Supplement of Atmos. Meas. Tech., 19, 603–615, 2026
https://doi.org/10.5194/amt-19-603-2026-supplement

Atmospheric Measurement Techniques
Open Access

Supplement of

Enhancing Accuracy of Indoor Air Quality Sensors via Automated Machine Learning Calibration

Juncheng Qian et al.

Correspondence to: Yuqing Dai (y.dai.2@bham.ac.uk) and Zongbo Shi (z.shi@bham.ac.uk)

Supplementary Materials

In this file, AirGriadient is denoted as AG, and AtmoCube is denoted as AC.

Table S1. Top 10 AutoML model training statistics for AirGradient sensors

Model_Subset	Model_ID	Test_RMSE	Test_MAE	Test_R^2	Train_RMSE	Train_MAE	Train_R^2	LB_RMSE	LB_MAE	LB_mean_residual_deviance	LB_Rank
Low (<50 µg/m³)	StackedEnsemble_BestOfFamily_1_AutoML_1_20251008_144707	3.452	1.45	0.569	3.668	1.14	0.721	4.562	1.558	20.816	1
Low (<50 µg/m³)	DeepLearning_grid_2_AutoML_1_20251008_144707_model_4	3.414	1.505	0.579	4.185	1.312	0.637	4.607	1.604	21.227	2
Low (<50 µg/m³)	DeepLearning_grid_2_AutoML_1_20251008_144707_model_3	3.673	1.955	0.512	4.825	2.03	0.517	4.642	1.548	21.544	3
Low (<50 µg/m³)	DeepLearning_grid_1_AutoML_1_20251008_144707_model_3	3.765	1.492	0.488	4.682	1.462	0.546	4.649	1.558	21.612	4
Low (<50 µg/m³)	DeepLearning_grid_1_AutoML_1_	3.54	1.724	0.547	4.535	1.584	0.574	4.673	1.946	21.838	5

	20251008_144707_model_2										
Low (<50 µg/m³)	DeepLearning_grid_3_AutoML_1_20251008_144707_model_4	3.491	1.681	0.56	3.937	1.373	0.679	4.683	1.735	21.934	6
Low (<50 µg/m³)	DeepLearning_grid_2_AutoML_1_20251008_144707_model_1	3.59	1.53	0.534	1.352	0.93	0.962	4.685	1.681	21.946	7
Low (<50 µg/m³)	DeepLearning_grid_3_AutoML_1_20251008_144707_model_3	3.583	1.815	0.536	4.691	1.862	0.544	4.689	1.78	21.982	8
Low (<50 µg/m³)	DeepLearning_grid_1_AutoML_1_20251008_144707_model_4	3.553	1.52	0.544	4.532	1.398	0.574	4.698	1.504	22.067	9
Low (<50 µg/m³)	StackedEnsemble_AllModels_1_AutoML_1_20251008_144707	3.41	1.406	0.58	3.463	1.092	0.751	4.749	1.54	22.556	10
High (⩾50	DeepLearni	85.74	67.492	0.73	77.232	61.739	0.862	85.468	67.948	7304.795	1

µg/m³)	ng_grid_3_AutoML_2_20251008_145912_model_3										
High (⩾50 µg/m³)	StackedEnsemble_BestOfFamily_1_AutoML_2_20251008_145912	81.534	62.367	0.755	54.901	45.028	0.93	86.229	65.222	7435.452	2
High (⩾50 µg/m³)	DeepLearning_grid_1_AutoML_2_20251008_145912_model_2	65.676	55.946	0.841	73.635	54.655	0.875	86.233	66.143	7436.123	3
High (⩾50 µg/m³)	DeepLearning_grid_2_AutoML_2_20251008_145912_model_3	84.987	65.181	0.734	85.45	67.694	0.831	87.139	67.996	7593.221	4
High (⩾50 µg/m³)	StackedEnsemble_AllModels_1_AutoML_2_20251008_145912	66.179	56.598	0.839	35.681	29.452	0.971	87.799	66.987	7708.709	5
High (⩾50 µg/m³)	DeepLearning_grid_1_AutoML_2_20251008_145912_mod	95.343	77.827	0.666	43.161	34.733	0.957	88.113	69.388	7763.824	6

	el_1										
High (⩾50 μg/m³)	DeepLearning_grid_3_AutoML_2_20251008_145912_model_1	140.436	84.304	0.274	18.061	13.396	0.992	90.677	72.281	8222.326	7
High (⩾50 μg/m³)	DeepLearning_grid_1_AutoML_2_20251008_145912_model_3	66.326	51.31	0.838	67.328	50.082	0.895	91.621	70.071	8394.498	8
High (⩾50 μg/m³)	DeepLearning_grid_1_AutoML_2_20251008_145912_model_5	112.725	94.296	0.533	90.516	66.322	0.811	92.224	72.416	8505.28	9
High (⩾50 μg/m³)	DeepLearning_grid_1_AutoML_2_20251008_145912_model_4	66.452	52.316	0.838	57.167	40.525	0.925	93.509	74.121	8743.909	10

