Supplement of

Site and Season Specific Calibrations Improve Low-cost Sensor Performance: Long-term Field Evaluation of PurpleAir Sensors in Urban and Rural India

Mark J Campmier[1], Jonathan Gingrich[2], Saumya Singh[1], Nisar Baig[3], Shahzad Gani[4,5], Adithi Upadhya[6], Pratyush Agrawal[7], Meenakshi Kushwaha[6], Harsh Raj Mishra[8], Ajay Pillarisetti[9], Sreekanth Vakacherla[7], Ravi Kant Pathak[8,10], Joshua S Apte[1,9*]

[1]Department of Civil and Environmental Engineering, University of California, Berkeley, CA, 94720, USA
[2]Department of Engineering, Dordt University, Sioux Center, IA, 51250, USA
[3]Department of Civil Engineering, Indian Institute of Technology Delhi, New Delhi 110016, India
[4]Centre for Atmospheric Sciences, Indian Institute of Technology Delhi, New Delhi 110016, India
[5]Institute for Atmospheric and Earth System Research/Physics, University of Helsinki, Helsinki, 00100, Finland
[6]ILK Labs, Bengaluru, India
[7]Center for Study of Science, Technology and Policy, Bengaluru 560094, India
[8] Indo Gangetic Plains-Centre for Air Research and Education, Hamirpur 210301, India
[9]School of Public Health, University of California, Berkeley, CA 94720, USA
[10]Department of Chemistry and Molecular Biology, University of Gothenburg, Gothenburg, Sweden
[*] Corresponding author. *E-mail address* apte@berkeley.edu

Table S1. Most relevant parameters selected through sequential feature selection (SFS) for each $PM_{2.5}$ channel (CF1, ATM, ALT) from Delhi collocation. Parameters: relative humidity (RH), temperature (T), and dew point (D). As expected, the $PM_{2.5}$ was selected as the most relevant, followed by RH features.

Ranking	CF1	ATM	ALT
1	$PM_{2.5}$	$PM_{2.5}$	$PM_{2.5}$
2	RH	RH^2	RH^2*T
3	$PM_{2.5}*RH$	$PM_{2.5}*RH^2$	$PM_{2.5}^2*RH$
4	$PM_{2.5}*RH^2$	$PM_{2.5}^2*D$	$PM_{2.5}^2*D$
5	RH^2	$PM_{2.5}^2$	$PM_{2.5}^2$
6	$PM_{2.5}^2*D$	$PM_{2.5}^3$	$PM_{2.5}^3$
7	$PM_{2.5}^2$	$PM_{2.5}^2*RH$	$PM_{2.5}*RH*T$
8	$PM_{2.5}^3$	RH	$PM_{2.5}*RH$
9	$PM_{2.5}^2*T$	$PM_{2.5}*RH$	D
10	$PM_{2.5}^2*RH$	$PM_{2.5}^2*T$	D^3
11	RH^3	RH^3	RH
12	$PM_{2.5}*RH*T$	$PM_{2.5}*RH*D$	$PM_{2.5}*RH^2$
13	$PM_{2.5}*T$	$PM_{2.5}*RH*T$	$PM_{2.5}*RH*D$
14	$PM_{2.5}*D^2$	$PM_{2.5}*D$	D^2
15	$PM_{2.5}*T^2$	$PM_{2.5}*D^2$	$PM_{2.5}*D^2$
16	$PM_{2.5}*T*D$	T^3	$PM_{2.5}*D$
17	T^3	PM2.5*T	$PM_{2.5}*T$
18	$PM_{2.5}*D$	T^2*D	T
19	T^2*D	D	$PM_{2.5}^2*T$
20	D	RH^2*D	RH^3
21	RH^2*D	T	RH^2
22	T*D	RH*T*D	$PM_{2.5}*T^2$
23	T^2	T*D	$PM_{2.5}*T*D$
24	$RH*D^2$	$RH*T^2$	RH^2*D
25	T	D^2	T^3
26	D^2	T^2	$T*D^2$
27	D^3	$PM_{2.5}*T^2$	RH*T
28	$PM_{2.5}*RH*D$	$PM_{2.5}*T*D$	$RH*D^2$
29	RH*T*D	RH*T	$RH*T^2$
30	$RH*T^2$	$T*D^2$	T*D
31	RH^2*T	RH^2*T	RH*D
32	RH*D	RH*D	RH*T*D
33	$T*D^2$	$RH*D^2$	T^2
34	RH*T	D^3	T^2*D

Table S2. Most relevant parameters selected through sequential feature selection (SFS) for each $PM_{2.5}$ channel (CF1, ATM, ALT) from Hamirpur collocation. Parameters: relative humidity (RH), temperature (T), and dew point (D). As expected, the $PM_{2.5}$ was selected as the most relevant, followed by RH features.

Ranking	CF1	ATM	ALT
1	$PM_{2.5}$	$PM_{2.5}$	$PM_{2.5}$
2	RH	RH	$PM_{2.5}$*RH*T
3	RH^3	RH^2	$PM_{2.5}^2$*D
4	RH^2	RH^3	$PM_{2.5}^2$*T
5	$PM_{2.5}^2$*T	$PM_{2.5}^2$*T	T^2*D
6	$PM_{2.5}$*D	T^3	T*D
7	RH^2*T	T^2*D	T^3
8	$PM_{2.5}$*D^2	T*D^2	D^3
9	RH^2*D	D^3	T*D^2
10	RH*D^2	T^2	RH*T^2
11	T^3	$PM_{2.5}$*D^2	D
12	D^3	RH*D^2	RH^2*D
13	T^2*D	D^2	RH^3
14	RH*T	T	T^2
15	T*D^2	$PM_{2.5}$*D	D^2
16	RH*D	D	RH*T
17	D	T*D	RH*D
18	RH*T*D	$PM_{2.5}$*T^2	RH*D^2
19	T^2	$PM_{2.5}^2$*RH	T
20	T	RH*T	RH*T*D
21	RH*T^2	RH*D	RH
22	D^2	RH^2*D	RH^2
23	T*D	RH^2*T	RH^2*T
24	$PM_{2.5}$*T^2	RH*T*D	$PM_{2.5}$*RH*D
25	$PM_{2.5}^2$*RH	RH*T^2	$PM_{2.5}$*D^2
26	$PM_{2.5}$*T	$PM_{2.5}$*T	$PM_{2.5}$*D
27	$PM_{2.5}$*RH^2	$PM_{2.5}$*RH^2	$PM_{2.5}$*T*D
28	$PM_{2.5}$*RH	$PM_{2.5}$*RH	$PM_{2.5}$*T^2
29	$PM_{2.5}$*T*D	$PM_{2.5}$*T*D	$PM_{2.5}$*RH
30	$PM_{2.5}$*RH*D	$PM_{2.5}$*RH*D	$PM_{2.5}$*T
31	$PM_{2.5}$*RH*T	$PM_{2.5}$*RH*T	$PM_{2.5}$*RH^2
32	$PM_{2.5}^2$	$PM_{2.5}^2$	$PM_{2.5}^2$
33	$PM_{2.5}^2$*D	$PM_{2.5}^2$*D	$PM_{2.5}^3$
34	$PM_{2.5}^3$	$PM_{2.5}^3$	$PM_{2.5}^2$*RH

Table S3. Most relevant parameters selected through sequential feature selection (SFS) for each $PM_{2.5}$ channel (CF1, ATM, ALT) from Bangalore collocation. Parameters: relative humidity (RH), temperature (T), and dew point (D). As expected, the $PM_{2.5}$ was selected as the most relevant for the ALT channel, followed by RH features. The other channels still select $PM_{2.5}$ but in combination with other features - likely an artifact of correlations between sensor performance, and diurnal temperature patterns. This is less noticeable in Hamirpur and Delhi where the temperature range is more seasonally dynamic.

Ranking	CF1	ATM	ALT
1	$PM_{2.5}$*T	$PM_{2.5}^2$*T	$PM_{2.5}$
2	$PM_{2.5}$*T*D	$PM_{2.5}^3$	$PM_{2.5}$*RH^2
3	$PM_{2.5}^2$*T	$PM_{2.5}^2$	$PM_{2.5}^2$*T
4	$PM_{2.5}$*RH^2	$PM_{2.5}$*RH^2	RH^3
5	$PM_{2.5}$	T^3	$PM_{2.5}$*T^2
6	D	T^2*D	$PM_{2.5}$*T
7	$PM_{2.5}^2$	$PM_{2.5}$*T^2	T^3
8	RH*T^2	$PM_{2.5}$	$PM_{2.5}^2$
9	T	$PM_{2.5}$*T	T^2
10	RH*T	T	$PM_{2.5}^2$*RH
11	$PM_{2.5}^2$*RH	T^2	T
12	RH	$PM_{2.5}$*RH*D	T^2*D
13	RH^2*T	$PM_{2.5}$*RH	$PM_{2.5}$*RH*D
14	D^3	$PM_{2.5}$*RH*T	T*D^2
15	RH^3	$PM_{2.5}$*D^2	RH^2*D
16	RH*D^2	$PM_{2.5}$*D	RH*D^2
17	RH^2*D	$PM_{2.5}$*T*D	$PM_{2.5}$*D^2
18	RH*D	D	$PM_{2.5}$*T*D
19	RH^2	D^2	$PM_{2.5}^2$*D
20	RH*T*D	T*D^2	RH*T^2
21	D^2	RH*T^2	$PM_{2.5}$*RH*T
22	T*D	D^3	D
23	T*D^2	T*D	$PM_{2.5}$*D
24	T^2	RH^2*D	$PM_{2.5}$*RH
25	T^3	RH	T*D
26	T^2*D	RH*D	D^2
27	$PM_{2.5}$*D^2	RH^3	RH
28	$PM_{2.5}^2$*D	RH*T*D	RH*D
29	$PM_{2.5}$*T^2	RH*D^2	RH^2*T
30	$PM_{2.5}$*RH*D	RH^2	RH*T
31	$PM_{2.5}$*RH*T	RH^2*T	D^3
32	$PM_{2.5}$*RH	RH*T	RH*T*D
33	$PM_{2.5}$*D	$PM_{2.5}^2$*RH	$PM_{2.5}^3$
34	$PM_{2.5}^3$	$PM_{2.5}^2$*D	RH^2

Table S4. Fitted coefficients from the hygroscopic growth model for each site and data channel. The broad variability in the β coefficient is indicative of the variable hygroscopic properties across India, and even among the two the Indo-Gangetic Plain sites, given the wide disparity between Delhi and Hamirpur fitted values.

