Correspondence to editor’s comments

Thank you for constructive comments, for which responses are given with the relevant part of the revised manuscript. In the revised manuscript, changes are marked in blue.

The revised version shows several improvements, in particular when it comes to addition of clarity, both in terms of writing as well as presenting details of the developed model.

However, one of the main criticisms pointed out by myself, and the other reviewer (in their words “There is no proof of the generalisation capability of the model, so it may well be that this model fails completely if it were applied to measurement data obtained under different conditions.“) is still not addressed at all. That is to say, the authors have developed a model (which is fine), but there is no way of knowing how well it generally performes. Since the idea is to develop a model for others to use (of the shelf), it should be made very precise what are the capabilities and restrictions of using the developed model.

Furthermore, I remain sceptical that deep learning provides here significant gains in model accuracy. Obviously, I might be wrong but based on the paper there is no way of telling how good the developed model is. The training data is small, and hence it would be very easy to train some other standard ML models and compare the performance of RND1.0 to those. The authors state that deep learning has previously shown to perform similarly or better than other ML methods, but I see no reason not to compare the developed model to some baseline models and gain some further perspective on the performance of the model.

1. We understand your concern. According to your comments, the uncertainty of RNDv1.0 associated with input variables is estimated from bootstrapping test (Table 4) and discussions are added to the revised manuscript.

In the revised manuscript, the relevant parts are as follows.

Line 235-248:

2.5. Influence of input variables on HONO concentration

A simple bootstrapping test was conducted to evaluate the relative importance of the input variable to HONO concentration. In this analysis, each variable was set to zero and MAE was calculated as an evaluation metrics (Kleinert et al., 2021). Of nine input variables, NO2 was found to have the most significant influence on HONO concentration, followed by RH, temperature, and solar zenith angle (Table 4). The highest MAE of 0.59 ppbv can be considered as the maximum uncertainty of RNDv1.0 due to the input variable.

The result of bootstrap test is in good agreement with those of our previous study (Gil et al., 2021), where more variables such as aerosol surface area and mixing layer height were incorporated into the model, highlighting the crucial role of precursor gases and heterogeneous conversion in HONO formation. Therefore, these results demonstrate that the RND model constructed from routinely measured criteria pollutants and meteorological parameters sufficiently captured the HONO variability in the urban atmosphere.

Line 261-269:

It is possibly due to the variability of HONO that is not fully captured by RNDv1.0 using 9 input variables. As stated above, heterogeneous reactions intimately involved in HONO formation are not considered in RNDv1.0. More importantly, the annual variability of criteria pollutants such as $PM_{2.5}$ has increased in recent years. Particularly in 2019, the monthly average $PM_{2.5}$ mass concentration was lower in April (21 μg m^{-3}) than in May (29 μg m^{-3}), unlike normal years. Given that the test result is within the uncertainty range of the model that is primarily determined by NO_2 (Table 4), RNDv1.0 will be applicable to urban environments under various conditions.

Table 4. The result of bootstrap test of measurement data used to train RNDv1.0 model. The greater the MAE, the greater the influence of variable.

Variable (X)	MAE (ppbv)
-	0.28
O_3	0.29
NO_2	0.59
CO	0.37
SO_2	0.34
Solar zenith Angle (SZA)	0.41
Temperature (T)	0.52
Relative humidity (RH)	0.52
Wind speed (WS)	0.34
Wind direction (WD)	0.29

2. As pointed out, the performance check of RNDv1.0 is necessary, especially due to the small data set for training. Given that HONO observations are extremely limited, however, we had no choice but to seek for practical evidence. Fortunately, HONO simulations in CMAQ model are available for the KORUS-AQ campaign during May~June 2016, and the two sets of HONO concentrations are compared in the revised manuscript (Figure 7). Figure 7 clearly shows that the conventional chemical transport model, CMAQ, fails to capture variations in the ambient HONO mixing ratio, whereas RNDv1.0 works reasonably (IOA = 0.90 and MAE = 0.3 ppbv for RNDv1.0, and IOA = 0.44 and MAE = 0.8 ppbv for CMAQv5.3.1).

In our previous study, HONO was estimated using a 1-layer ANN model with more input variables such as NO mixing ratio, boundary layer height, and surface area of sub-micron particles (Gil et al., 2021). The results of RNDv1.0 was then compared to those of ANN model (Figure A below). While the simulations of ANN model tended to trace better the observed HONO concentrations than RND model, the performance of RNDv1.0 was better than that of ANN model (r^2 = 0.7 and MAE = 0.3 ppbv for RNDv1.0, and r^2 = 0.6 and MAE = 0.4 ppbv for ANN model) (Gil et al., 2021). The accuracy of RNDv1.0 is also better than that of ensemble ML models (RF, BPNN, GBDT, r^2 = 0.7, MAE = 0.3 ~ 0.5 ppbv) (Cui et al., 2021).

In the revised manuscript, the relevant parts are as follows.

Line 226-233:

Next, the HONO calculated in RNDv.1.0 was compared with observations and results from CMAQ (Community Multi-scale Air Quality Model, v5.3.1) simulations during the KORUS-AQ study (May~June 2016) (Figure 7). More information on CMAQ modeling can be found elsewhere (Appel et al., 2021). While the results of RNDv.1.0 reasonably traced the observed variations (IOA = 0.90), the CMAQ severely underestimated the measured HONO concentration (IOA = 0.44). These results demonstrate the performance and efficacy of RNDv1.0 in calculating the ambient HONO mixing ratio that are poorly reproduced in conventional operating models.