Table S2. Top 10 AutoML model training statistics for AtmoCube sensors

Model_Subset	Model_ID	Test_RMSE	Test_MAE	Test_R2	Train_RMSE	Train_MAE	Train_R2	LB_rmse	LB_mae	LB_mean_residual_deviance	LB_ Rank
Low (<50 µg/m³)	DeepLearning_grid_3_AutoML_1_20251008_161954_model_1	6.57	2.895	0.803	2.828	1.368	0.972	6.782	2.529	45.989	1
Low (<50 µg/m³)	StackedEnsemble_BestOfFamily_1_AutoML_1_20251008_161954	6.253	2.94	0.822	2.995	1.663	0.969	7.231	2.478	52.285	2
Low (<50 µg/m³)	DeepLearning_grid_2_AutoML_1_20251008_161954_model_4	13.264	3.79	0.197	6.022	1.886	0.874	7.353	2.559	54.068	3
Low (<50 µg/m³)	StackedEnsemble_AllModels_1_AutoML_1_20251008_161954	6.242	2.704	0.822	3.936	1.599	0.946	7.432	2.364	55.233	4
Low (<50 µg/m³)	DeepLearning_grid_1_AutoML_1_20251008_161954_model_3	3.968	1.801	0.928	7.361	2.067	0.812	7.961	2.299	63.38	5
Low (<50 µg/m³)	DeepLearning_grid_1_A	4.318	1.876	0.915	7.481	1.986	0.805	8.109	2.385	65.755	6

	utoML_1_20251008_161954_model_4										
Low (<50 μg/m³)	DeepLearning_grid_2_AutoML_1_20251008_161954_model_1	6.4	2.973	0.813	3.578	1.6	0.955	8.159	2.991	66.577	7
Low (<50 μg/m³)	GBM_grid_1_AutoML_1_20251008_161954_model_1	6.54	2.881	0.805	6.186	2.093	0.867	8.213	3.02	67.453	8
Low (<50 μg/m³)	DeepLearning_grid_3_AutoML_1_20251008_161954_model_3	15.851	14.642	-0.147	16.739	15.26	0.026	8.27	3.637	68.386	9
Low (<50 μg/m³)	DeepLearning_grid_3_AutoML_1_20251008_161954_model_4	5.592	2.483	0.857	5.623	1.889	0.89	8.296	2.648	68.831	10
High (≥50 μg/m³)	StackedEnsemble_AllModels_1_AutoML_2_20251008_162752	121.022	100.61	0.629	44.53	34.225	0.939	89.695	74.541	8045.133	1
High (≥50	DeepLearnin	136.761	118.11	0.526	73.709	60.697	0.834	94.745	80.044	8976.538	2

µg/m³)	g_grid_1_AutoML_2_20251008_162752_model_2										
High (⩾50 µg/m³)	DeepLearning_grid_3_AutoML_2_20251008_162752_model_3	129.325	109.921	0.576	56.071	45.928	0.904	96.485	78.438	9309.44	3
High (⩾50 µg/m³)	StackedEnsemble_BestOfFamily_1_AutoML_2_20251008_162752	129.279	110.052	0.577	64.196	52.846	0.874	99.53	79.996	9906.26	4
High (⩾50 µg/m³)	DeepLearning_grid_2_AutoML_2_20251008_162752_model_3	129.605	106.901	0.575	48.165	36.582	0.929	100.524	82.119	10105.011	5
High (⩾50 µg/m³)	DeepLearning_grid_1_AutoML_2_20251008_162752_model_3	114.367	96.043	0.669	68.377	54.058	0.857	102.895	80.471	10587.304	6
High (⩾50 µg/m³)	DeepLearning_grid_3_AutoML_2_20251008_162752_model_	137.281	114.282	0.523	35.036	20.896	0.963	106.762	87.51	11398.144	7

	1										
High (⩾50 μg/m³)	GBM_lr_annealing_selection_AutoML_2_20251008_162752_select_model	139.195	111.078	0.509	67.122	53.151	0.862	109.135	87.475	11910.463	8
High (⩾50 μg/m³)	DeepLearning_grid_1_AutoML_2_20251008_162752_model_4	114.042	97.226	0.671	57.249	40.911	0.9	111.219	83.622	12369.733	9
High (⩾50 μg/m³)	GBM_grid_1_AutoML_2_20251008_162752_model_7	139.607	101.849	0.506	17.735	14.743	0.99	115.542	91.521	13350.025	10