		Delhi	Hamirpur	Bangalore
ATM	α	1.011	0.887	0.907
	β	0.424	0.157	0.082
CF1	α	0.671	0.602	0.714
	β	0.379	0.126	0.107
ALT	α	4.824	3.838	3.598
	β	0.202	0.132	0.056

;

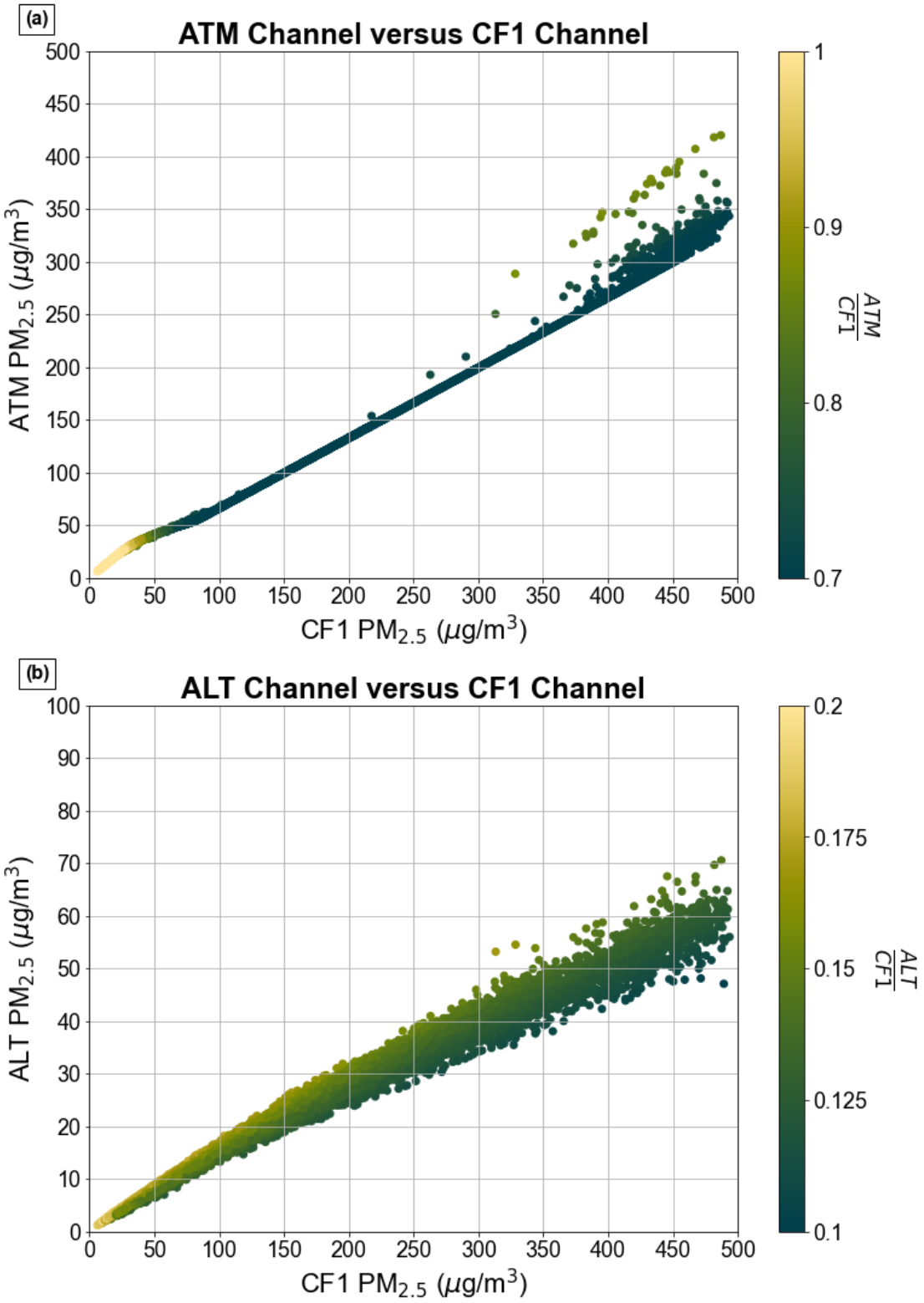

Figure S1. Panel (a) shows correlation between PurpleAir ATM channel and CF1 channel. Values are 1:1 correlated until about 35 μg/m^3 where the ratio gradually transitions to 0.7:1 at around 50 μg/m^3 (CF1). Panel (b) shows correlation between PurpleAir ALT channel and CF1 channel. Since ALT does not include any density assumptions, and contains some noise from the particle distribution model, the ratio follows a distribution centered at around 0.15:1.

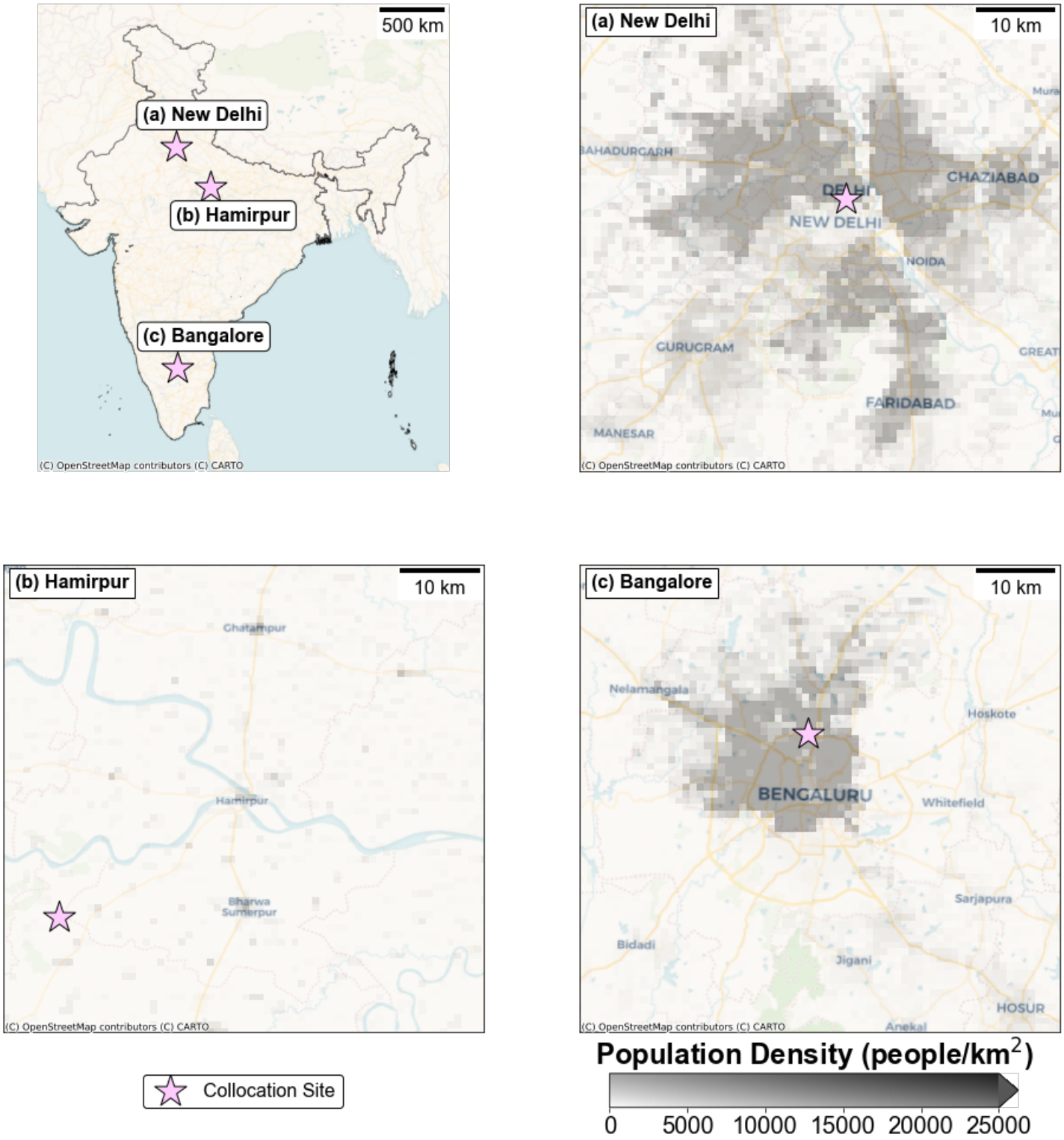

Figure S2. Illustration of collocation sites (PurpleAir-Reference pairs) across India. The top left panel depicts the overall map of the country, with New Delhi and Hamirpur in the northern Indo-Gangetic Plain region, and Bangalore in the south. Panels (a – c) highlight the locations of the collocation sites within a given settlement, with shading representing the population density from the Global Human Settlement Layer (Schiavina et al., 2022). Panel (a) shows the central location of the US Embassy collocation site within the dense urban center of New Delhi. Panel (b) shows the extremely sparse IGP-CARE collocation site on the outskirts of Hamirpur. Panel (c) shows the CSTEP collocation site within the dense urban core of Bangalore.

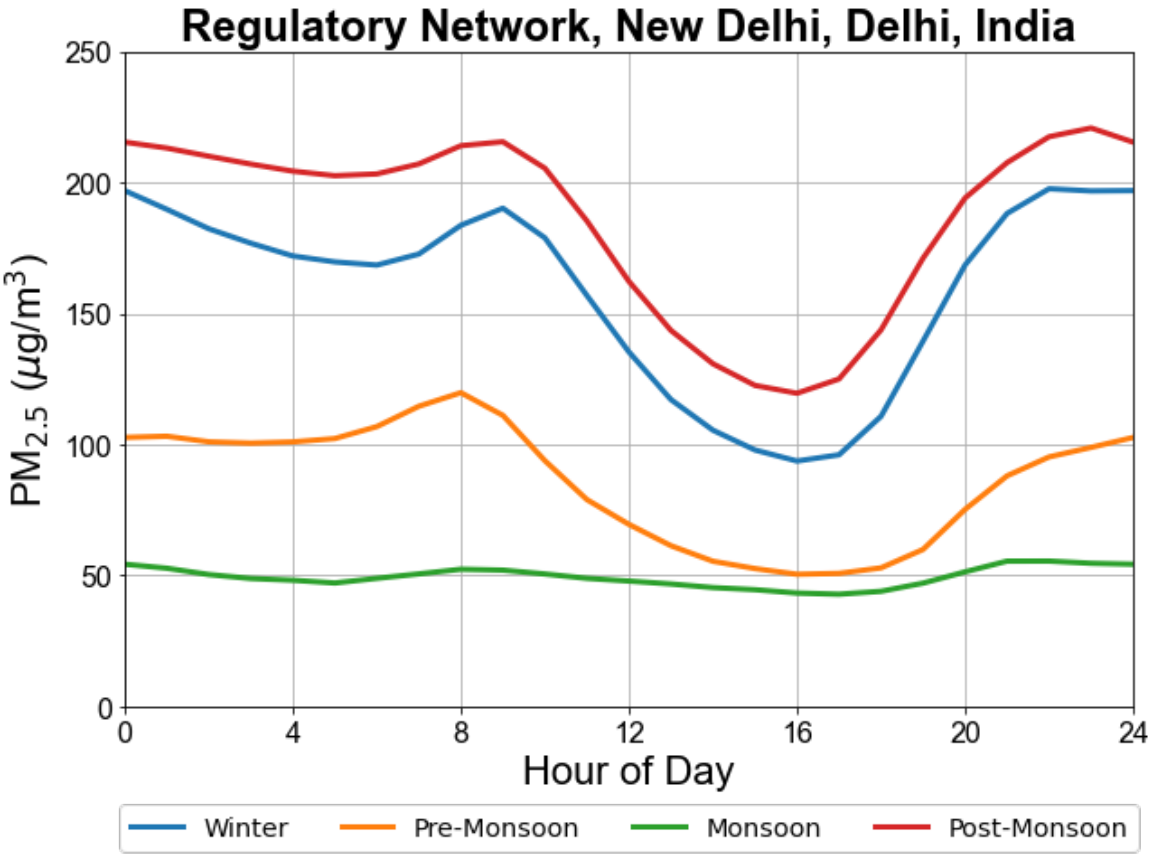

Figure S3. Mean diurnal profiles of $PM_{2.5}$ by season averaged across all reference sites in New Delhi. Mean concentrations are highest in the post-monsoon and winter seasons, with strong morning and evening peaks. The monsoon period features the lowest mass concentrations (albeit still above World Health Organization guidelines), as well as much less dynamic range due to the rains and ventilation conditions during the monsoon season.

