

Project: 1163
Project title: C2Phase: Closure of the Cloud Phase
Principal investigator: Corinna Hoose
Report period: 2025-07-01 to 2026-06-30

Subproject A: Statistical emulation of the aerosol effects on convective precipitation for an Earth System Model

Aerosol-cloud interactions are the most uncertain aspect in global warming, so we want to improve their representation in climate models. In detail, we want to include the aerosol effect on precipitation into the convective parameterization of a coarse grid model.

In the last years report we presented the extraction of the training data and an emulator for predicting the amount of precipitation itself (see Figure 1). We also showed that the prediction of the aerosol effect is not possible with only one set of experiments and thus, we requested extra computation time to enable the double run strategy: Run every experiment twice, once with a specific amount of aerosol and once with this amount halved to get the relative effect of aerosol in the experiments. This was working quite well even though the prediction remains a complex problem. As shown in Figure 2 the prediction with only one set of experiments is nearly impossible (blue dots – no skill visible), this is getting much better with the double run strategy (red and black dots – still not at the grey 1:1 line but much closer). In addition to the second set of experiments, we further improved the selection of input features for the training (black dots only). Now the R^2 score for predicting the relative effect of CCN on convective precipitation is ~ 0.25 . This means that our model is a lot better than just assuming the mean effect of CCN. Nevertheless in theory a perfect model would have a score of one. We assume that the low values are due to the high complexity of convective systems and the prediction will not improve any further. This is linked to the fact, that the CCN effect on convection was observed to vary from a suppression of -93% to an enhancement of 700% (Tao et al (2012), *Rev. Geophys.*, DOI 10.1029/2011RG000369). Surprisingly, linear models like Linear Regression and Huber Regression are performing better than more complex models like Random Forest, XGB and Histogram Gradient Boosting, so we focused on these simpler models.

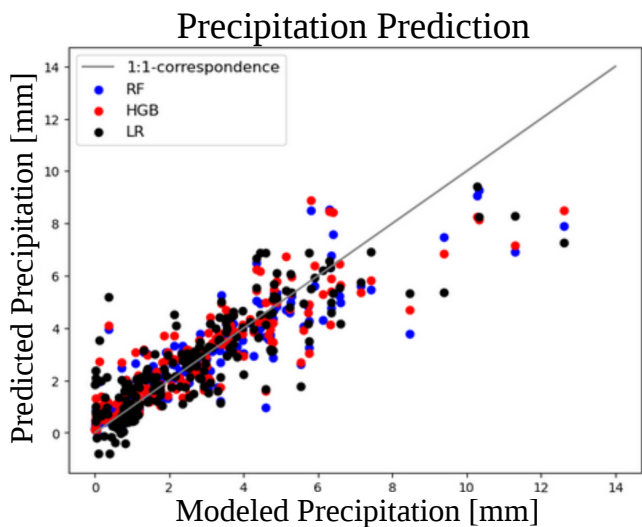


Figure 1: Leave-one-out cross validation for different ML architectures (Random Forest, Histogram Gradient Boosting and Linear Regression) predicting Precipitation (but not the CCN impact).

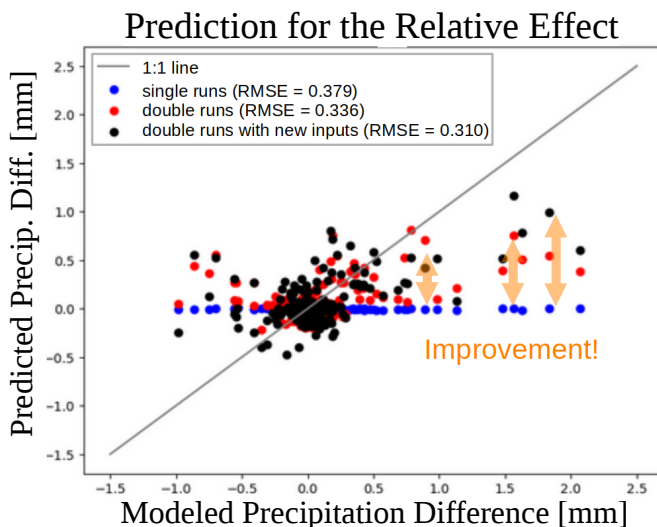


Figure 2: Cross-Validation for the different methods emulating the CCN effect. The improvement from the old method (blue dots), where only one set of experiments is used to calculate the CCN effect, is very visible. Even through the improvement in the RMSE is not so big, it can be seen that for the old method the predicted effect is always almost zero, independent of the modeled effect. But with the double run method, this is no longer the case. Taking some inputs into account further improved the prediction.