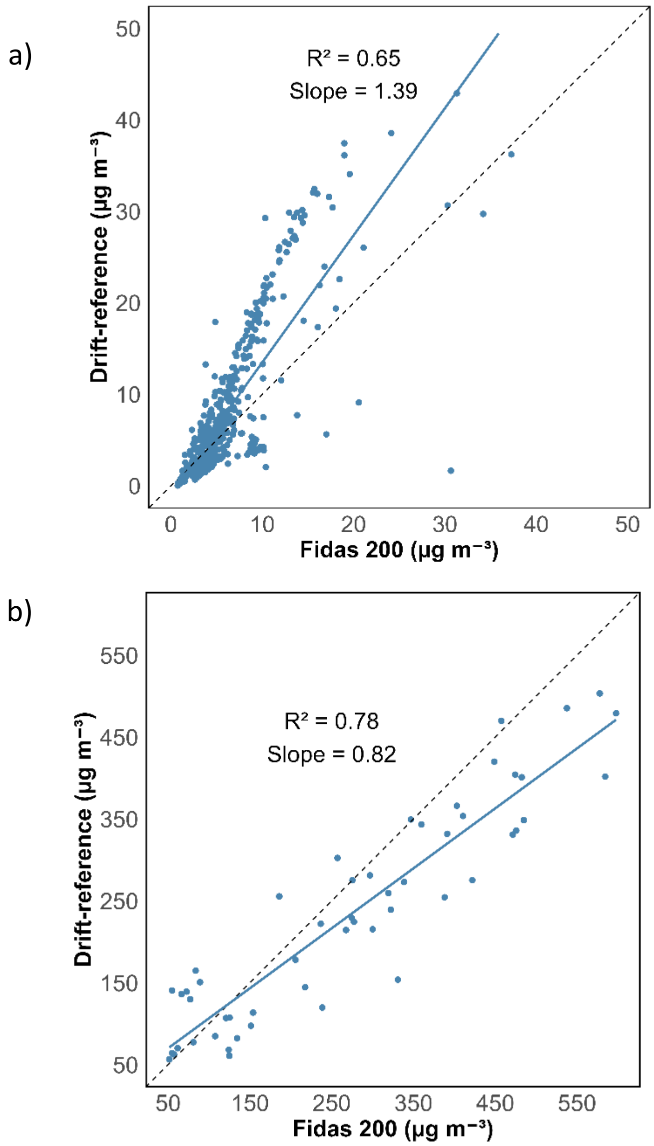


Figure S1. Relationship between all drift-reference sensors average concentration and Fidas 200, a) below 50 µg m^{-3}, b) above 50 µg m^{-3}.

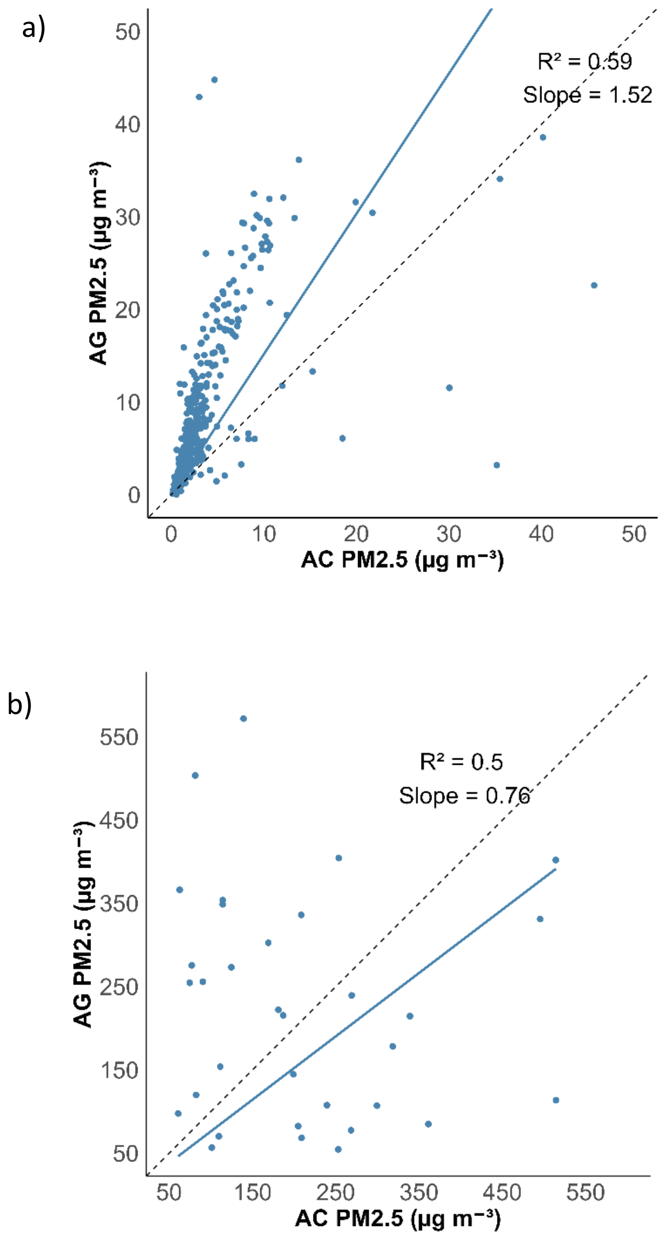


Figure S2. Inter-relationships between AG and AC sensors, a) below 50 µg m^{-3}, b) above 50 µg m^{-3}.