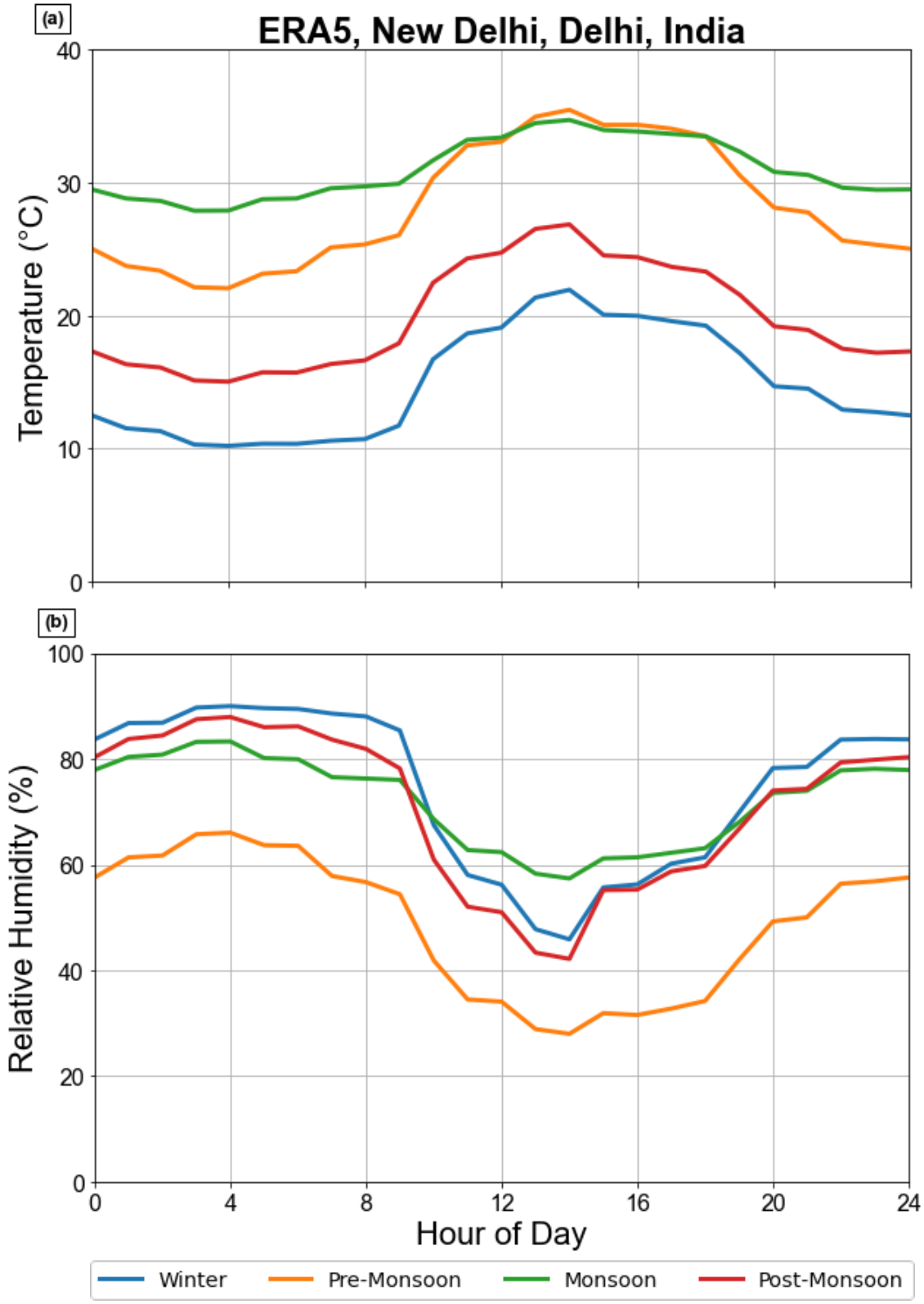

Figure S4. Panel (a) shows mean diurnal profiles of relative humidity by season with the ERA5 pixel containing the Delhi collocation site. Panel (b) shows mean diurnal profiles of temperature by season with the ERA5 pixel containing the Delhi collocation site.

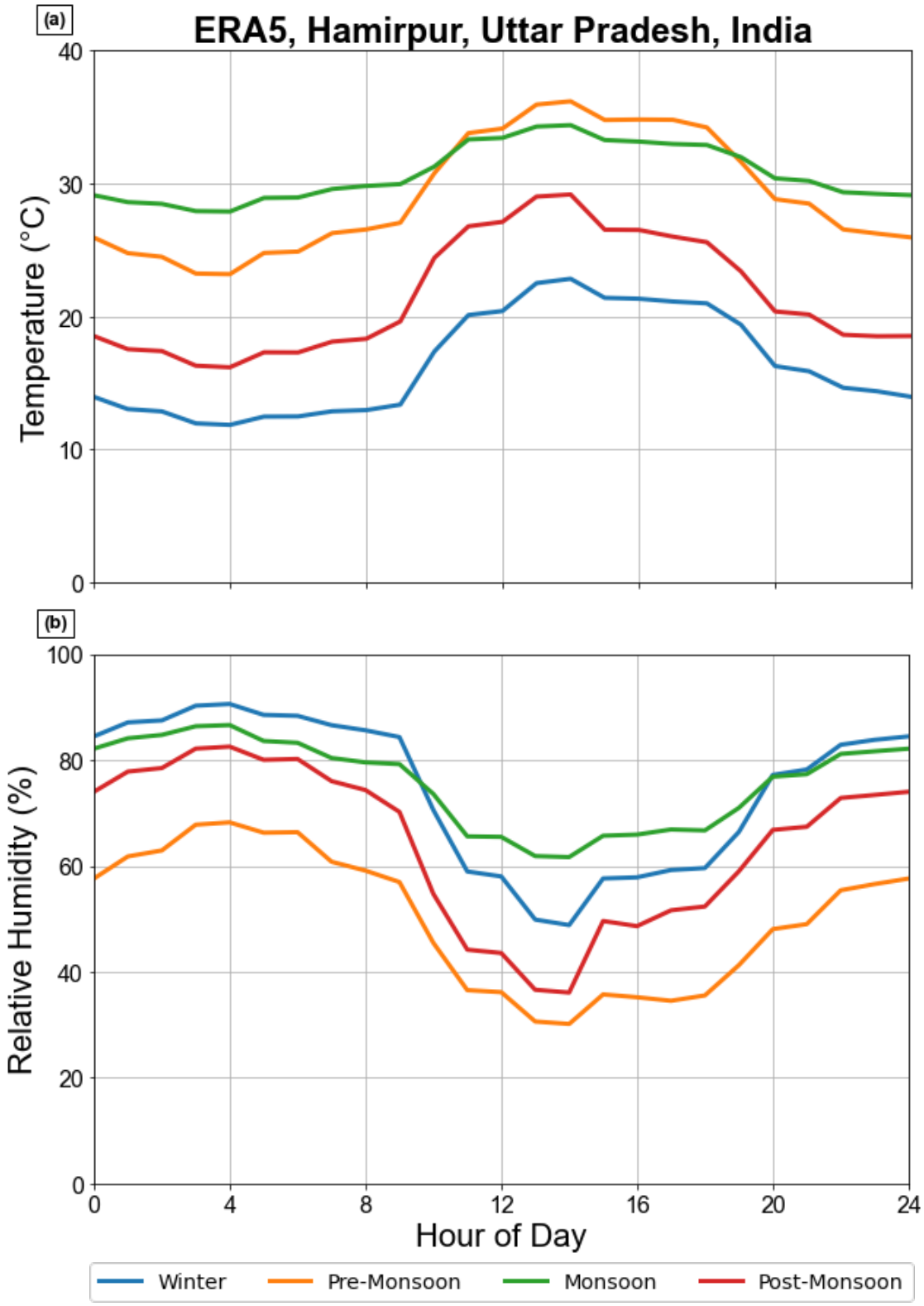

Figure S5. Panel (a) shows mean diurnal profiles of relative humidity by season with the ERA5 pixel containing the Hamirpur collocation site. Panel (b) shows mean diurnal profiles of temperature by season with the ERA5 pixel containing the Hamirpur collocation site.

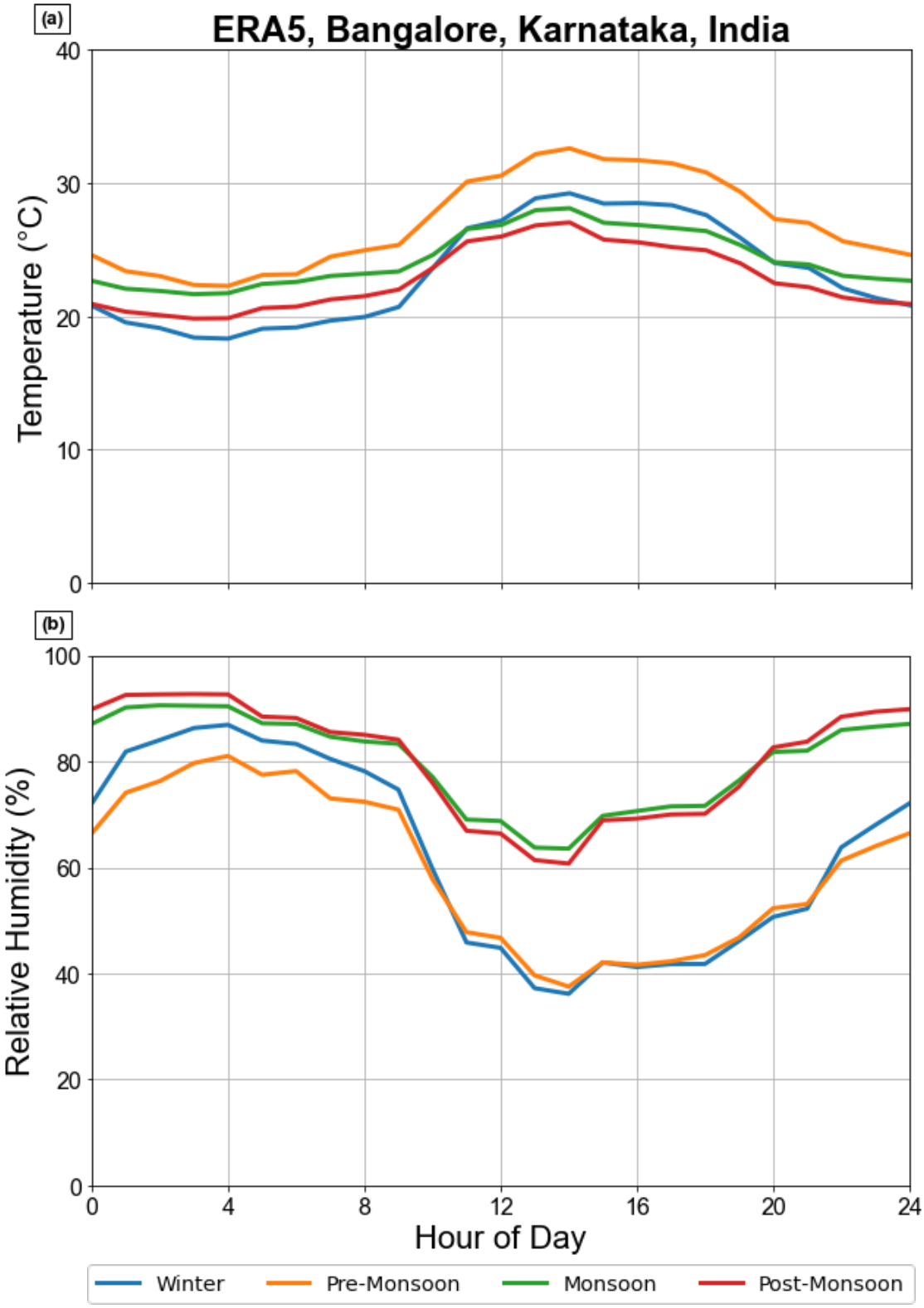

Figure S6. Panel (a) shows mean diurnal profiles of relative humidity by season with the ERA5 pixel containing the Bangalore collocation site. Panel (b) shows mean diurnal profiles of temperature by season with the ERA5 pixel containing the Bangalore collocation site.

