

Thank you again to the editor and reviewer for taking the time to improve our paper. We have addressed the comments below and finalised the manuscript.

The authors have addressed the comments of the reviewers reasonably thoroughly, and have strengthened the manuscript. While it is unfortunate that they are not in a position to address the model biases they highlight, the additional analysis of NO_x they have added provides useful new insight, and I feel that the paper is now suitable for publication in ACP.

Comments and remaining minor issues:

The abstract still states "Results from this study demonstrate the applicability of hourly output from UKESM1 for human and ecosystem health-based impact assessments" despite the large (factor of two) systematic bias. The results should not be used for any impact assessment without strong bias correction, so the statement is misleading and needs adjusting or removing. This point has been adequately addressed in the conclusions, but not in the abstract.

Thank you for pointing this out. We have amended the abstract:

'Results from this study ~~demonstrate the applicability of hourly output from~~ [reveal that hourly ozone concentrations from UKESM1](#) ~~for require bias correction before use for~~ human and ecosystem health-based impact assessments, ~~increase confidence in model projections, and highlight areas that would benefit from further observations.~~

The response argues that there is a decoupling of the ozone bias and DOR performance. While it is true that these features are decoupled, it doesn't necessarily follow that they are independent, and that the DOR is correct for the right reasons. However, I accept that the authors are not in a position to explore the mechanisms more deeply, and feel that the new analysis added in section 3.4 goes some way to addressing the issue.

Thank you for your comment here. We agree that the DOR is not necessarily correct for the right reasons but we hope we can encourage more researchers to consider reporting and investigating this metric.

With access to few reliable measurements, it is reasonable to argue that more are needed. However, I remain unconvinced that this would address any of the issues with the model raised in the manuscript. It will improve quantification of model biases, but provide no understanding of how these biases arise or a pathway to remove them. However, this may be a topic for future work.

We have adapted the text to avoid stating that our study shows more measurements are needed. Instead, we highlight that this study has brought together all these observations and therefore provides an overview of ozone in the tropics. E.g. '[In this study, we bring together in situ ozone observations from 13 sites across the tropics for the first time.](#)'

Figure 4: This max/min arrow plot remains difficult for the reader to interpret. A much neater solution would be a horizontal line starting at the minimum time and ending at the maximum time, thus highlighting the period of ozone increase. This would cross midnight for measurements at Bukit and Daintree.

Thank you for this suggestion, we have taken it on board and changed the figure.

Figure S7 provides a nice exploration of spatial variation and site sampling error, but it is difficult to make out which line/location is which. Please consider an alternative coloring scheme based on saturation rather than hue, with the saturation dependent on how well the model captures the observed ozone concentration or seasonality (e.g., darker colors represent the observations better).

This is a great idea! We have improved the figure.

The addition of the comparison with other models (Fig S10) is very helpful.

Thanks.

Line 79: Nascimento reference missing

Added.

Line 127: DOR is defined in the abstract, but has not yet been defined in the paper.

Thank you for pointing this out, the definition has now been added at first mention.

Line 196: The new subtitle "Figures and data" is not very informative. I do not think that subtitles 2.3.1 and 2.3.2 are needed.

We certainly agree. The subtitles have been removed.

Line 258: "high accuracy" is an overstatement given an RMSE of 5-7 ppb and r^2 of 0.64.

This has been changed to 'reasonably well'.

Line 391: "Further discussion.... in Section 4.3" This sentence refers forward in the text to a point the reader hasn't yet reached; either bring the discussion into this section, or drop the forward reference.

We have dropped the forward reference.

Line 394: The performance of other models is squeezed in in the final sentence of the section, and is out of place here. Given the later discussion of CMIP6 model performance (lines 435-442), the final sentence here can be removed.

This section identifies that Watukosek is badly represented by ukesm1. I think it is important to also point out here that other models represent the seasonal cycle much better. UKESM1 also struggles in Yangambi, but other models are in fact worse. This is especially important as the Yangambi data has not been made public until now so this multimodel comparison is novel. Although it is only one sentence right at the end, I think this is an appropriate place to make this comparison.

Line 434: as line 391 above, although this may be addressed by replacing "Sect 4.2 and 4.3" with "in the following sections".

We have changed the sentence as you suggest to avoid forward reference.