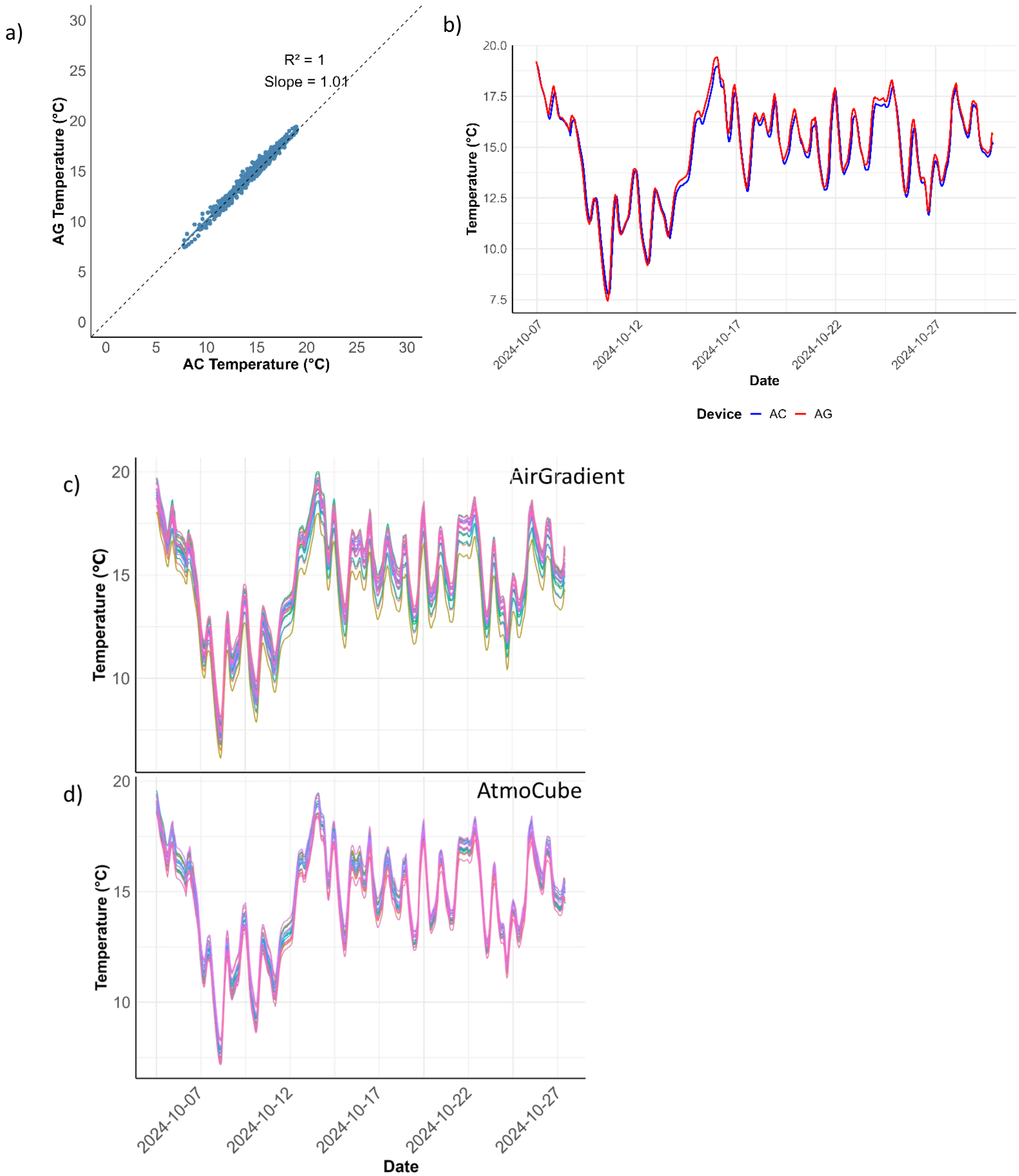


Figure S3. Inter-relationship of measured temperature of AC and AG, a) scatter plot, b) timeseries, c) timeseries of AG temperature measurements, d) timeseries of AC temperature measurements