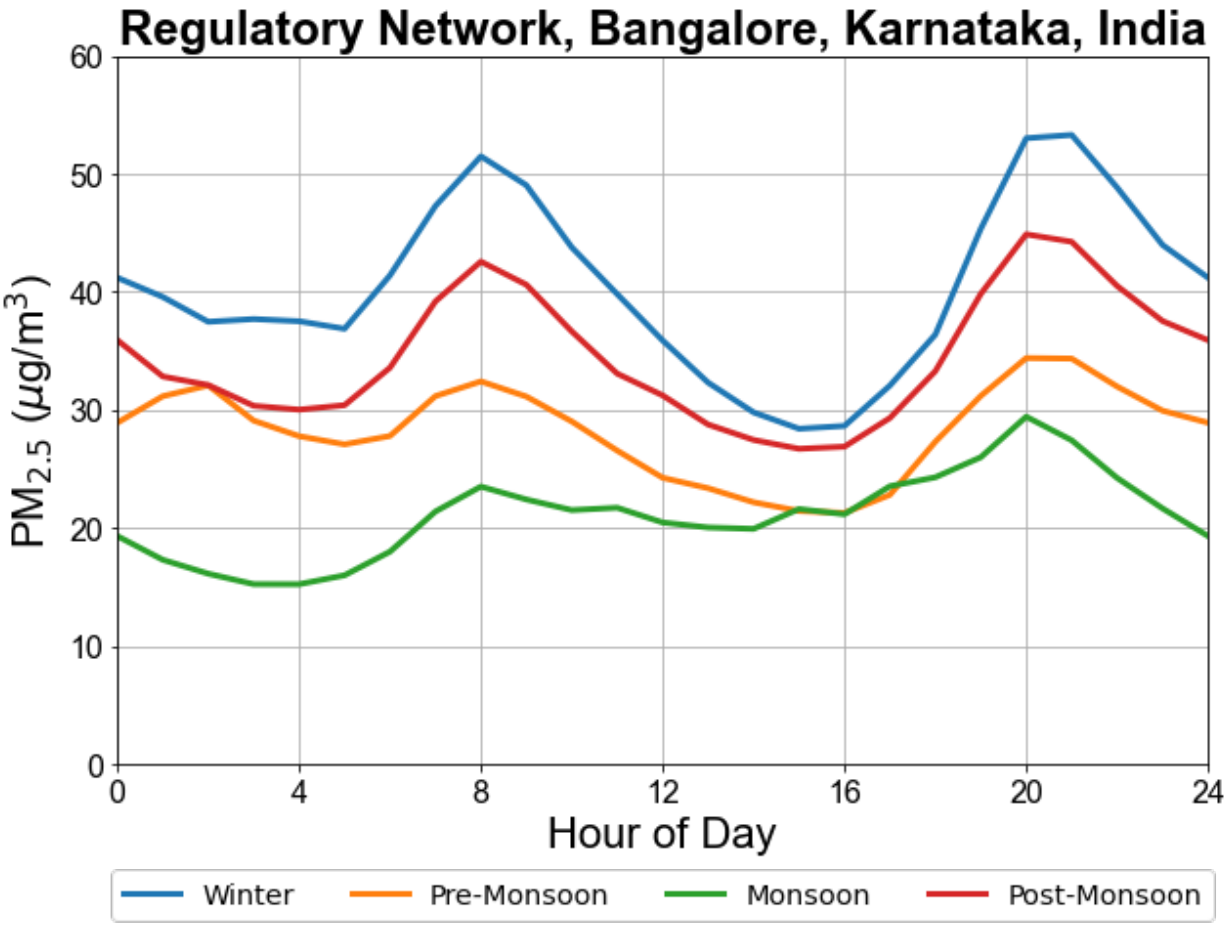

Figure S7. Mean diurnal profiles of $PM_{2.5}$ by season averaged across all reference sites in Bangalore. Mean concentrations are highest in the post-monsoon and winter, with strong morning and evening peaks. The monsoon period features the lowest mass as well as much less dynamic range due to the rains and ventilation conditions during the monsoon season.

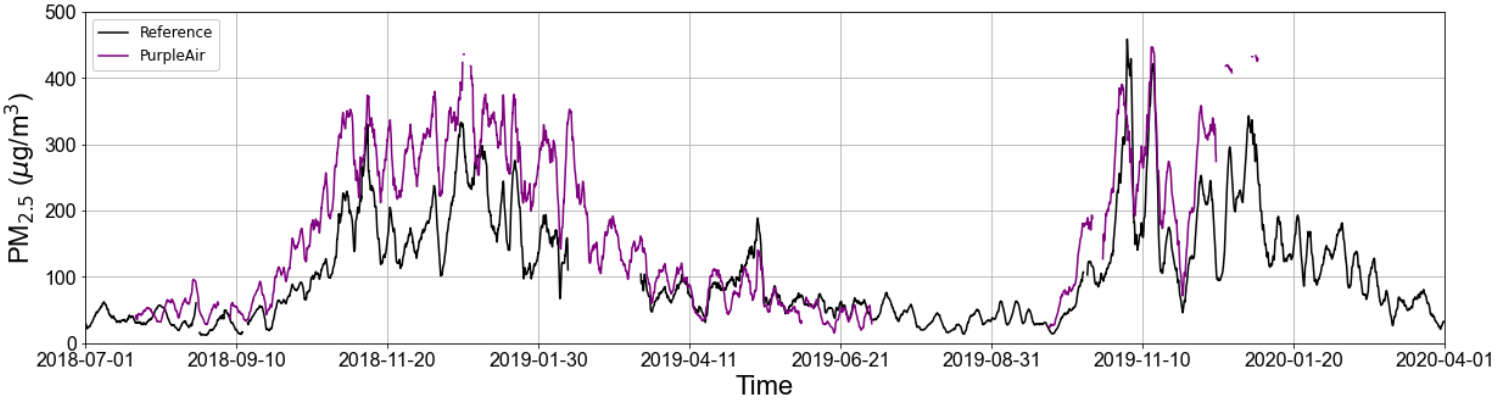

Figure S8. Time series of Delhi PurpleAir (CF1) and reference (BAM-1020) signal throughout the 21-month collocation period with 72-hour rolling mean smoothing. Gaps indicate insufficient data.

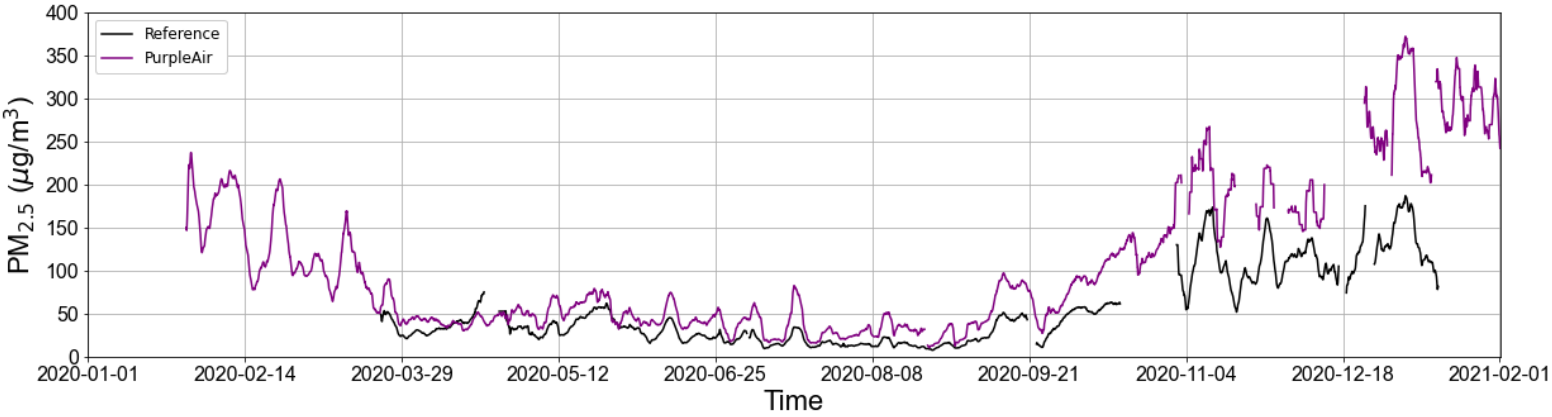

Figure S9. Time series of Hamirpur PurpleAir (CF1) and reference (BAM-1022) signal throughout the collocation period with 72-hour rolling mean smoothing. Gaps indicate insufficient data.

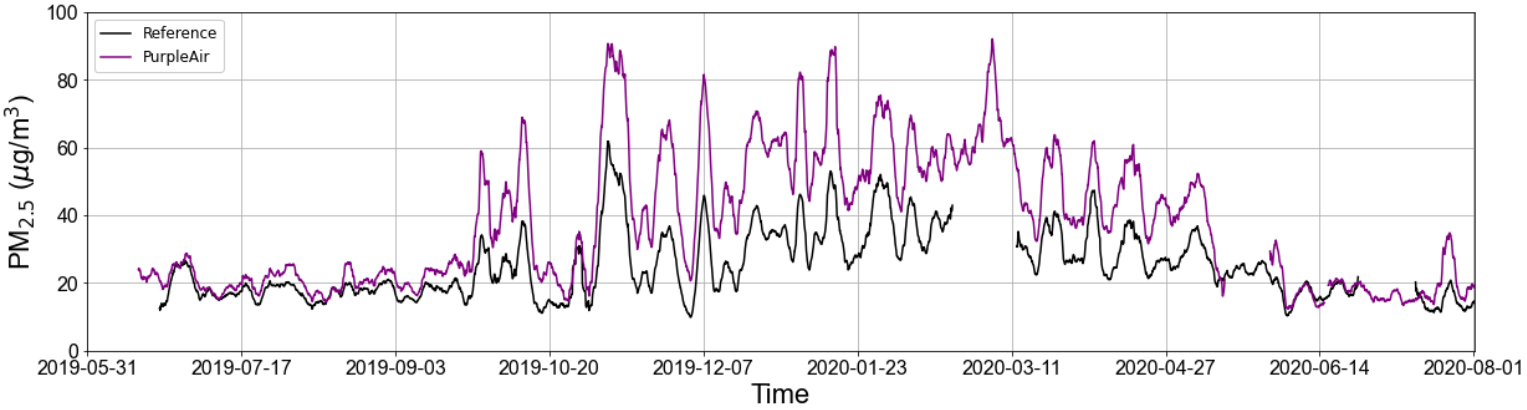

Figure S10. Time series of Bangalore PurpleAir (CF1) and reference (BAM-1022) signal throughout the collocation period with 72-hour rolling mean smoothing. Gaps indicate insufficient data.

