Geosci. Model Dev. Discuss., 2, C423–C429, 2009
www.geosci-model-dev-discuss.net/2/C423/2009/

Geoscientific Model Development Discussions

Interactive comment on "Automatic generation of large ensembles for air quality forecasting using the Polyphemus system" *by* D. Garaud and V. Mallet

D. Garaud and V. Mallet

damien.garaud@inria.fr

Received and published: 6 November 2009

1 General Comments

"The paper summarizes an innovative approach to running an air quality ensemble system using an automatic ensemble member generation. The system allows for uncertainties in emission data height, a large source for uncertainty, as well as chemistry and deposition and chemistry model physical parameterizations. The system appears to be designed carefully and using flexible modules that allow for easier system configuration."

We thank the reviewer for her or his comments.

1.1 Meteo Ensemble

“Meteorological uncertainties play a very large role as shown in previous studies. The authors chose to include this by randomly perturbing the input meteorology. I recommend describing why this approach is better than using inputs from members of a met. ensemble system. The latter approach would ensure mass continuity that is not guaranteed with the random perturbation approach. In addition, various vertical diffusion and cloud fraction schemes are allowed. These uncertainties could also be included from the input from a met. ensemble directly. Therefore, tests driven by a meteorological ensemble would provide a useful alternative to the approach shown that may produce inconsistencies.”

We totally agree. We plan to work on the integration of meteorological ensembles. We added a question about this point in the conclusion.

However the mass continuity is already ensured: in the numerical model, the vertical wind is diagnosed from the horizontal wind.

1.2 Emissions

“Emission magnitude certainties are not included directly, thereby ignoring a large source of uncertainty.”

The uncertainty in the emission magnitudes is taken into account. See the options 26, 27 and 28: the emissions of three main categories (biogenic, NOx and VOCs) are perturbed independently of the vertical distribution (option 8).

GMDD

2, C423–C429, 2009

Interactive Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

1.3 Ensemble Assessment

"Finally, the system should be evaluated with probabilistic matrices as uncertainty information is one of the most useful outputs from an ensemble system. It is recommended that reliability diagrams and ranked histograms be included to assess the system's reliability potential."

The purpose of the paper is to describe our approach to generate large ensembles that take into account all uncertainty sources. To make this clearer, we added further perspective in the abstract, the introduction and the conclusion.

We decided not to show probabilistic indicators. The first reason is that the ensemble is only roughly calibrated, as we now mention in the paper:

"a few ensembles were generated in order to roughly calibrate the uncertainty parameters, based on comparisons with observations (not reported here)."

The second reason is that we would need a much longer paper to deal with the probabilistic evaluation properly. We decided to explain in details the generation procedure, so that we can refer to this paper later. We added an analysis of the structure of the ensemble, to give insights in the outcomes of the method and to demonstrate what variability can be obtained.

We agree the probabilistic evaluation is a key point. We are currently writing a (less technical) paper on the probabilistic evaluation and the calibration of the ensemble.

2 Specific Comments

- Eqtn 2: *define "n".* It is defined below the equation. It is the upward-oriented normal to the ground.

GMDD
2, C423–C429, 2009
Interactive Comment

- Eqtn 7: *Von Kármán constant.* Fixed.
- Pg. 896, sect. 2.1.7: *how are low and medium distributions for SNAP determined?* These distributions are based on the EMEP distributions (for their own vertical distribution) and our default distributions.
- Pg. 898, section 2.2.2: *Why was 40 m chosen as the lowest model layer? Some chemical models have shown that lower layers improve performance.* We have not noticed this for simulations at European scale. After your comment, we may add 30 m as an alternative height, when we generate a new ensemble.
- Pg. 898, section 2.2.3: *How is mass continuity assured here but winds are perturbed in section 3. How different are these options than ingesting vertical motions directly from the met model and then adjusting for mass-conservation?* Only the horizontal wind is perturbed, and the vertical wind is diagnosed afterward. We modified the text to make this clear. In practice, we never use the vertical winds of the meteorological model (even when no ensemble modeling is involved). The chemistry-transport model has its own advection scheme and its mesh, so we need to recompute the vertical wind. This point has been studied in the literature, including on Polyphemus platform Sportisse et al. (2007).
- Pg. 899, Section 3.1: *Aren't some fields being perturbed twice therefore possibly over-weighting their contribution. For example, winds are perturbed here but different mixing schemes are also used that are driven by winds? Can you provide statistics from unperturbed vs perturbed members to evaluate the benefit of these perturbations, if any?* The fields are perturbed independently (before the numerical time integration) so as to avoid what is pointed out. For example, the vertical diffusion coefficient is computed with unperturbed winds, and then perturbed. As we explain in section 3.2, first we carry out the preprocessing with the computation of all input fields, and then we perturb the fields before the numerical time integration.