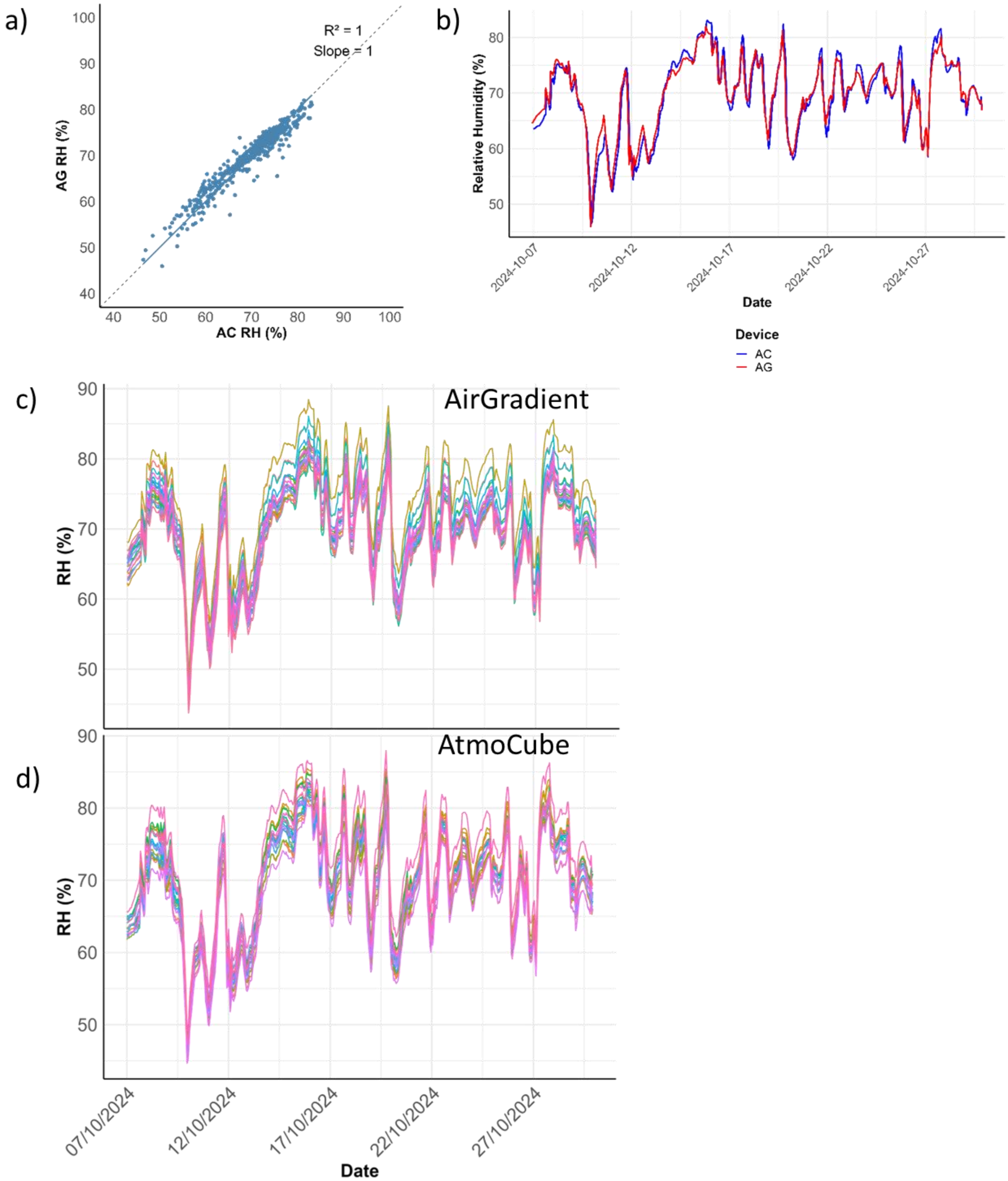


Figure S4. Inter-relationship of measured relative humidity of AC and AG, a) scatter plot, b) timeseries, c) timeseries of AG relative humidity measurements, d) timeseries of AC relative humidity measurements