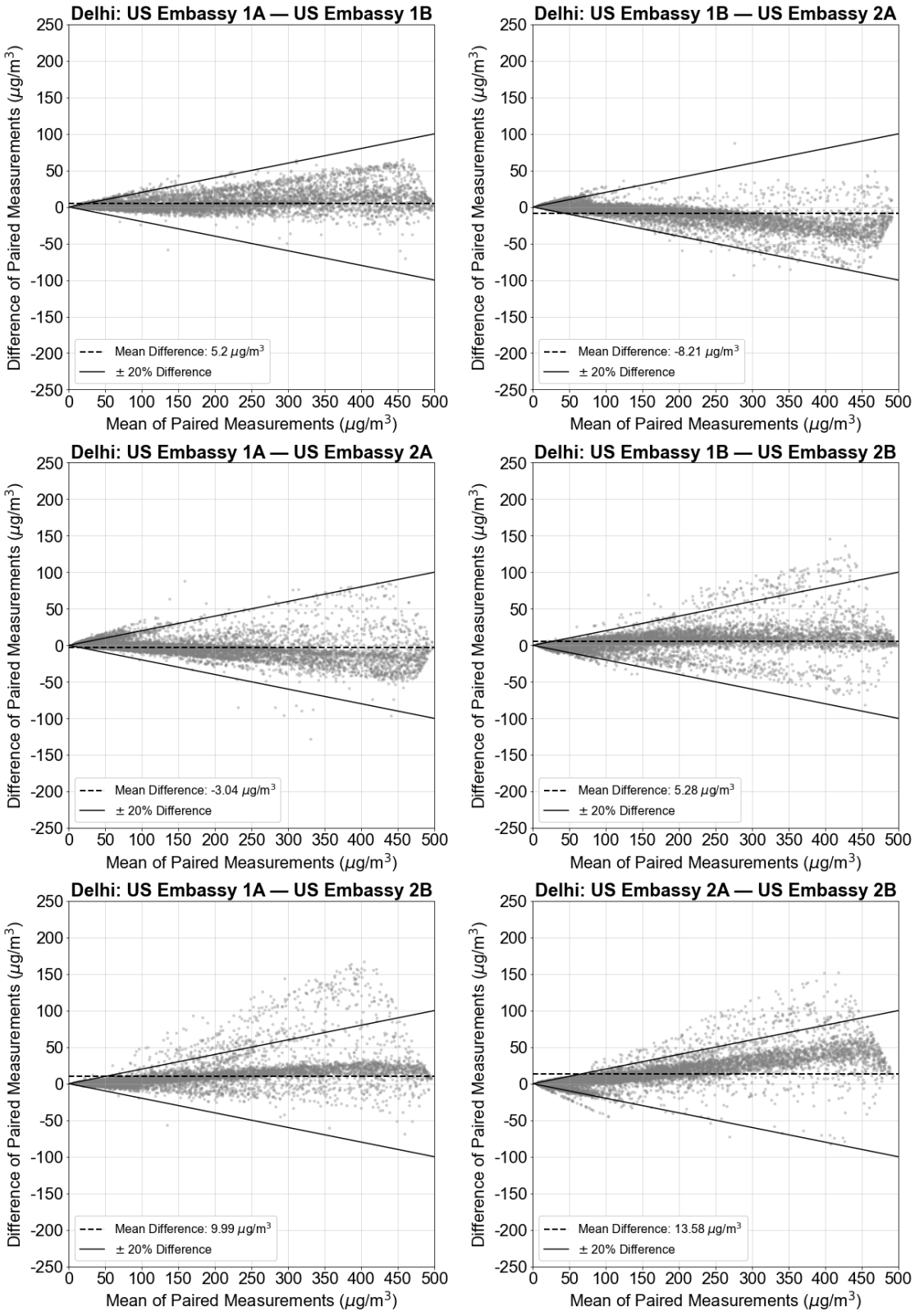

Figure S11. Bland-Altman plots for PA at the Delhi collocation site, the US Embassy, indicating the structure of pairwise differences in four individual Plantower channels from two PurpleAir sensors as function of concentration. The bulk of hourly scatter points are within 20% difference.

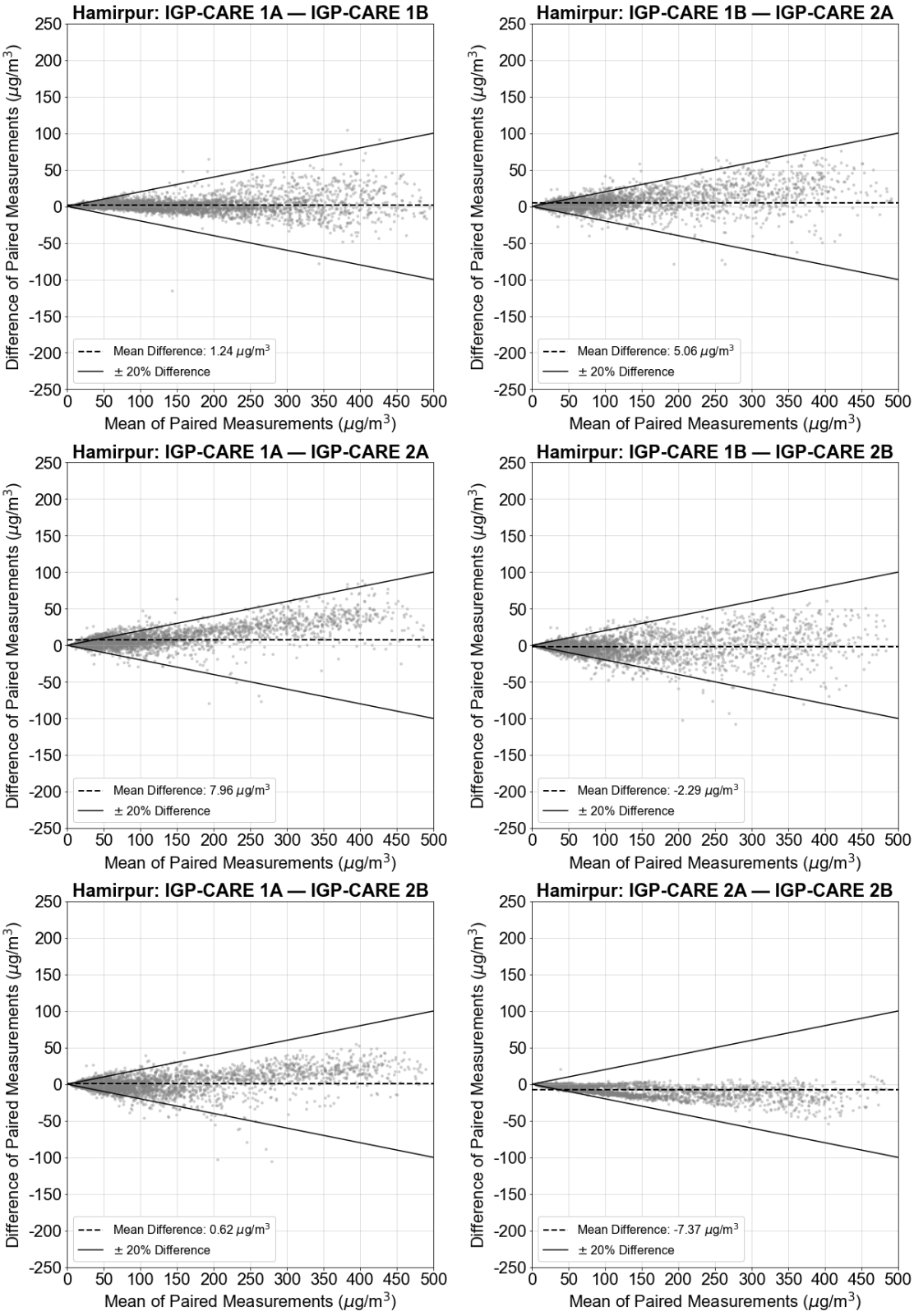

Figure S12. Bland-Altman plots for PA at the Hamirpur collocation site, IGP-CARE, indicating the structure of pairwise differences in four individual Plantower channels from two PurpleAir sensors as function of concentration. The bulk of hourly scatter points are within 20% difference.

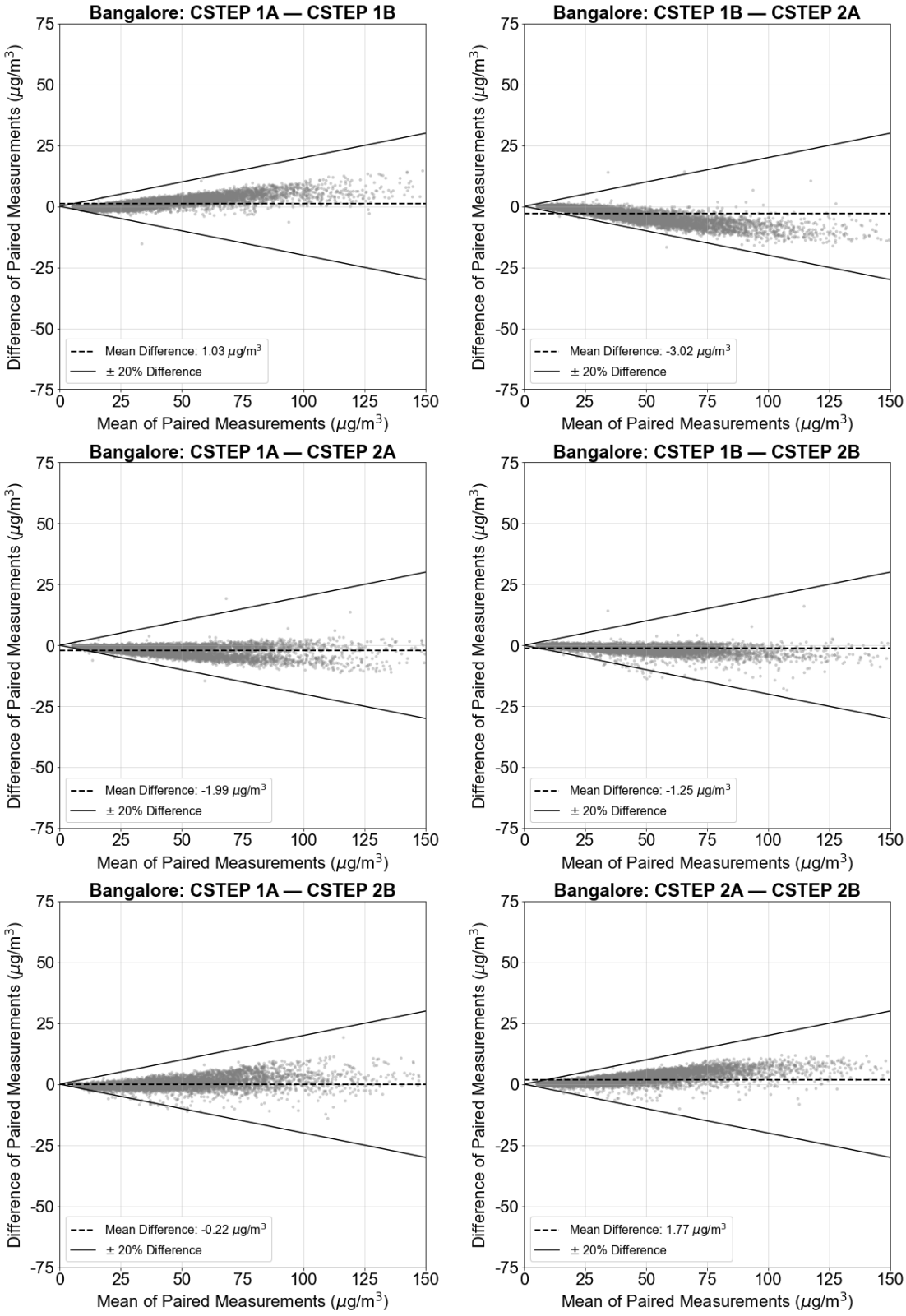

Figure S13. Bland-Altman plots for PA at the Bangalore collocation site, CSTEP, indicating the structure of pairwise differences in four individual Plantower channels from two PurpleAir sensors as function of concentration. The bulk of hourly scatter points are within 20% difference.

