Correction of approximation errors with random forests applied to modelling of aerosol first indirect effect

A. Lipponen[1], V. Kolehmainen[1], S. Romakkaniemi[1] and H. Kokkola[2]

[1]Department of Applied Physics, University of Eastern Finland, P.O.Box 1672, FI-70211 Kuopio, Finland
[2]Finnish Meteorological Institute, Kuopio unit, P.O.Box 1627, FI-70211 Kuopio, Finland
Keywords: cloud, model, error, correction.
Presenting author email: Antti.Lipponen@uef.fi

In numerical simulations of atmospheric models, often a reduced model has to be used due to computation time or resource limitations. The use of a reduced model, however, induces errors to the simulation results. These errors are referred to as the approximation errors.

In this study, we consider the approximation errors caused by coarse discretizations of aerosol particle number and volume size distributions in a sectional aerosol model such as SALSA (Kokkola et al., 2008). We model the approximation error as an additive noise process in the simulation model and employ the random forest (Breiman, 2001) regression algorithm for constructing a computationally low cost predictor for the approximation error. In this way, the overall simulation problem is decomposed into two separate and computationally efficient simulation problems: solution of the reduced model and prediction of the approximation error realization. Given the trained random forest model, the accurate model can be approximated by the sum of the reduced model and the predicted approximation error realization.

The proposed approach is tested for handling the approximation errors due to a reduced coarse sectional representation of aerosol size distribution in the cloud droplet activation parameterization by Abdul-Razzak and Ghan (2002). Here a set of approximation error realizations between the accurate and reduced cloud droplet formation parametrizations is first constructed. Second, these approximation error realizations and the corresponding input variables of the parametrizations are used as the training data for the random forest algorithm and the random forest predictor is constructed and used in the simulations.

In this study, the approximative and sufficiently accurate parametrizations consisted of 7 and 70 size sections for the aerosol particle number and volume size distributions, respectively. The cloud droplet number concentration (CDNC) samples simulated with the accurate parametrization as functions of the CDNC samples computed both with the approximative and approximation error corrected parametrizations are shown in Figure 1. The mean squared and mean relative errors of the CDNC samples corresponding to the approximative and approximation error corrected parametrizations are shown in Table 1. The results show a significant improvement in the accuracy of the simulation compared to the conventional simulation with a reduced model. The approach proposed in this study is rather general and the extension of it to different physical simulation models is a straightforward task.

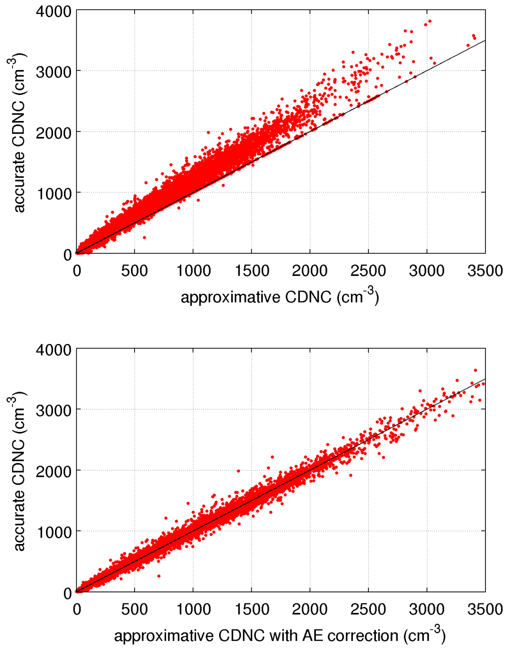

Figure 1. The cloud droplet number concentrations (CDNC) computed with the accurate parametrization as functions of approximative (up) and corrected (down) CDNC values. The solid black lines in the figures represent the identity lines.

Table 1. Mean squared errors and mean relative errors of the approximative and corrected CDNC values.

Parametrization	Mean squared error (cm^{-6})	Mean relative error (%)
Approximative	1.70×10^4	20.3
Approximation error corrected	1.27×10^3	8.4

This work was supported by the Academy of Finland (project 119270 and Centre of Excellence programs 1118615 and 250215) and by the strategic funding of the University of Eastern Finland.

Abdul-Razzak, H. and Ghan, S.J. (2002) *J. Geophys. Res.* **107**, doi:10.1029/2001JD000483.

Breiman, L. (2001) *Machine Learning* **45**, 5-32.

Kokkola, H., Korhonen, H., Lehtinen,, K.E.J., Makkonen, R., Asmi, A., Järvenoja, S., Anttila, T., Partanen, A.-I., Kulmala, M., Järvinen, H., Laaksonen, A., and Kerminen, V.-M. (2008) Atmos. Chem. Phys. 8, 2469-2483.