Atmos. Chem. Phys. Discuss.,
https://doi.org/10.5194/acp-2019-753-RC2, 2019

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Improving the prediction of an atmospheric chemistry transport model using gradient boosted regression trees" *by* Peter D. Ivatt and Mathew J. Evans

Anonymous Referee #1

Received and published: 3 November 2019

This paper describes the first application of a decision tree algorithm to predict the bias in an atmospheric chemistry transport model. It describes the approach taken, demonstrates its successful use for tropospheric ozone, and investigates its sensitivity to choice of training data and other options. The study is innovative in approach, and will certainly catalyse new research in application of machine learning techniques to atmospheric models. The paper is concise and well-structured, although there are some weaknesses in written style and presentation that need to be cleaned up. Overall the paper is appropriate for publication in ACP once the existing deficiencies have been addressed.

Printer-friendly version

Discussion paper

Specific Comments

Some drivers of model bias are likely to be non-local, particularly for longer-lived variables such as ozone. How might this be addressed within the framework of the current approach? Other biases will be due to limitations in how representative observations are of the scales resolved in the model (the comparison with coastal sites over Japan demonstrates this). How could these be addressed? It is notable that the observation data have already been filtered for urban and mountain sites. How sensitive are the results to the choice of which types of environment to exclude (distinct from the choice of location, which is covered well in section 7)?

l.59: "standard emissions configuration" It would help the reader to define standard (e.g., by citing a reference), or to drop this phrase if the subsequent list of inventories effectively covers it.

l.87: The phrase "most typical of" is unclear here, as flight measurements and sondes are different. The characteristics and sampling of the observations were similar?

l.98: "We have tended to favour" suggests prior work in this area, in which case it should be cited.

l.113: How is an "adequate level of complexity" defined for a particular problem?

l.115: A clause or sentence is required to define what "more interpretable" means in this context

l.157: The inability of the approach to correct the bias in some circumstances might also suggest that the parameters used for training are incomplete, and that some sources of bias are non-local.

l.191: The gain associated with a particular variable is not clearly defined here, and the feature importance (used on the y-axis in Fig 7) is undefined. Please add a sentence here to explain how these concepts are derived.

l.260: This final sentence presents an important but speculative conclusion, and would be much stronger if some evidence is provided from the study to back it up. The preceding analysis is vague about the reasons for the differences observed, so at least one concrete example of how model failings could be explained is needed.

l.264: How is the direct prediction implemented? Is model ozone used as one of the variables? If so, it appears unsurprising that "ozone + predicted (bias)" differs little from "predicted (ozone + bias)", and the main interest is therefore in what the differences tell us about the robustness of the decision algorithm used.

l.286: "Are all the variables needed?" It should be clear from the gain analysis shown in Fig 7 which variables are unimportant. If it isn't, why not, and if it is, what do you find?

l.294: The importance of particular variables is indicative not causative, and this would make it very difficult to extract information on the reasons for model biases. The nighttime bias of the model is clear from a simple comparison with observations alone (Fig 2), and while identification of NO3 as important provides good evidence that the approach successfully recognises this, it is not clear that it is possible to reverse this process.

Numerical precision: one decimal place is sufficient for the biases, and two decimal places for the R values. Greater precision is not warranted in the abstract or body of the text.

The word "significantly" is used frequently in a colloquial sense rather than a statistical sense (e.g., l.174, l.177). For clarity, please use a different word or provide statistical metrics where appropriate.

Typos and minor corrections

l.14: shows -> show

l.30: "etc" better explained or removed

l.102: multi-variant -> multivariate

l.112: "this this"

l.115: underlies -> underlie

l.121: "5 kfold" -> "5 k-fold" or perhaps just "5-fold"

l.144: calculated the -> calculated with the

l.157: the reality -> reality

l.212: dot -> dots

Interactive comment on Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2019-753, 2019.

Printer-friendly version

Discussion paper