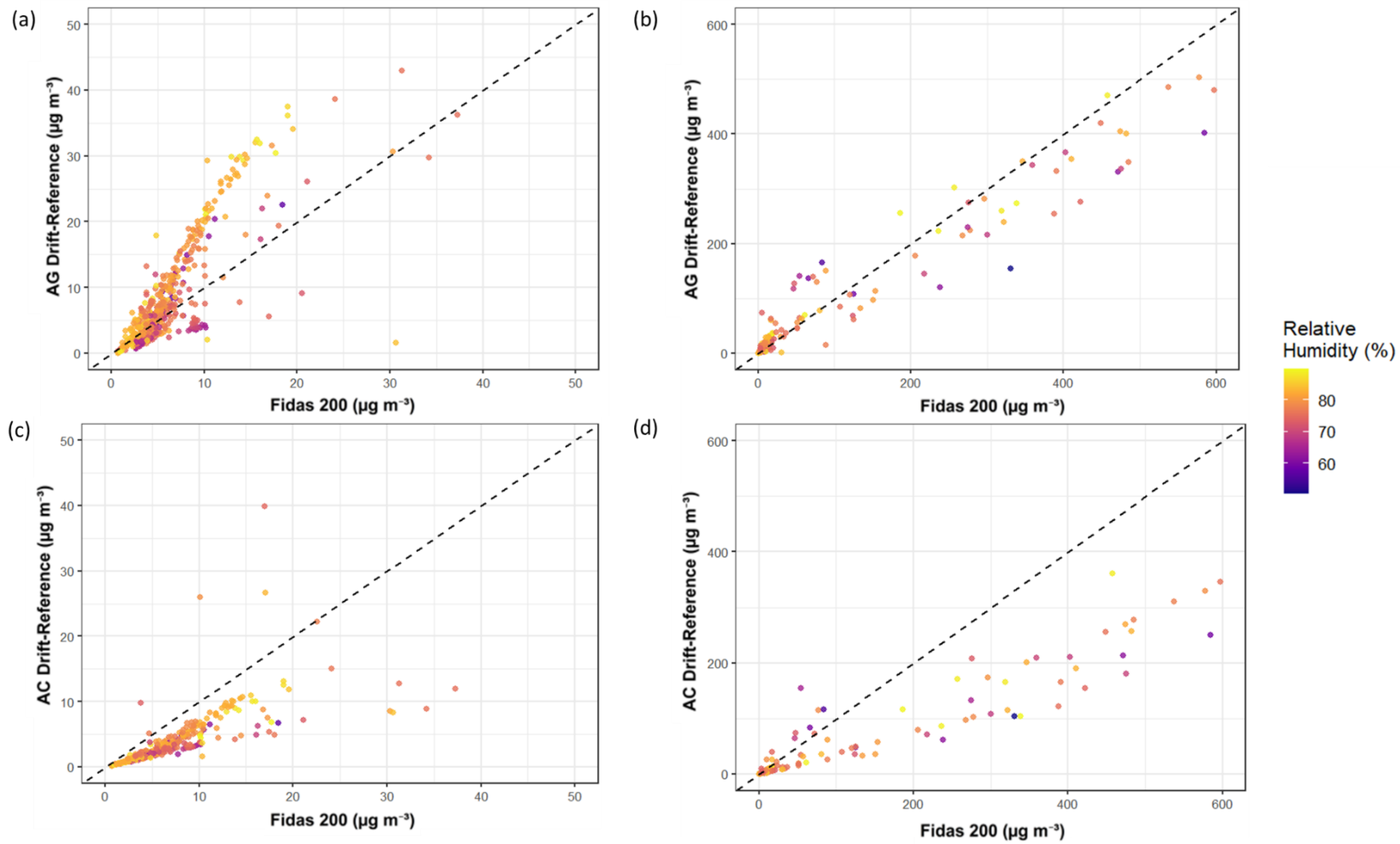


Figure S5. Pre-calibration $PM_{2.5}$ readings with relative humidity levels, a) AG below 50 µg m^{-3}, b) AG above 50 µg m^{-3}, c) AC below 50 µg m^{-3}, and d) AC above 50 µg m^{-3}.

