Nonlin. Processes Geophys. Discuss.,
https://doi.org/10.5194/npg-2019-26-SC1, 2019

Nonlinear Processes in Geophysics
Discussions
Open Access
EGU

Interactive comment on "Simulating model uncertainty of subgrid-scale processes by sampling model errors at convective scales" *by* Michiel Van Ginderachter et al.

Michael Tsyrulnikov

michael.tsyrulnikov@gmail.com

Received and published: 28 June 2019

In general, this is an interesting paper, which tackles the very hard "model-error problem" and has many high-quality elements.

What I like in the manuscript?

1. The careful examination of the spinup process when comparing tendencies from the two models. Without eliminating or radically reducing the initial transient there is no chance to detect the (very small in one time step) model error.

2. The flux formulation, which guarantees conservation of the spatial mean values of the perturbed fields. (I would note that the perturbations of the fluxes need to vanish at the boundaries to ensure the conservation.)
3. The careful way the perturbations from the archive are added to the model fluxes.
4. It's interesting that the error appears to be additive rather than multiplicative (section 2.3.5).

What don't I understand or don't fully agree with in the manuscript?

1. Actually, the authors study an artificially introduced model error. They do so by switching off the deep convection parameterization and examining the differences in the tendencies between the model-*with*-the-parameterization (the "truth") and the model-*without*-the-parameterization (the "target"). In other words, they study the error in the model-*without*-the-parameterization (with respect to the model-*with*-the-parameterization). In practice, however, we are interested in studying model errors/uncertainties *due to* a parameterization, not uncertainties due to the absence of a parameterization. Here I agree with Referee #1. Correspondingly, in practice, model-error perturbations are to be added to the (operational) model *with* the parameterization.

   However, I think this doesn't make the results presented in the manuscript irrelevant. They are still interesting – if regarded as an example of how an uncertainty in the model can look and how can it be modeled.

2. Essentially, the proposed technique is a substitute for a deep convection parameterization. The technique needs to be trained on a model with the parameterization or, as suggested by Referee #1, on a fine-grid model. The results (section 3) show that the substitute is quite successful. This is interesting by itself, but is

this approach practical? I think, it can be practical in the case when a physics parameterization is absent and the technique is trained on a high-resolution model.

3. The lack of spinup in Fig.3 (top) is remarkable. But it is shown for the *domain average* water vapor flux (so that there is no spinup in the *bias*). How about the **RMS** flux error: is it still small at the first steps after the deep convection parameterization is switched off? (I mean, switching off a parameterization should change the model attractor so that a transition of the model state from the old attractor to the new one is expected to take place, which should manifest itself as an adjustment process.)

4. The perturbation generation scheme is meaningful (as I mentioned above) but seems ad-hoc. In particular, it is unclear why different criteria are selected in sampling the archive and in perturbing the fluxes ($M_{\mathsf{cut-off}}$ versus OMEGA/MOCON). To justify the technique, you could require that the perturbations are sampled from the same conditional distribution of the error given the current state of the column in question.

Specific comments

1. Fig.4. Given the limited number of cases (72) and the very large standard deviations as compared to the mean values, the mean values seem to be not statistically significant. Even if we assume that the cases are independent and the distribution is Gaussian, the significance level can be easily seen to be well above the acceptable level of 0.05.

2. Figs. 3 and 4 display results for the *mean* fluxes. Figs. 5–7 and Table 1, however, seem to show results for individual profiles. I think it should be clarified which conclusions you draw from the mean-flux results and why the mean-flux results are relevant at all? I mean, in the perturbation generation scheme, you

seem to take individual profiles. But the distributions of individual profiles and the distributions of their domain averages should be very different.

3. I don't understand how the vertical and cross-variable correlations as well as the probability densities you have estimated are used in the perturbation generation scheme? Sampling profiles from an archive does not require any knowledge of their probability distributions. How do you actually use the distributions? From the manuscript, it seems that these are not used.

Minor comments

1. It is not everywhere clear at which level the fluxes are evaluated.

2. Fig.5. "In blue the fitted exponential distribution."

   I don't see anything blue in Fig.5.

3. P.11–12. "while the correlation coefficient of the flux error of meridional wind w.r.t. the other errors is close to zero. This is because, contrary to the other variables, meridional momentum flux error doesn't have a dominant sign."

   Actually, the correlation coefficient is insensitive to biases.

I am sure the paper will be improved and eventually published!

Interactive comment on Nonlin. Processes Geophys. Discuss., https://doi.org/10.5194/npg-2019-26, 2019.