Corrigendum to Atmos. Chem. Phys., 20, 12853–12869, 2020
https://doi.org/10.5194/acp-20-12853-2020-corrigendum

Atmospheric Chemistry and Physics
Open Access
EGU

Corrigendum to "Using machine learning to derive cloud condensation nuclei number concentrations from commonly available measurements" published in Atmos. Chem. Phys., 20, 12853–12869, 2020

Arshad Arjunan Nair and Fangqun Yu

Atmospheric Sciences Research Center, State University of New York, Albany, New York 12203, USA

Correspondence: Arshad Arjunan Nair (aanair@albany.edu)

Published: 4 November 2021

During submission the following statement was regrettably omitted, which we provide here.

In Sect. 3.2.2, subsection "RFRM application: measured predictors as input", at the end of the first paragraph, please append the following:

> While the RFRM is trained on speciated $PM_{2.5}$, only submicron (PM_1) measurements are made at the SGP site. When input as predictors, it is assumed that most of the $PM_{2.5}$ mass is held by aerosols with a diameter of less than 1 µm.

We thank Dr. James J. Schwab for bringing this to our attention.

Appendix A

For potentially arising queries in the reader's mind, we provide additional discussion.

A1 Is the assumption $PM_1 \approx PM_{2.5}$ reasonable?

The aerosol mass size distribution ($dM/d\log D_p$ vs. D_p) or PMSD typically has two major modes: the (accumulation) fine mode and coarse mode. For $PM_{2.5}$, i.e., mass of aerosols with a diameter of less than 2.5 µm, the fine mode – typically maximum around 0.5 µm – contributes to most of $PM_{2.5}$. *Typical* refers to ambient conditions where condensational growth and coagulation determine the PMSD (for this aerosol size range) rather than transport from sources. For the SGP surface layer, the GCAPM-simulated $PM_1 : PM_{2.5}$ ratio is illustrated in Fig. 1. Considering simulated ($PM_{2.5}$–PM_1) $\ll PM_1$ and the absence of $PM_{2.5}$ speciation measurements, we assume that $PM_1 \approx PM_{2.5}$ when measurements are input as predictors to the RFRM in Sect. 3.2.2, subsection "RFRM application: measured predictors as input".

A2 Why not train the RFRM to use PM_1 as input?

Validating with measurements is not the primary goal of this paper. Regardless, the RFRM can be retrained to use speciated PM_1 as input predictors. Consider the two cases:

1. $PM_{2.5}$-RFRM: RFRM-ShortVars is trained on GEOS-Chem-APM-modeled $PM_{2.5}$ speciation (and other predictors as in RFRM-ShortVars) and uses measured PM_1 as input to derive [CCN0.4]. This is identical to RFRM-ShortVars in the published paper.

2. PM_1-RFRM: RFRM-ShortVars is trained (instead) on modeled PM_1 speciation and uses measured PM_1 as input to derive [CCN0.4].

Figure 2 compares their predictions (daily-aggregated; cf. Fig. 12b). Figure 3 compares each RFRM's predictions with observations (cf. Fig. 13a). Due to accounting for rare coarse-mode mass, there is a slight difference (improvement) in the RFRM performance between the two cases.

Published by Copernicus Publications on behalf of the European Geosciences Union.

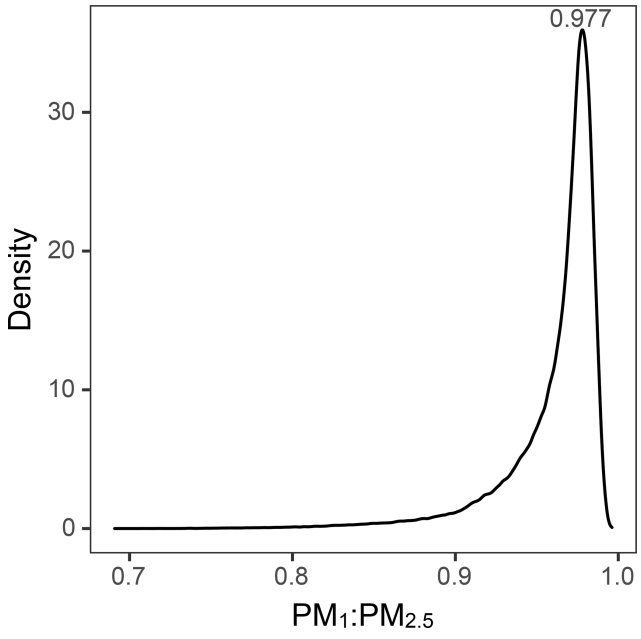

Figure 1. Density plot for the GCAPM-simulated $PM_1 : PM_{2.5}$ ratio for the SGP surface layer.