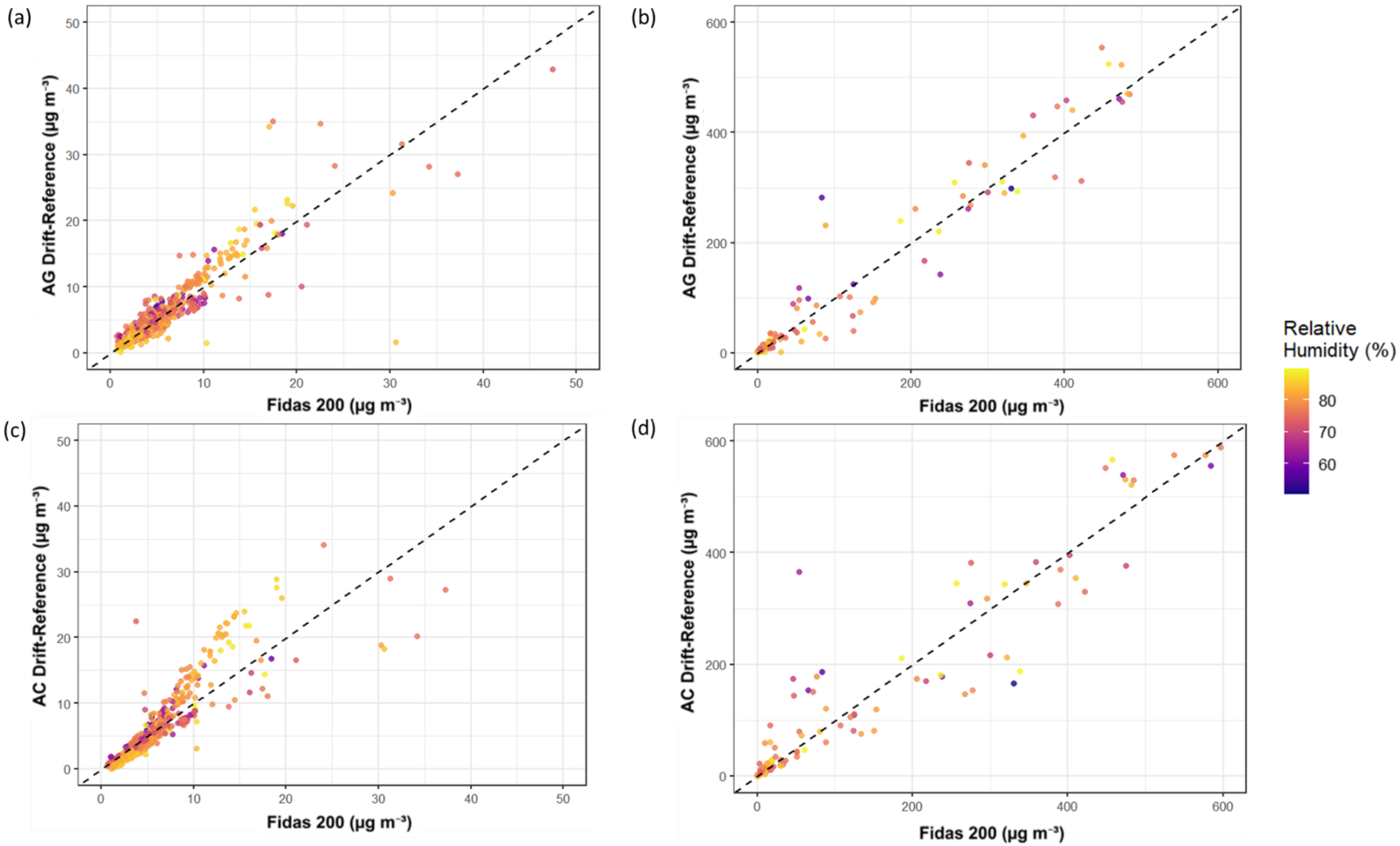
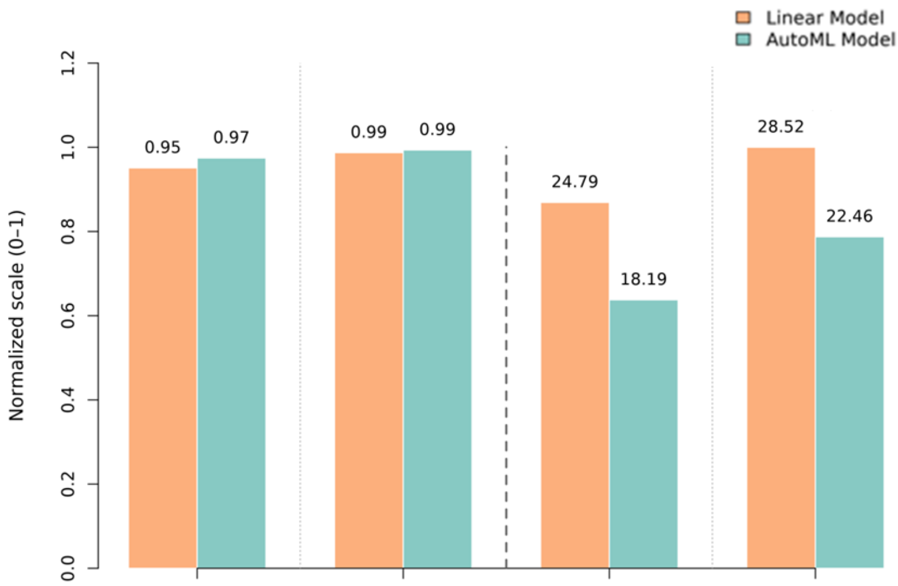


Figure S6. Post-calibration $PM_{2.5}$ readings with relative humidity levels, a) AG below 50 μg m^{-3}, b) AG above 50 μg m^{-3}, c) AC below 50 μg m^{-3}, and d) AC above 50 μg m^{-3}.