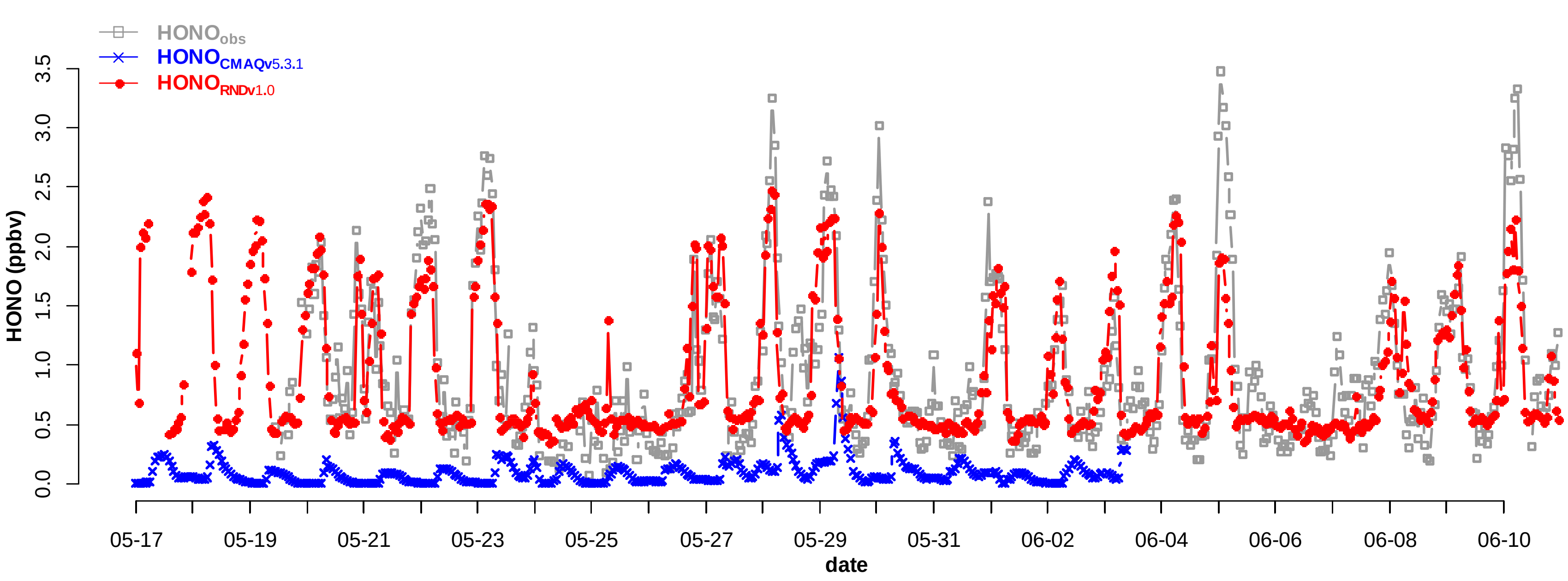

Figure 7. During the KORUS-AQ campaign (May-June 2016), HONO mixing ratios calculated using RNDv1.0 (red dot) are compared with those observed (gray square) and calculated using CMAQv5.3.1 (blue cross).

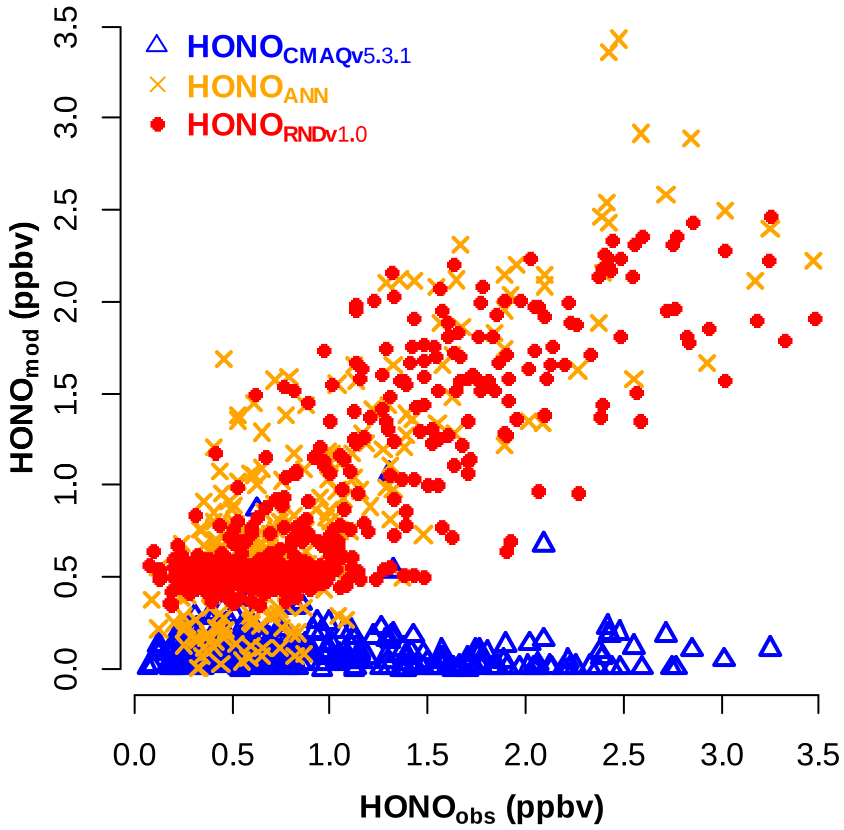

Figure A. HONO concentrations calculated in RNDv1.0 (red dot) are compared with observations and those calculated in one-layer ANN model (orange cross) and in CMAQv5.3.1 (blue triangle) during the KORUS-AQ campaign (May~June 2016).