GMDD

2, C423–C429, 2009

Interactive Comment

- Pg. 906: line 20: *May be an unfair comparison as single realization will have higher resolution and more sophisticated physics than any ensemble member.* No, the reference members are the most sophisticated and the best tuned (from experience) models. As we explain in section 4.2.3, it is noteworthy that the best model (in terms of RMSE) does not include all the best parameterizations.

- Pg. 908, Fig. 5: *Any reason that reference model 5 has very low ozone in the north compared to the ensemble member? Is the ensemble solution even a remote possibility?* Yes, it is fairly possible, at least for certain periods. This is due to high NOx emissions, especially over U.K., that titrate ozone.

- Pg. 908, Fig. 9: *The night-time ozone range seems beyond the realm of possibility? System configuration evaluation using Talagrand and equally likely diagrams would help to identify members that are not realistic and should be removed.* We worked on that issue. It happens that the uncertainty is underestimated: there are many observations outside the envelop of the ensemble. Therefore the range may not be large enough. This may seem surprising at first, but it is fairly possible because it is difficult to model ozone in the night. It is consistent with the high RMSEs we observe on this target. Such a range can also be seen on other ensembles (not generated with Polyphemus).

- Pg. 908, Fig. 10: *A plot of probability of ozone exceedencing a certain value for the summer and reliability diagrams are crucial for understanding the utility of the system for probabilistic predictions.* We agree. But in this paper, we do not aim at proving the utility of our ensemble for probabilistic predictions. We aim at explaining the approach and at giving insights in the structure of the generated ensemble. Estimation of the uncertainties (of the target concentrations) and probabilistic forecasts are two big and complex topics that we decided not to address directly here. We understand this is frustrating, but introducing properly these topics would require much more space than we can afford. Another

GMDD

2, C423–C429, 2009

Interactive Comment

paper is being written on the subject, and of course, it includes all the adequate indicators mentioned in the review.

- Table 1: *Describe how ranges of variables like time step were determined. Can some sensitivity results be performed to show the relative importance of these quantities (e.g.: Min Kz)?* The ranges are often based on experience. For instance, the minimal values for the vertical diffusion coefficient are used (or have been used) by different models in the community, and these gave good results in previous Polyphemus simulations too. The values for the exponent p in the Troen & Mahrt parameterization are suggested in Troen and Mahrt (1986), and they both gave good results. As for the sensitivity study, one may refer to Mallet and Sportisse (2006b) which covers many options.

- Table 6: *Can this table be plotted as a time-series instead to show little structure.* We do not understand what is suggested here. The table is a reference for the global performance of the models; it is provided for the reader to see the performance range, the performance of the cited models, ... There is no time series involved. If it is suggested that we plot in four figures the four indicators against time, the figures will not be easy to read because of the number of models and the total number of timesteps.

References

Sportisse, B., Quélo, D., and Mallet, V.: Impact of mass consistency errors for atmospheric dispersion, Atmospheric Environment, 41, 6;132–6;142, 2007.

Troen, I. and Mahrt, L.: A simple model of the atmospheric boundary layer; sensitivity to surface evaporation, Boundary-Layer Meteorology, 37, 129–148, 1986.

Mallet, V. and Sportisse, B.: Uncertainty in a chemistry-transport model due to physical parameterizations and numerical approximations: An ensemble approach applied to ozone modeling, Journal of Geophysical Research, 111, doi:10.1029/2005JD006149, 2006b.

GMDD

2, C423–C429, 2009

Interactive Comment

Interactive comment on Geosci. Model Dev. Discuss., 2, 889, 2009.

GMDD

2, C423–C429, 2009

Interactive Comment