Geoscientific
Model Development
Discussions

Open Access

EGU

Interactive comment on "Machine learning models to replicate large-eddy simulations of air pollutant concentrations along boulevard-type streets" *by* Moritz Lange et al.

Anonymous Referee #1

General Comments

Lange and coauthors do an admirable job of presenting a difficult concept to make clear, that is, their algorithm is essentially a model of a model. The study makes clear the benefits of such an absurd sounding task, that the CFD models in question are computationally very expensive. And so the authors present a much faster algorithm which produces similar results, with the intent of providing an easy-access tool for first-pass examination of urban planning on local-scale air quality. The authors were very thorough in their assessment of the available algorithms and justify their choices. There are a lot of open questions over how to improve the model performance in the street

canyon, and these are only highlighted by the Drift detection discussion, but I believe these can be left to future work. Overall I think the manuscript is worthy of publication.

Specific Comments

I have one minor technical concern: the use of a scaling variable feels a bit suspect on first reading. After a closer reading, it feels to me like a presentation problem. That is, the scaling variable is presented as something apart from the model when it really is a part of the model, since it will be needed to make future predictions. I would prefer to see a discussion of how the scaling variable is chosen for the cross-validation / evaluation cases: is it simply the value found in the minimization of RMSE on the training data carried forward, or are the authors recomputing the scaling variable for the evaluation data? I suspect the former, and I would like to see that confirmed in the methods section 2.5. If the latter, then there needs to be more description of this process and how the authors intend to compute the multiplicative factor a for novel datasets.

Technical Corrections

The paper is very well written and there were no glaring grammatical errors.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2020-200, 2020.