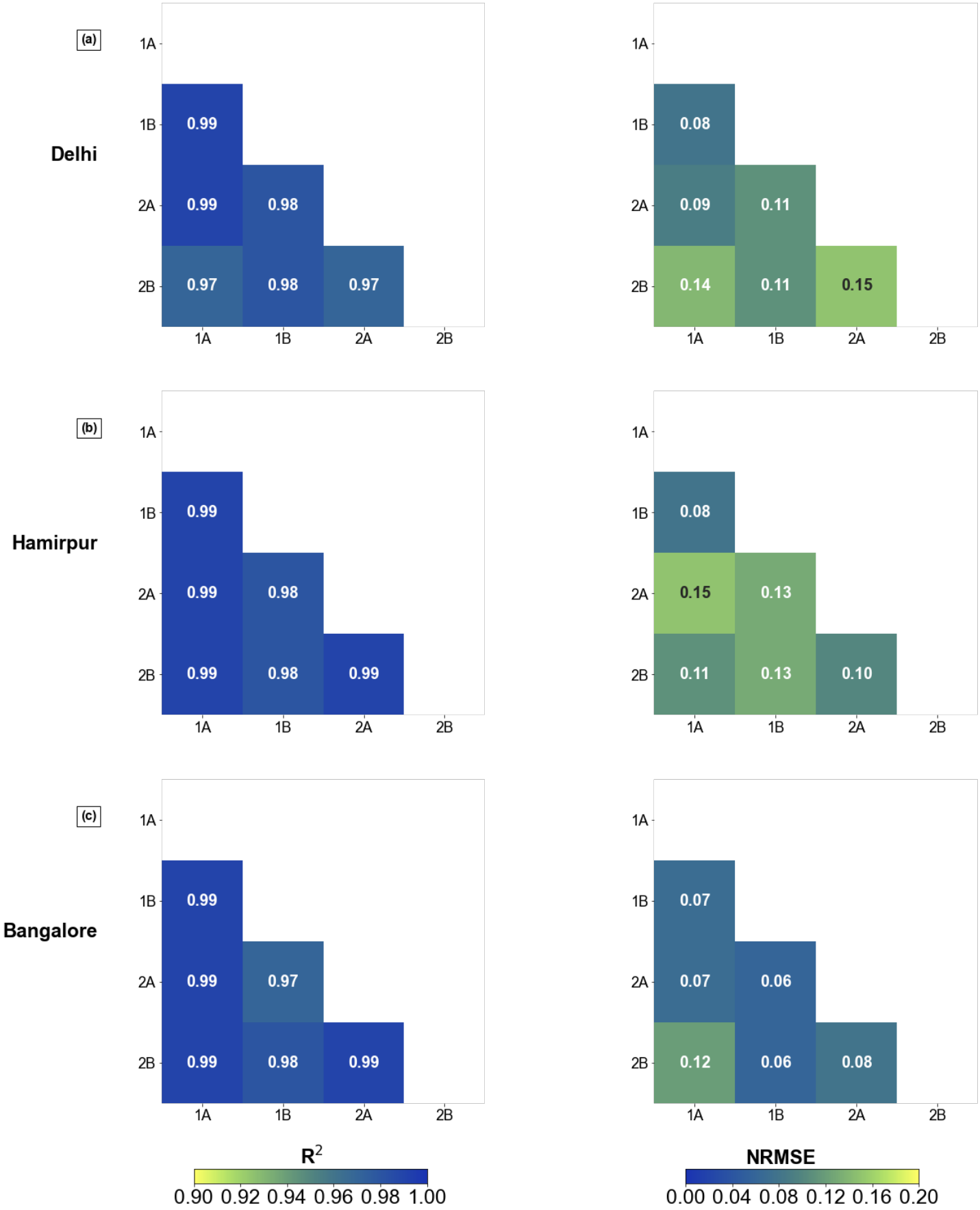

Figure S14: Intra-Plantower R^2 (left) and NRMSE (right) for the two collocated PurpleAir units in Delhi (panel [a]), Hamirpur (panel [b]), and Bangalore (panel [c]). All units demonstrate high precision.

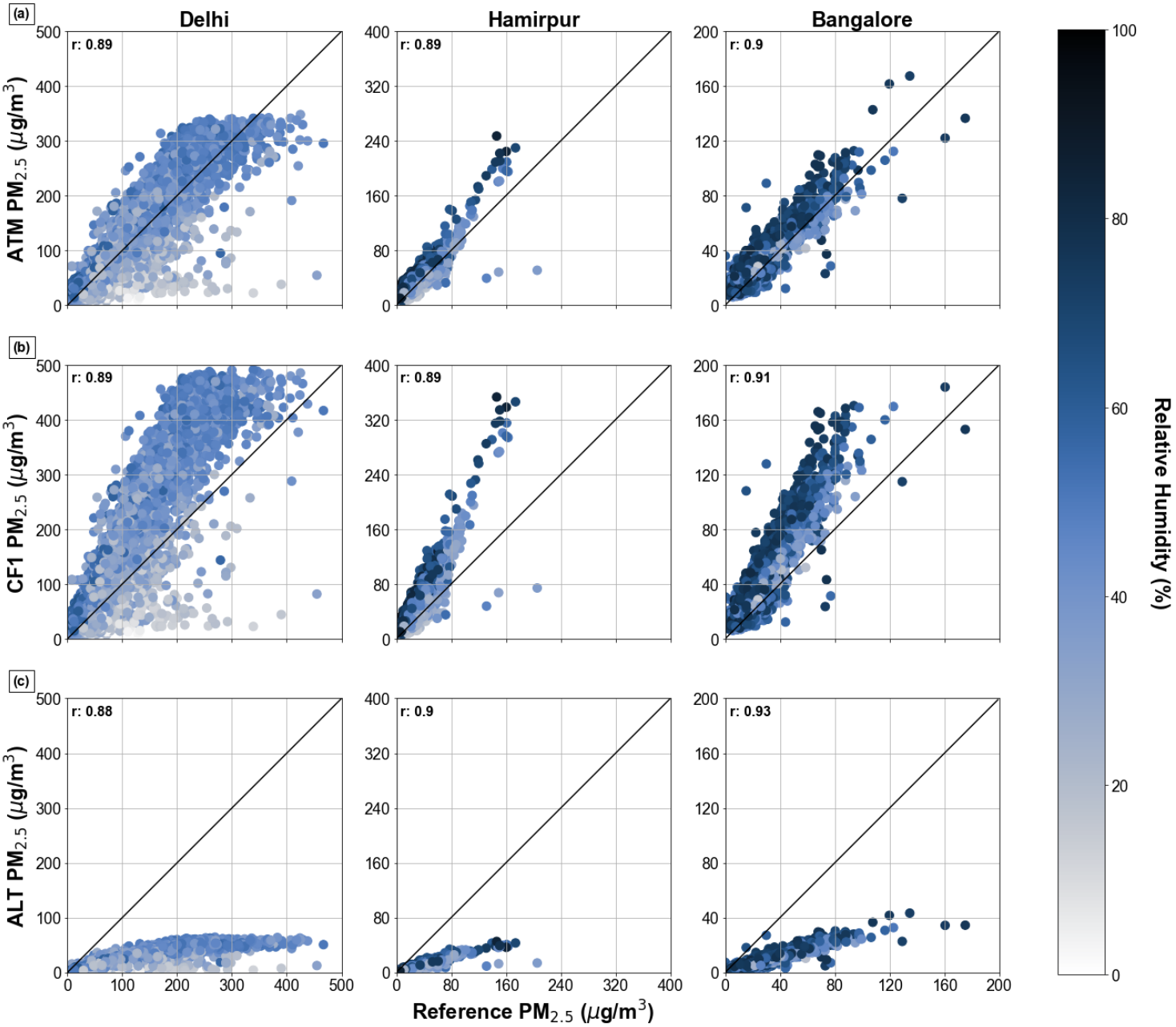

Figure S15: Scatter plots of PurpleAir raw performance versus collocated reference signal for each hourly block average colored by PurpleAir RH and segregated by site and the three PurpleAir PM2.5 channels ATM (panel [a]), CF1 (panel [b]), and ALT (panel [c]). Correlation coefficient (Pearson's r [r]) is presented on the top left of each subplot. Solid lines represent the 1:1 line in each subplot.

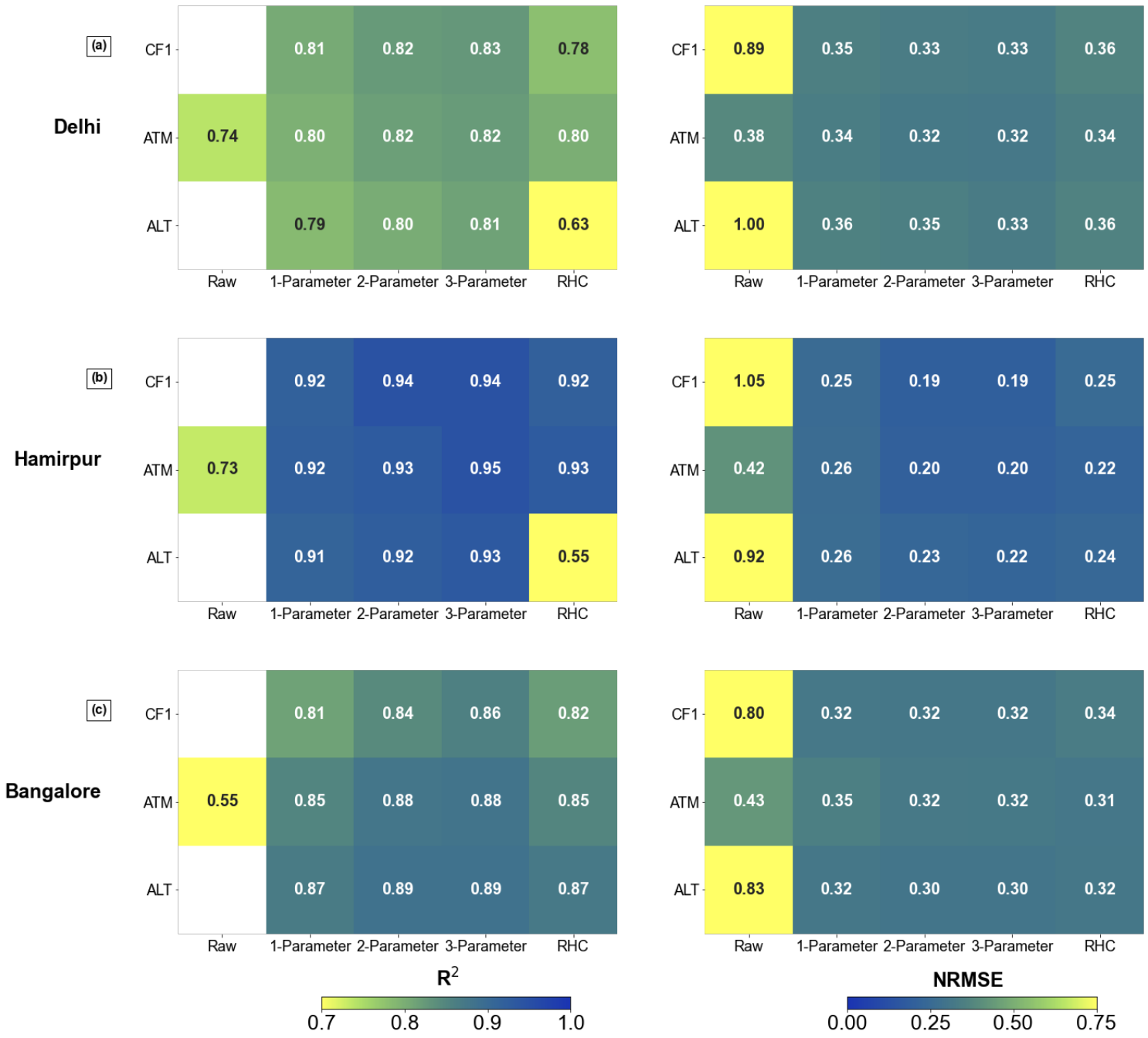

Figure S16: Regression metrics, R^2 (left) and NRMSE (right) for raw data, 1-parameter model, 2-parameter model, 3-parameter model, and the theory-driven hygroscopic growth model (RHC) for each $PM_{2.5}$ channel (CF1, ATM, ALT) for each site (Delhi [panel (a)], Hamirpur [panel (b)], and Bangalore [panel (c)]). The largest improvements are from the raw data to the 1-parameter model, with only marginal improvements in the 3-parameter and RHC models.

