Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2018-263-RC2, 2019

Geoscientific Model Development Discussions

Interactive comment on "Calculating the turbulent fluxes in the atmospheric surface layer with neural networks" *by* Lukas Hubert Leufen and Gerd Schädler

Anonymous Referee #2

Received and published: 28 January 2019

Review of: Calculating the turbulent fluxes in the atmospheric surface layer with neural networks ==========

The paper describes a boundary-condition for specifying turbulence fluxes in climate models, as a function of readily available features. The BC is entirely empirically specified: an ANN is trained from experimentally measured data, and the resulting predictions are compared with those of a standard model (MOST).

The paper is generally well written, and addresses the potentially interesting and important topic of wall-modelling in complex flows, but suffers from a lack of analysis of the data and performance of the numerical method proposed. Coarsely speaking this

work fits the pattern of: (i) apply machine-learning to data-set, (ii) report the fit. Little insight into the data, physics, or performance of the algorithms is gained by the reader - and the authors' own results suggest a much simpler model would fit their data equally well. A path they do not investigate. For these reasons I recommend rejection.

Major comments:

- The most serious shortcoming, is that of more-or-less uncritically applying ANNs to the data-set, without examining their suitability, and to what extent the data can be explained by simpler models. In particular Figure 4 shows performance of your ANNs which is almost independent of number of neurons. In fact you show a 1-layer, 1-neuron "network" performs basically the same as 1-layer, 12-neurons, or a deep network with 2-layers with 7-7 neurons (7 inputs). This result strongly suggests that almost all the predictive power of ANN for this data is contained in a linear fit. At the very most a linear-fit-plus-bounds would have essentially the same predictive power (given your use of the tanh activation function).

Given this, it seems redundant and unnecessarily complicated to use the heavy-machinery of ANNs, with its associated costs: obfuscation of the functional relationship, expense of evaluating the network, and lack of statistical/noise modelling.

Indeed, data that can be reproduced with a single perceptron strongly suggests an extremely simple main relationship between features and output, which is an opportunity to discover simple physical relationships and main-effect parameters.

Note that Figure 3 is not a defense against these criticisms - as the authors themselves state, the relevant plot for the usefulness of the ANN in climate models is Figure 4, not Figure 3. Indeed the difference between these figures indicates that the more complex networks are overfitting the data from the available towers.

In my opinion this paper should not be published without an analysis of the data and a comparison with simpler models. Data analysis could include:

o Sobol index analysis to identify effects of individual features and coupled features on the output (ANOVA). o Active-subspace analysis to identify main-directions in the input space that contribute most to the output (these will most likely correspond to the weights in your 1-layer, 1-neuron network). o Correlation analysis on the input-space, are inputs independent? This and last two points will contribute to dimensional reduction. o Parametric/Non-parametric estimation of noise in the output.

Simpler models could include:

o Linear regression, with a variety of noise models. o Linear-regression-plus-bounds (only if the above fails) o Input variable elimination (via ANOVA) prior to linear regression/ANN. o Low-order polynomial fits/gene-expression programming to obtain simple explicit expressions capturing the functional relationship.

Only if ANNs do significantly better than linear models is the current work worth publishing.

- I'm not convinced by the assertion that there is a significant computational speed advantage to be gained by replacing MOST with an ANN. I'm not familiar with global climate models, but I see the intent to use it as a BC in an 3d LES simulation. In similar simulations in engineering problems, wall-modelled BCs (e.g. involving solution of an ODE at the wall) account for a negligible component of the total CPU time, and never more than $\sim$5%, with most time spent in the volume. Please explain what is special about your models that causes this situation to be reversed. Please quantify the time spend by your code in various parts of the calculation, so the reader can see the relation of the ground-modelling to other time-consuming parts of the code.

- Assumptions: Repeated reference is made to the assumptions made in the derivation of MOST. I would appreciate in Section 2 an enumeration of all assumptions made, perhaps with some comment on their validity and their role in simplifying the MOST model.

- Relatedly, the training data-sets contain the turbulence fluxes. Are these fluxes measured directly, or are they computed from measurements of u* and theta*, or is a more complex model used to map from measured quantities to turbulent fluxes. What assumptions are made in this map? What modelling assumptions are inherent to your ANN approach?

- Title should be "turbulence fluxes" not "turbulent fluxes". They are fluxes-of-turbulence, not fluxes-which-are-turbulent.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-263, 2018.