Reply to referee comment #1

This paper presents a modeling study to evaluate the impact of ozone damage on Chinese wheat yields using the developed DO3SE-Crop model and field experimental data. Overall, this is a comprehensive study that provides valuable insights into the effects of ozone on food security. However, there are areas that could be improved.

The title of the paper could be modified to reflect a broader scope of the model's purposes and applications, rather than focusing solely on a specific site in China.

This is a really good suggestion - thanks. We have now changed the title to 'Development of the DO_3SE-crop model to assess ozone effects on wheat crop phenology, biomass and yield'.

A primary concern is that only data from 2008 was used to train the model. This raises questions about the representativeness and robustness of the model when applied to other years and regions.

We trained the model on 2008 data, and then tested the model on unseen data from 2007 and 2009 specifically to test the ability of the model to capture variability between years. Since we found that the model was able to explain 7.9-8.7% (Line 881-884) of the variation we judge this to be a good indication that the model is indeed able to capture between year variability reasonably well. However, we agree that the application of the model to other regions is untested, due to the fact that data from other sites/regions are simply not available. We have made both the description of the calibration and evaluation of the model clearer by substantially re-writing section 2.2 and also make the issues associated with the model's regional representation clearer in the text (Line 929-934) suggesting that this requires future investigation as new datasets become available.

The extremely high correlation between the modeled and observed yields (R^2 of 0.99) suggests potential overfitting, which could lead to biased prediction results. The authors should address this issue to be more convincing.

Line 256-262 We have made an update to the model which now accumulates stomatal ozone flux (acc_{fst}) from only 200°C days before the simulated timing of anthesis. This update was made to accurately capture ozone uptake at a time when the canopy was fully established. Correlations between modelled and observed yields for this updated model are now an R^2 of 0.92 (n=4) and 0.68 (n=20) for the training and test data respectively. Since we would expect a higher correlation for the training dataset (since we calibrate to fewer data points) and since the test dataset shows a reasonable fit to a larger set of unseen data, we believe this provides the evidence that that model is not over-fitting. We make this clearer in the discussion (where previously we have only reported results for the trained data) so that we now say - Line 881-884 'Importantly, when applied to the test dataset (i.e. excluding 2008 data for the Y2 and Y16 cultivar) the model, was found to simulate the grain dry matter under ambient and elevated O_3 treatments to within 7.9-8.7% of the observed values (R^2 =0.68, 76 g/m^2 see Fig. 9)'

It appears that the ozone-induced yield losses derived from field observations are significantly larger than the corresponding simulation results. The reasons behind this discrepancy should be investigated and discussed.

We find that we tend to underestimate the O_3-induced relative yield loss (RYL) by between -2.76 and 15.34 (observed less modelled RYL) across all years and cultivars. Further analysis of these data (see new tables S2a. and b.) show that average differences between observed and modelled RYL

estimates for all cultivars are similar between years (ranging from 4.94 to 6.73) but that average differences between cultivars are more variable with the sensitive cultivars (Y2 and Y19) ranging between 5.02 and 9.0; and tolerant cultivars (Y16 and Y15) ranging between 2.66 and 5.54. This would suggest that O_3-induced yield losses can be more reliably modelled for tolerant cultivars which may suggest that additional processes causing O_3-induced yield losses in sensitive cultivars are not captured, such as O_3 altering the allocation of C to different plant parts (Feng et al., 2008) or O_3 inducing additional respiratory costs via the upregulation of defence mechanisms (Biswas et al., 2008).

To address this, we have added text and two tables in the supplementary (section S6; text and Table S2a. and S2b) giving the explanation above and have added a short paragraph summarising these findings to the end of the results section at Line 775-784 and the beginning of the Discussion section at Line 790-801.

In the methods section, the authors cite numerous related studies for the formulas of different modules. While the sensitivity of these empirical parameters has been evaluated, the uncertainties arising from the process need to be elaborated upon.