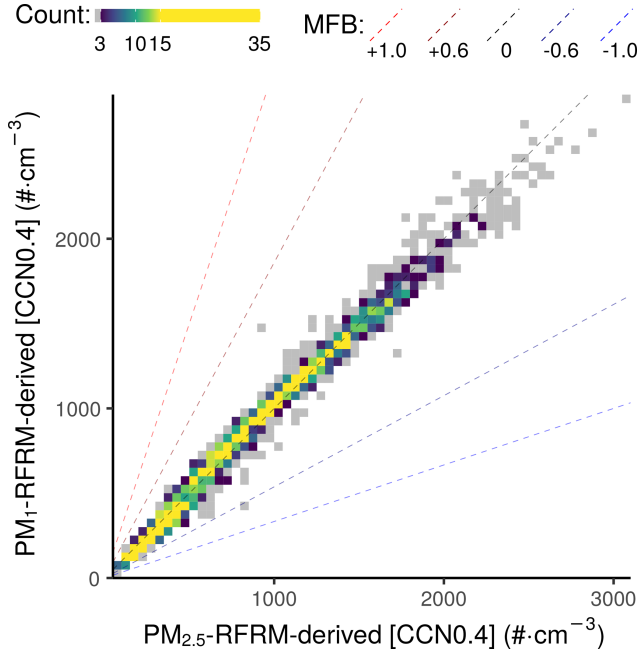

Figure 2. Comparison of the RFRM derivation of [CCN0.4].

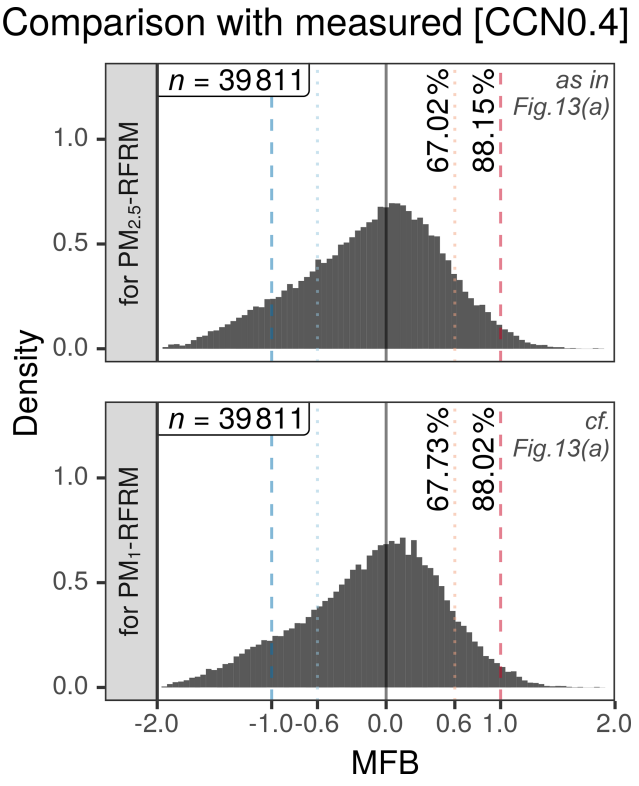

Figure 3. Mean fractional bias (MFB) of the RFRM-derived [CCN0.4] compared to SGP-measured [CCN0.4]. The histograms show the pairwise counts by MFB (total is inset top left in each panel). The lines indicate an MFB of 0 (black), +1 (dashed red), −1 (dashed blue), +0.6 (dotted red), and −0.6 (dotted blue). The percentage of RFRM-derived values in good ($|\text{MFB}| < 0.6$) and fair ($|\text{MFB}| < 1$) agreement are shown close to the +0.6 and +1.0 MFB lines, respectively.