In addition to finalizing the emulator, we were working on the implementation of the emulated effect into the convective parameterisation of a coarse resolution model. The (very preliminary) results for the online coupled version are presented in the request document, since this work is ongoing. Here we will show results from the offline coupled emulator as first step of applying it to the global model. For this we calculated the input fields for the emulator from the output of a global model run. An example is given in Figure 3.

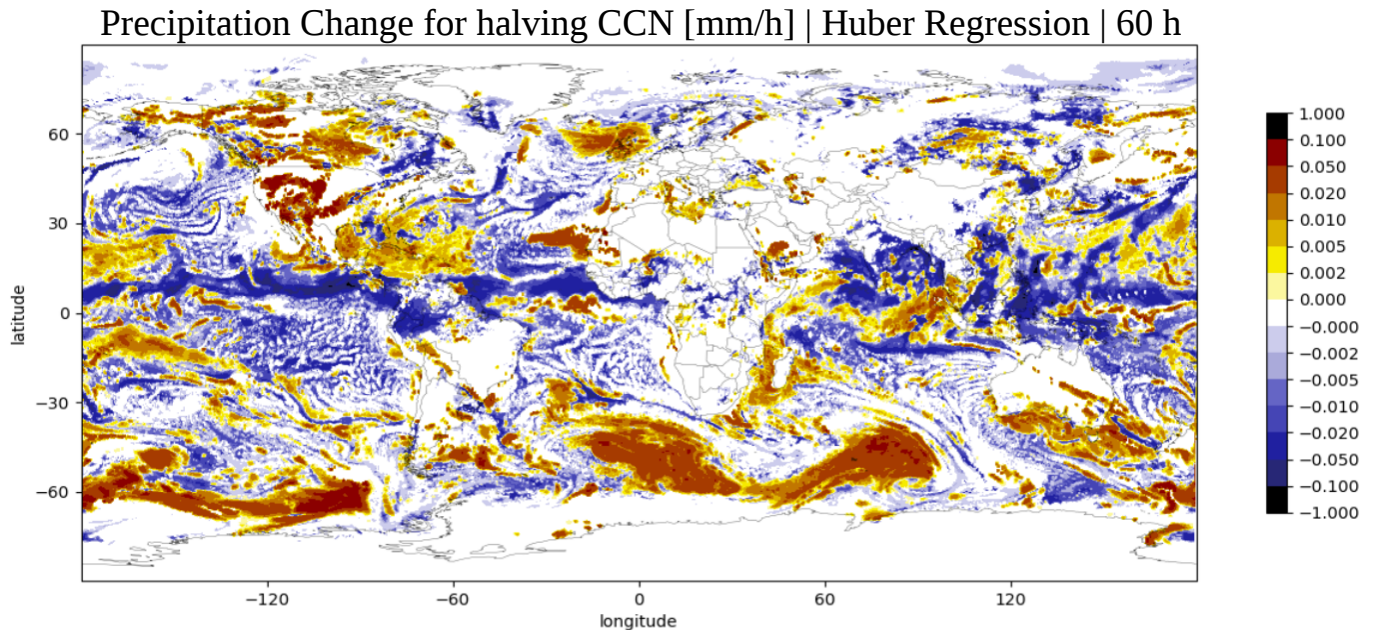


Figure 3: Spatial distribution (snapshot after 60 h) of the emulated effect in the global setting for a Huber Regression (with performs slightly better than the Linear Regression). The blue colors in the tropics indicate less precipitation for less CCN, which is a positive feedback (more CCN \leftrightarrow more precipitation). The red colors over southern ocean corresponds to a negative feedback (more CCN \leftrightarrow less precipitation).

For subprojects B and C, we further investigated impacts of secondary ice production (SIP) processes on the dynamics and microphysics of convective clouds in different aerosol environments. For this, we simulated multiple convective cloud cases found in aircraft/satellite and ground based observations across the globe using the ICON-NWP model with the 2 moment microphysics scheme, with nested (R3B7 to R3B10) grids. The SIP processes implemented in ICON include: 1) the Hallett-Mossop rime-splintering process, 2) raindrop freezing and shattering, 3) ice-ice collisions, and 4) sublimation of snow and graupel. The results are discussed in an article which is currently under review in Atmospheric Chemistry and Physics (Waman et al. 2026; <https://doi.org/10.5194/egusphere-2025-6129>).