how can you be sure that the best set of parameters minimizing the error between observed and simulated data have been found? Please explain.

We have now added additional text explaining and justifying the rationale behind the selection of our calibration method to return close to the optimum set of parameterisations. We have included two additional references that support this approach.

Line 627-640: The manual calibration process consisted of three stages as explained above, as well as comparisons with established information on wheat growth from the literature. By reducing the number of parameters involved in the calibration, the chance of equifinality (multiple combinations of parameters yielding similar results) was minimised (Beven, 2006). The parameters identified by the sensitivity analysis were varied within realistic ranges to obtain a parameterization that closely approximates wheat physiological processes. Multiple parameterizations were tested to avoid convergence on local minima in R^2 and RMSE. While further fine-tuning of the parameter ranges could potentially improve yield prediction, it might also disrupt simulations of other key plant processes, such as carbon allocation or photosynthesis. The calibration approach balances the need for accurate output simulation with the physiological realism required for wheat growth under the conditions of this study. Though it is difficult to claim that the absolute optimal parameter set has been achieved, this limitation is common to any model calibration (Wallach, 2011). The current parameterisation represents a physiologically realistic simulation of wheat growth under the conditions of the present study using a robust calibration method.

References - Wallach, D.: Crop model calibration: A statistical perspective. *Agronomy Journal*, *103*(4), 1141–1153. https://doi.org/10.2134/agronj2010.0432, 2011.

Beven, K.: A manifesto for the equifinality thesis. *Journal of Hydrology*, *320*(1–2), 18–36. https://doi.org/10.1016/j.jhydrol.2005.07.007, 2006.

• Lines 887-889, page 31.The WOFOST model (Nguyen et al., 2024) should be included in the list of models being capable to simulate both the instantaneous and long-term O3 impact on wheat grain yield.

Line 935: Modified the sentence.

• Page 43. The reference “Nguyen et al., 2024” is incorrect. Please check.

Page 49: Corrected the reference

References

Dentener, F., Emberson, L., Galmarini, S., Cappelli, G., Irimescu, A., Mihailescu, D., Van Dingenen, R., van den Berg, M., 2020. Lower air pollution during COVID-19 lock-down: improving models and methods estimating ozone impacts on crops. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 378, 20200188. https://doi.org/10.1098/rsta.2020.0188

Nguyen, T.H., Cappelli, G.A., Emberson, L., Ignacio, G.F., Irimescu, A., Francesco, S., Fabrizio, G., Booth, N., Boldeanu, G., Bermejo, V., Bland, S., Frei, M., Ewert, F., Gaiser, T., 2024. Assessing the spatio-temporal tropospheric ozone and drought impacts on leaf growth and grain yield of wheat across Europe through crop modeling and remote sensing data. European Journal of Agronomy 153, 127052. https://doi.org/10.1016/j.eja.2023.127052