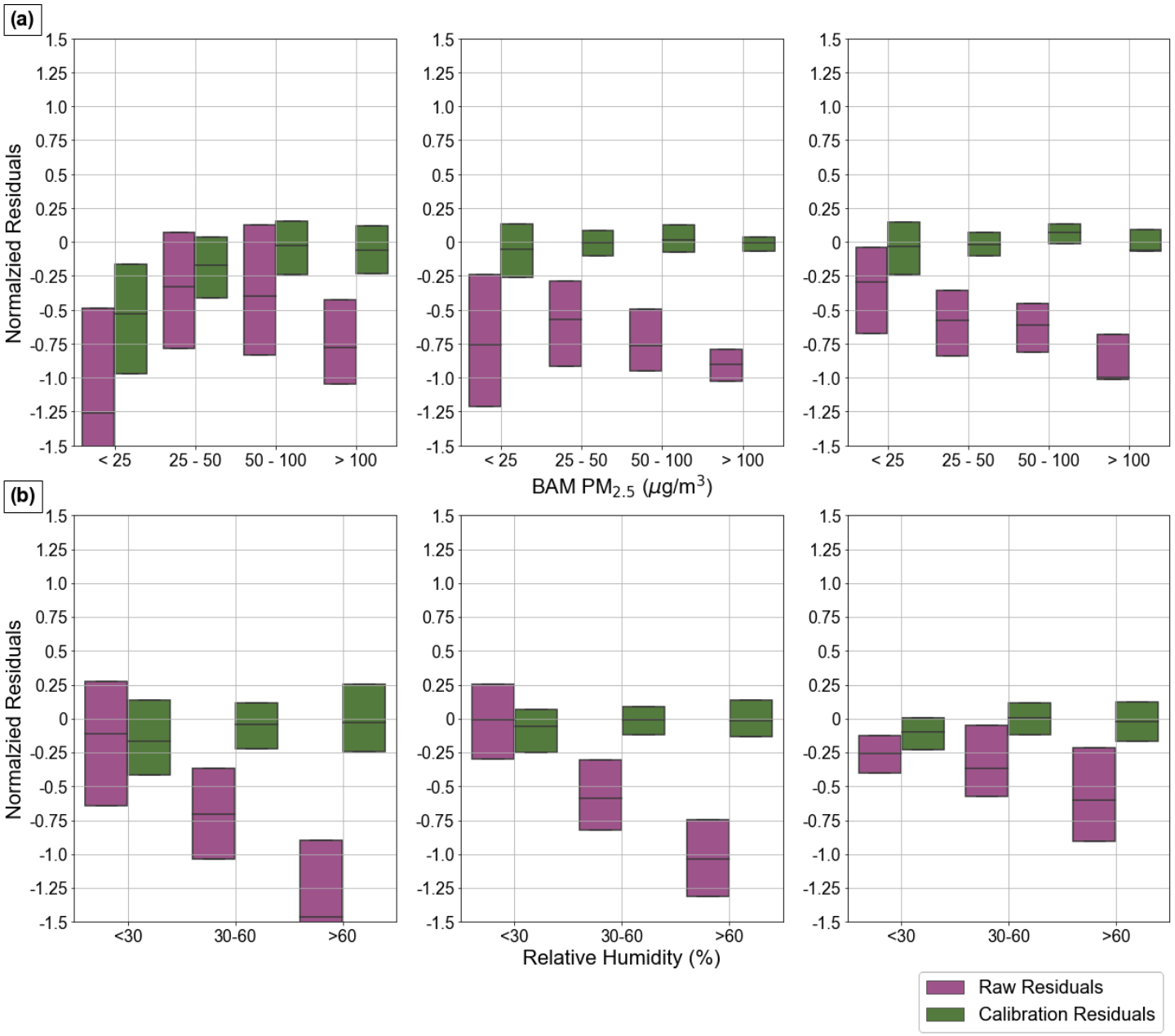

Figure S17. Normalized Residual distributions for the raw PurpleAir data (CF1) and the calibration Models for each site. The panel (a) bins by reference $PM_{2.5}$ and panel (b) bins by PurpleAir relative humidity. The calibrated data effectively reduces bias correlated with all factors.

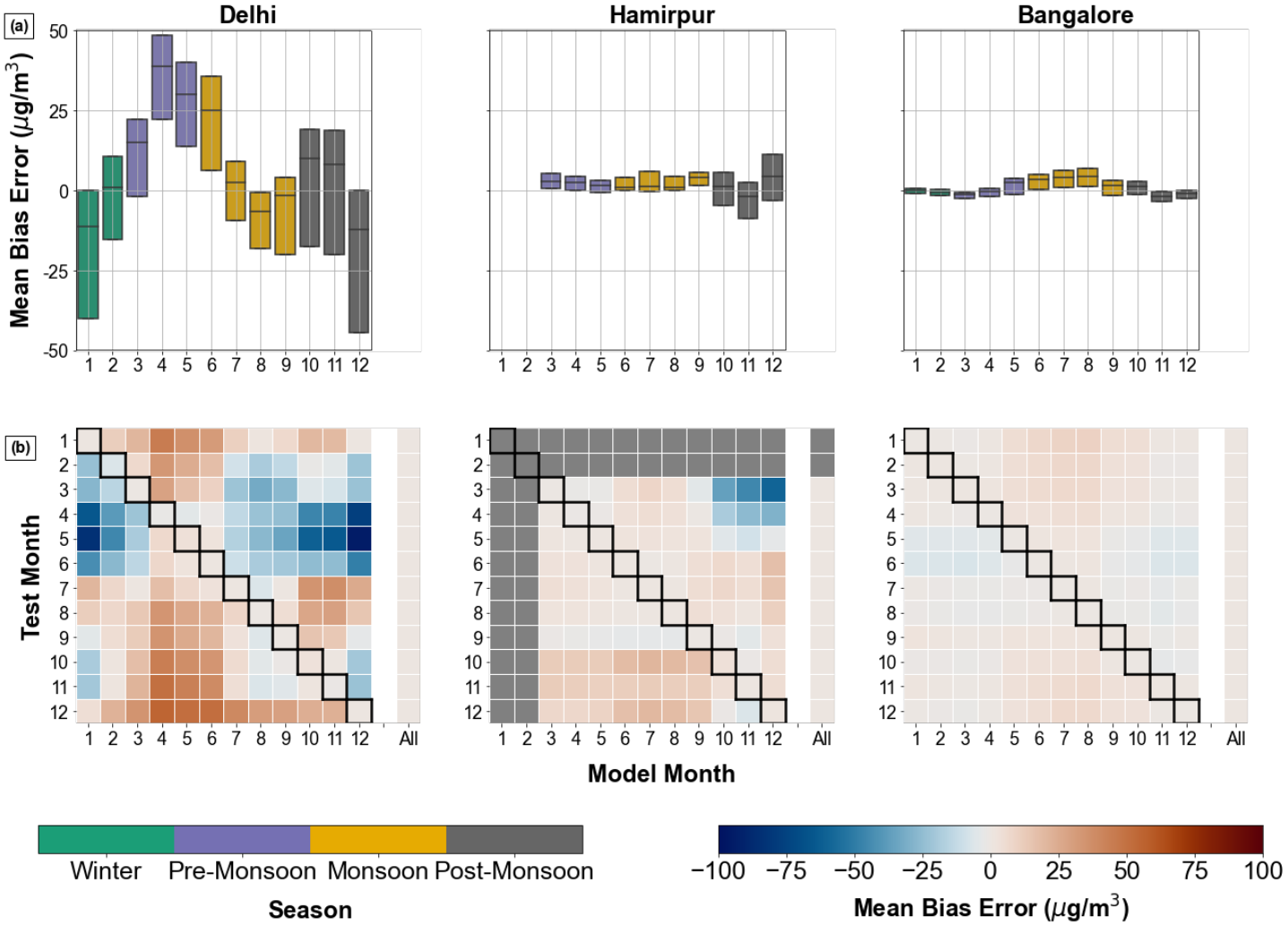

Figure S18. Panel (a) depicts box plots of the distribution of mean bias for a given model starting month of a 4-week rolling ordinary least squares (ROLS) model on all other windows. The results of testing each window on equations 2 – 5 is denoted by "All." The bottom, solid line and top of the boxes represent the 25th, 50th, and 75th percentiles respectively. Panel (b) presents the median Mean Bias of a 4-week ROLS model trained starting in the month (colored by season) on the x-axis and evaluated on all other windows as binned by starting month on the y-axis. The results of testing each window on equations 2 – 4 is denoted by "All." Gray boxes represent months without sufficient data. Models trained in the pre-monsoon (April, May) underpredicted in other seasons, contrary to the typical pattern of overprediction – with an especially strong effect in Delhi due to the higher mass concentrations relative to Bangalore. Finally, as a point of comparison, we present the performance of our long-term calibration in individual months at each site in the column in (b) titled "All". Consistent with our observation that 4-week models trained in a single month generally do not perform as well in other months, we also note that in general, monthly models perform somewhat better for a given month than the long-term model does.