We believe this comment relates to the fact that although we focus attention on defining the values (and ranges) of key parameters, we implicitly assume that we have included all, or at least most, of the important model processes that determine O_3 effects on crop development, growth and yield. To address this, we include mention of the various ways in which O_3 can impact crop growth and yield (referencing a seminal paper by Feng et al. 2008 that performed a comprehensive meta-analysis for wheat) in the discussion to show that we are aware there are additional processes not included in DO_3SE-Crop that might influence crop responses to O_3 which could form the basis of future studies. We also refer to some of these 'missing' processes perhaps being responsible for the underestimation in RYLs, especially for sensitive cultivars (see point above)- please see Line 796-805.

We also appreciate that although we have identified the key parameters through application of a sensitivity assessment, we have not included any uncertainty analysis (e.g. the certainty with which we can define a range of values for each parameter) and how any uncertainty in these values and ranges may influence results. We note this in the methods section at Line 562-563 (in Section 2.2) but state it is outside the scope of this paper by including the following sentence.

Line 562-563 '*We note that assessing the probability distribution of these ranges would also be useful but consider this outside the scope of the current paper due largely to data limitations*.

Uncertainties may arise from assumptions, such as "We then assume that these values are consistent across cultivars and years" (Line 580). These need to be addressed in more detail.

We now make clear in the text that assumptions that parameterisation will work well for different cultivars and years need to be made with some degree of caution by including the following text at Line 623-627 in the section on crop phenology.

Line 628-632 '*We then assume that these values are consistent across cultivars and years. Figure 3 suggests this is a reasonable assumption since the phenology module captures the timing of*

anthesis and harvest for unseen cultivars and years within 2 to 4 days and 1 to 6 days respectively, of the observed timings. However, we appreciate that assuming these phenology parameters will work for a wider variety of cultivar types (e.g. early or late sown and/or maturing) and years with rather different meteorological conditions needs to be done with caution'.

The figures in the paper could be better presented to improve visualization and readability. For instance, the colors of different legends in Figure 3 are difficult to distinguish.

Line 683: This figure (now Figure 5) has been updated for better visualization and readability.

The large differences between the modeled and observed Anet and gO3 in Figure 4 need to be explained, as the model tends to underestimate these values. Additionally, correct the figure caption for panels c and d, as they appear to be swapped. Ensure consistency in font sizes, such as in Figure 5c and d.

We have provided an explanation for the discrepancy between the modelled and observed Anet and gO3 values in Figure 4 by including the following sentence at Line 651-663.

'In Fig. 4a and b, the steep decline in modelled A_{net} and g_{O3} is not seen in the observed dataset. This discrepancy may occur since the simulated A_{net} and g_{O3} values represent sunlit parts of the upper canopy which comprise both green and senesced leaf material. In contrast, observed A_{net} and g_{O3} values are measured specifically on the flag leaf and most likely for only the green parts of the leaf and most likely only for the green parts of the leaf, since the LI-6400 photosynthesis system mounted with a 6400–40 leaf chamber fluorometer (used to measure A_{net} and g_{O3} in the Xiaoji experiment, Feng et al., 2016) will not provide values for senesced leaf material. See also Figure 4 which combines A_{net} and g_{O3} with observed relative chlorophyll content and clearly shows the leaf is senescing as predicted by the model.'

Minor Comments

Correct typographical errors in the abstract, such as "...regions of ufor risk assessment."

Line 11-18 - Corrected the typo. This now reads: -

'A substantial body of empirical evidence exists to suggest that elevated O3 levels are causing significant impacts on wheat yields at sites representative of highly productive arable regions around the World. Here we extend the DO3SE model (designed to estimate total- and stomatal-O3 deposition of for risk assessment) to incorporate a coupled Anet-gsto model to estimate O3 uptake, an O3 damage module (that impacts instantaneous Anet and the timing and rate of senescence), and a crop phenology, carbon allocation, and growth model based on the JULES-Crop model. The model structure allows scaling from the leaf to the canopy to allow for multiple leaf populations and canopy layers'

In Line 50, O3 should be in subscript form.

Line 52 - Corrected the typo.

In Line 104, "Evaluation" should be in lower case.

Line 16 – Changed the heading.

Ensure references are complete and correctly formatted, including publication year, journal name, etc. For example, the reference to Yang, L. et al. should include all pertinent information.

All references have been checked for formatting.

Add missing parentheses in Figure 5b.

Line 715 - Parentheses added in the ‘now Fig 6’.