

Interactive comment on “Magnitude, Trends, and Impacts of Ambient Long-Term Ozone Exposure in the United States from 2000–2015” by Karl M. Seltzer et al.

Anonymous Referee #2

Received and published: 2 September 2019

General Comments

The manuscript titled “Magnitude, Trends, and Impacts of Ambient Long-Term Ozone Exposure in the United States from 2000-2015” by authors Seltzer et al., submitted for publication in Atmospheric Chemistry and Physics, discusses a topic of interest to the air quality and health impacts community. The study develops a machine learning approach to generate an observation-based estimate of long-term ozone exposure and calculates resulting human-health and crop yield impacts.

The technique the authors present is novel, the analysis is scientifically rigorous, and the results quantitatively reveal interesting insights regarding the various ozone expo-

C1

sure metrics in use. The paper represents a significant contribution within the scope of ACP, and the presentation quality is sufficiently clear, though some slight improvements can be made. I consider this manuscript well-suited for publication in ACP after the following minor comments are addressed.

Specific Comments

1) Further details regarding the ANN method and the calculations of human health and crop yield impacts are needed to be fully understood by the reader. Regarding the ANN method (Section 2.2):

-How was the chosen architecture decided upon? Was any testing performed to optimize the number of hidden layers or nodes to achieve the best performance while maintaining manageable computation time?

-A table listing all the input parameters to the ANN would be helpful. I think all inputs are probably stated in the paragraph starting at L115, but a table format would be much easier to digest by the reader. This could include a description of the variable, the data source, and units, such that the text in this paragraph could be greatly simplified. Relatedly, from what source are the monthly methane concentrations?

-A clearer description of the format of the input variables is needed. Specifically, how many measurement stations are included in the TOAR dataset used for training (if this differs by metric, perhaps provide one as an example)? What exactly constitutes a “single” input to the ANN? Is it a single site, with the meteorological conditions and emissions for that location and time? Or are 2-D or 3-D arrays of the inputs fed in, perhaps all sites for a particular hour, to provide spatial or temporal context? In particular, I was left wondering about the statement starting “Since emissions from East Asia. . .” at L124. How are Asian emissions input to the ANN such that the spatial separation/time lag between Asia and the U.S. site is conveyed? Is it conceivable that the ANN will recognize the heightened importance of Asian emissions for the U.S. West coast vs. the East coast?

C2

2) Regarding the calculation of human health and crop yield impacts (Section 2.5):

-Quite a bit of jargon is included here (e.g., GBD and ICD are undefined acronyms), and it is sometimes unclear where values (e.g., of the TMRELs in L213) were obtained. Are these taken from the literature, or from your analysis?

3) Finally, for the upstream impact metric, AOT40, it is stated in the discussion and conclusions that the 40 ppb O₃ threshold results in “a disproportionate influence of meteorological variability on its magnitude” (L321). I appreciate the need to adhere to convention/widely-used metrics, but could you also explore the effect of reducing this threshold? My understanding of O₃ human-health impacts is that there is no safe level, so it seems justifiable to explore crop-loss metrics with thresholds below 40 ppb as well.

Technical Corrections

L77: Not all of these acronyms are familiar; are there distinctions between the networks that the reader should be aware of? Urban vs. rural, difference in meas. technique, etc.? If so, please state.

L79 and throughout: inconsistencies with tense. (“The first. . . metric comes from. . .” vs. earlier “Daily O₃ observations. . . were retrieved from. . .”) Improving the consistency of past vs. present tense will improve the readability of the manuscript.

L90: It would help a reader unfamiliar with these metrics to clarify how the acronym is derived – “M12” meaning 12-hr mean? “AOT”. . . ?

L112: “. . .in each hidden layers. . .” should be “layer”

L141: “. . .step towards to. . .” should be “towards an”

L157: It's unclear why the CTM output is referred to as a “pseudo-observational dataset” in Section 2.3. Perhaps later (Section 3.1) when the CTM output is sampled at the monitoring locations, then this is warranted. But, when just describing the

C3

output from GEOS-Chem, it would be more accurately described as “model output.”

L160: Why are these particular years chosen?

L167: For the scaling method used to get from the first GEOS-Chem pressure level to the surface, do you mean the “aerodynamic resistances to dry deposition” described in Zhang et al. (2012)? Found using their eq. 5? Might want to call this out by name, since they don't refer to this as “scaling” (doing ctrl-f for “scaling” in that article finds a use of scaling factors different from this purpose).

L181: “This evaluations. . .” double use of “evaluate” in this sentence.

L182: “We then applied this framework. . .” This is a bit vague; please be more specific about what is done to arrive at the “Met. Adjusted” results shown in the Figs.

L232: I'm generally accustomed to reading a description of the figure in the text prior to getting into the discussion about it. I suppose this is a matter of writing style, though, and leave it up to the authors whether to include more explicit descriptions of this and forthcoming figures in the main text.

L238: The second factor contributing to deviations, accumulation of errors given that AOT is a cumulative index, makes sense for the green vs. yellow lines where deviations grow later in the time series. But, this doesn't describe the red vs. blue line deviations “early in the time series,” described above (L232). Any idea why those errors, before adjusting for meteorology, would appear largest early in the time series?

L276: Could you elaborate on what you mean by “different” – what would you expect if you excluded the urban core stations, e.g.?

L305: Perhaps give the numbers for the mean trends following adjustment? Use of “marginally” implies a small effect, but “played a role” implies some (unspecified) effect. The reader may be left wondering: did the met. factors play a large or small role?

L311: “were reaching” implies, to me, mixing ratios were increasing to this amount, but

C4

I think you instead mean that the values were decreasing over time, correct? Perhaps change this wording, to make that clear.

L321: "This is due to the sensitivity of particular meteorological variables on extreme O3 episodes. . . ." This sentence doesn't quite make sense, please clarify.

L349: The evolving differences between the T2016 and J2009 metrics is a very interesting result; do you care to expound on which metric might be more valid for certain applications? Or if there are alterations to the current metrics that you would suggest?

L408: "months" should be "monthly"

L415: semicolon should be replaced by a comma

Figure 1: Please state what the dashed yellow/green lines in all panels represent (linear fit?)

Figure 5: In case one wanted to compare the top and bottom panels (M12 Maize vs. Soybean, AOT40 Soybean vs. Wheat), consider using consistent y-axis ranges. Also, is there a reason the time series for AOT40 Maize is not shown?

Interactive comment on Atmos. Chem. Phys. Discuss., <https://doi.org/10.5194/acp-2019-672>, 2019.