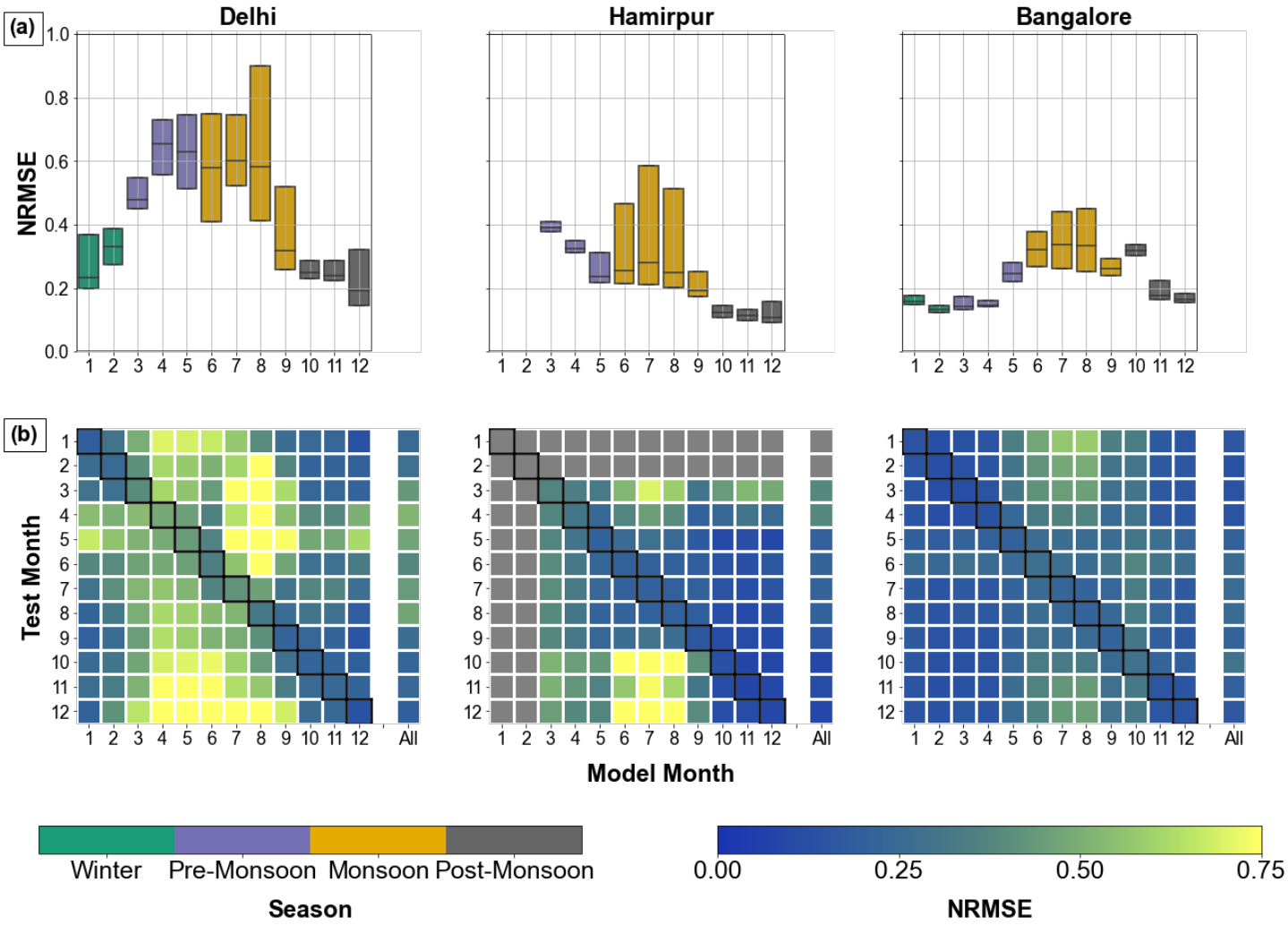

Figure S19. Panel (a) depicts box plots of the distribution of NRMSE for a given model starting month of a 4-week ROLS model on all other windows. The bottom, solid line and top of the boxes represent the 25th, 50th, and 75th percentiles respectively. Panel (b) presents the median NRMSE of a 4-week rolling ordinary least squares (ROLS) model trained starting in the month (colored by season) on the x-axis and evaluated on all other windows as binned by starting month on the y-axis. Gray boxes represent months without sufficient data. The strongest performance is always within the training month (i.e., the boxed diagonal cells), followed by months within the same training season. Models perform worst within the Pre-Monsoon (April, May) and early Monsoon (June, July), regardless of training month. Finally, as a point of comparison, we present the performance of our long-term calibration in individual months at each site in the column in (b) titled "All". Consistent with our observation that 4-week models trained in a single month generally do not perform as well in other months, we also note that in general, monthly models perform somewhat better for a given month than the long-term model does.

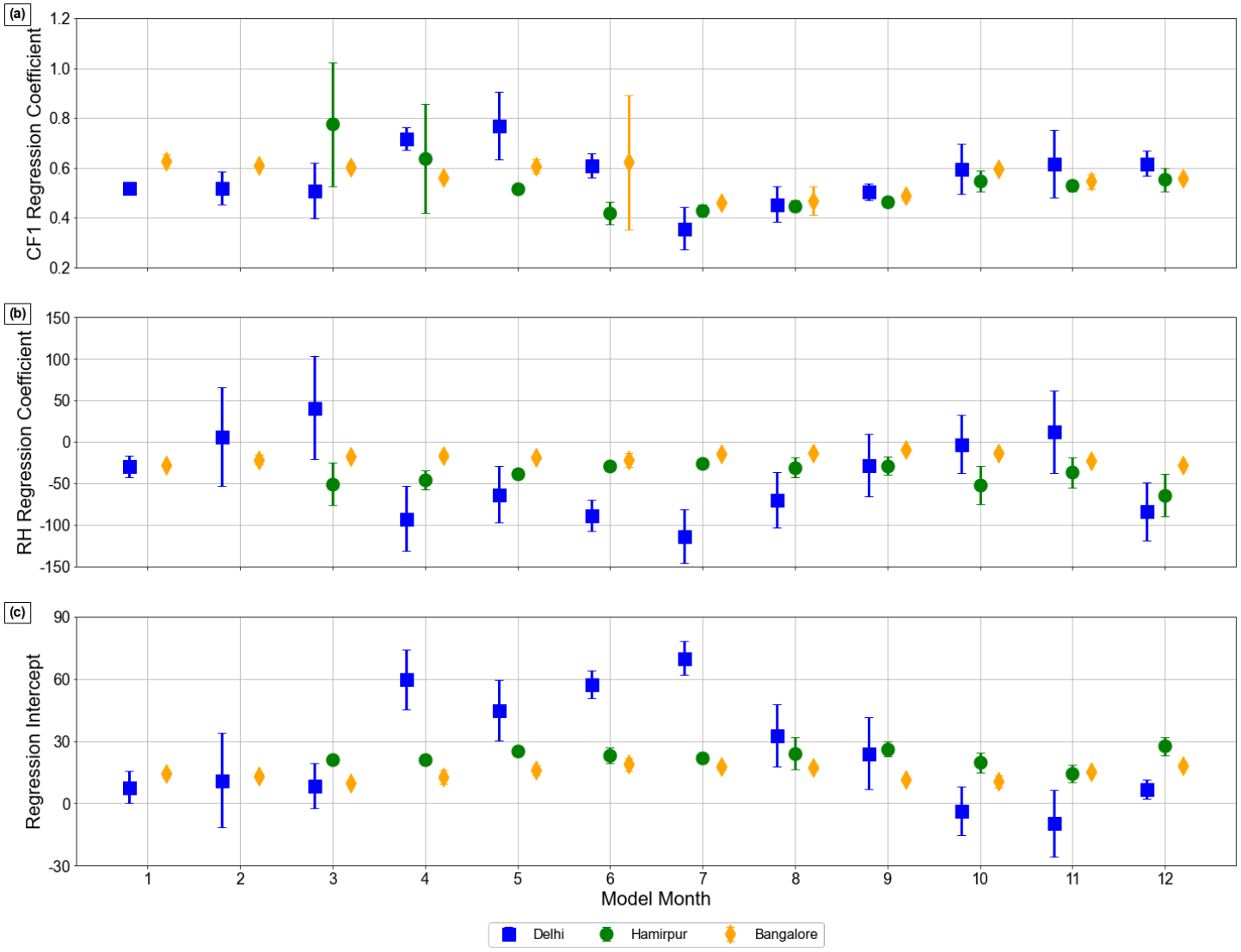

Figure S20. Model results from 4-week rolling ordinary least squares (ROLS) exercise, with model start month on the x-axis and coefficient results on the y-axis. Panel (a), panel (b), and panel (c) show CF1 regression coefficient, RH regression coefficient, regression intercept respectively. Scatter points represent medians with upper and lower bounds defined by the interquartile range (75^{th} percentile minus 25^{th} percentile). All sites demonstrate seasonality, with lower magnitude CF1 coefficients, higher magnitude RH coefficients, and higher magnitude intercept terms during the monsoon (June-Sept). Monthly swings are especially pronounced in Delhi, where the diversity of sources, and dynamic particle size distribution leads to relatively erratic model formulation compared to the more stable Hamirpur and Bangalore coefficient values.

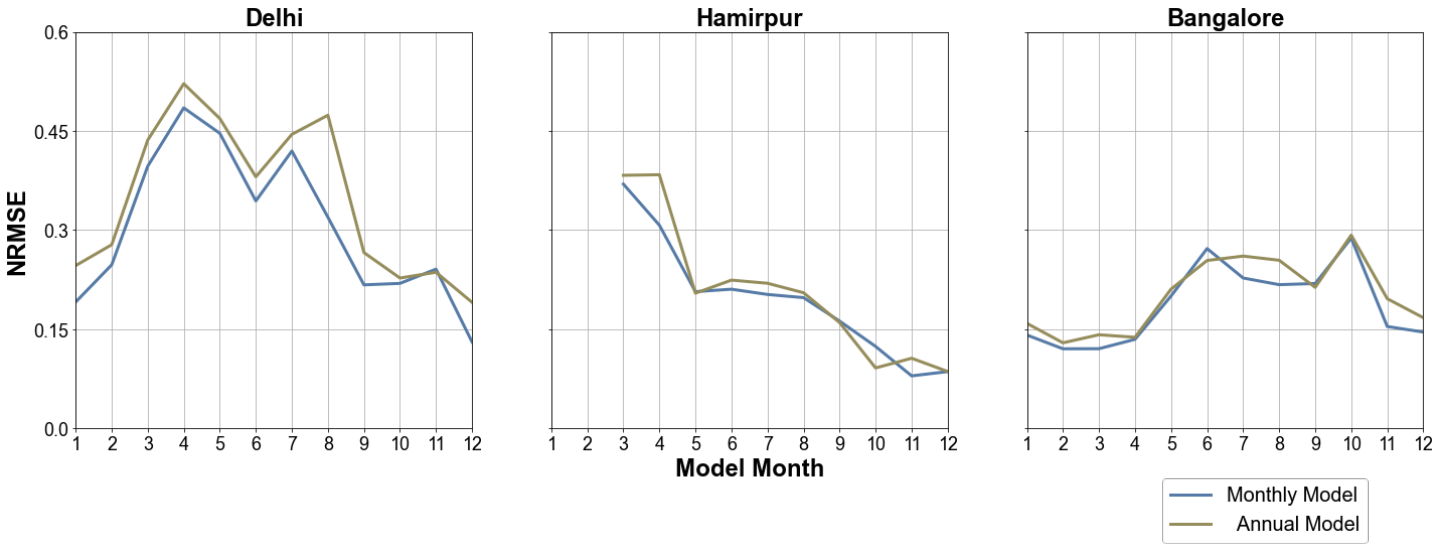

Figure S21. Comparison of calibration model performance (NRMSE) in each month for two calibration approaches: the long-term annual model applied to a given month, as compared to a 4-week rolling ordinary least-squares (ROLS) evaluated in the month for which it is developed. The ROLS models generally have lower NRMSE annual models in any month since they are solely trained in the given month. However, annual model performance is generally within 5% of the ROLS model performance.

References

Schiavina, M., Melchiorri, M., Pesaresi, M., Panagiotis, P., Freire, S., Maffenini, L., Goch, K., Tommasi, P., and Kemper, T.: GHSL Data Package 2022, https://doi.org/10.2760/19817, 2022.