(a)



(b)

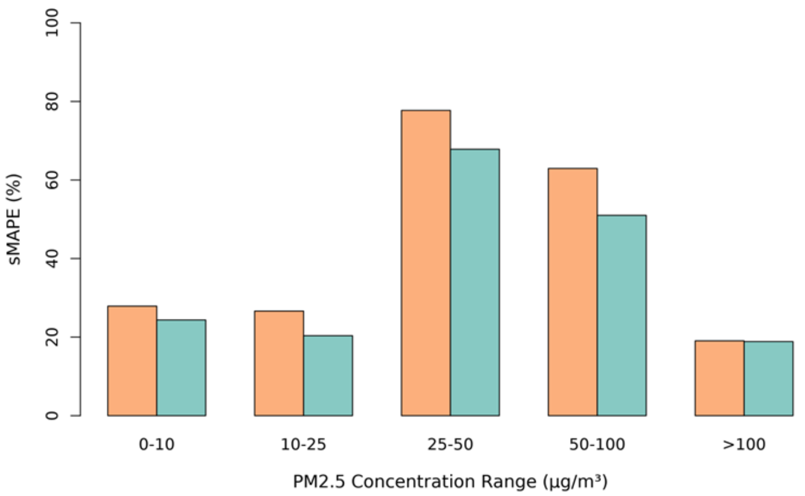


Figure S7. Comparison of AutoML model and the multivariate linear regression model for AirGradient ONE, a) performance metrics, and b) error by concentration range.

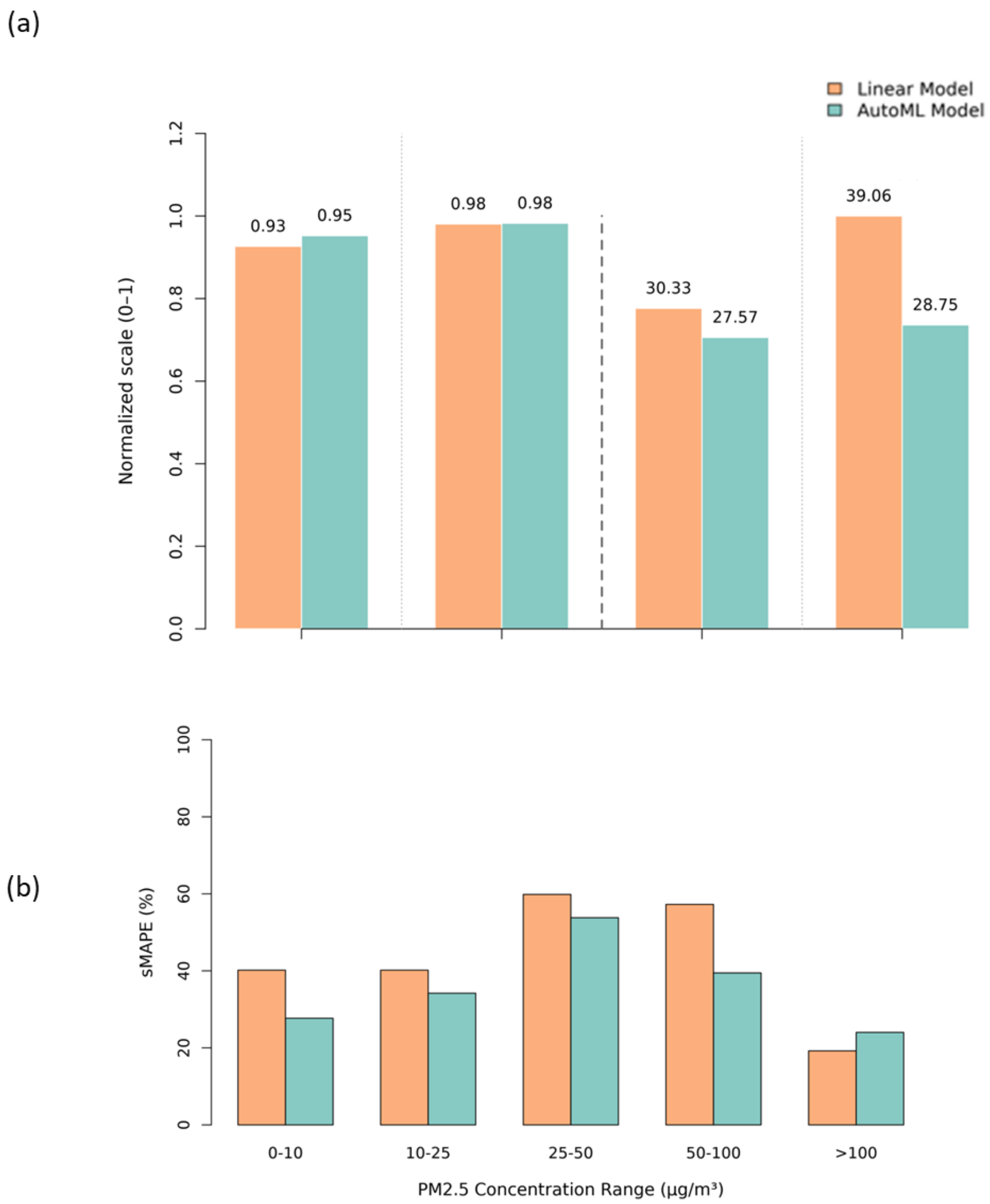


Figure S8. Comparison of AutoML model and the multivariate linear regression model for AtmoCube, a) performance metrics, and b) error by concentration range.