Supplement of Atmos. Meas. Tech., 16, 4723–4740, 2023
https://doi.org/10.5194/amt-16-4723-2023-supplement

Atmospheric Measurement Techniques
Open Access

Supplement of

Development of low-cost air quality stations for next-generation monitoring networks: calibration and validation of NO_2 and O_3 sensors

Alice Cavaliere et al.

Correspondence to: Alice Cavaliere (alice.cavaliere@ibe.cnr.it)

Table S1. Summary of pre–deployment calibration and field validation procedures performed in the study.

City	Lat. (deg N)	Lon. (deg E)	Time period	Time step	Process	Poll.	AQ id	No. records	RH (%)	intT (°C)	Ref. ($\mu g m^{-3}$)
Florence	43°47’52”	11°11’31”	2017/07/19 2017/09/30	3min	Calibration	O3	AQ1	19223	17–96	21–49	0–220
							AQ2	17275	17–99	21–48	1–220
						NO2	AQ1	19223	17–96	21–49	2–114
							AQ2	17275	17–99	21–48	2–114
Montale	43°54’57”	11°00’26”	2018/06/19 2019/08/22	1H	Validation	O3	AQ1	7344	10–98	1–58	2–185
							AQ2	9303	0–98	1–54	2–185
						NO2	AQ1	7383	10–98	1–58	0–88
							AQ2	9340	0–98	1–54	0–88

Table S2. Hyperparameters used for non–parametric models. For any hyperparameters not specified in the table, refer to the default ones of the library scikit–learn version 1.2.2.

Pollutant	Model	AQ id	Hyperparameters
O3	**RF**	AQ1,AQ2	$trees = 100$
	GB	AQ1,AQ2	$boostingstage = 100$
	SVM	AQ1	$C = 100$
		AQ2	$C = 10$
	RBF	AQ1	$\epsilon = 0.1$
		AQ2	$\epsilon = 1$
	MRF	AQ1,AQ2	$trees = 100$, $maxdepth = 10$
	MGB	AQ1,AQ2	$boostingstage = 100$
NO2	**RF**	AQ1,AQ2	$trees = 100$
	GB	AQ1,AQ2	$boostingstage = 100$
	SVM	AQ1	$C = 10$
		AQ2	$C = 100$
	RBF	AQ1	$\epsilon = 1$
		AQ2	$\epsilon = 0.01$
	MRF	AQ1,AQ2	$trees = 100$, $maxdepth = 10$
	MGB	AQ1,AQ2	100 boosting stage

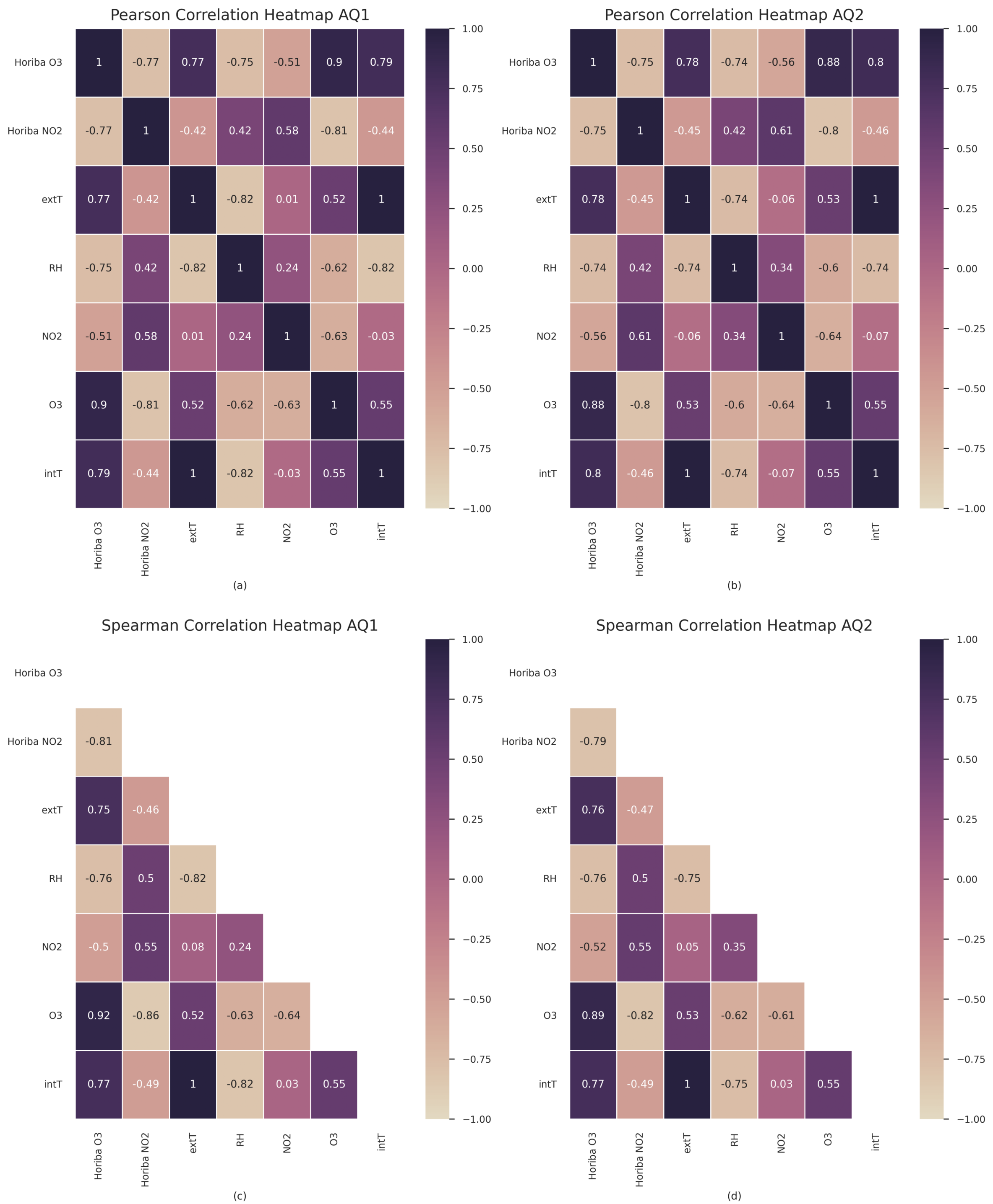

Figure S3. Correlation matrices for AQ1 and AQ2 calibration: Pearson correlation for AQ1 (a) and AQ2 (b); Spearman correlation for AQ1 (c) and AQ2 (d).

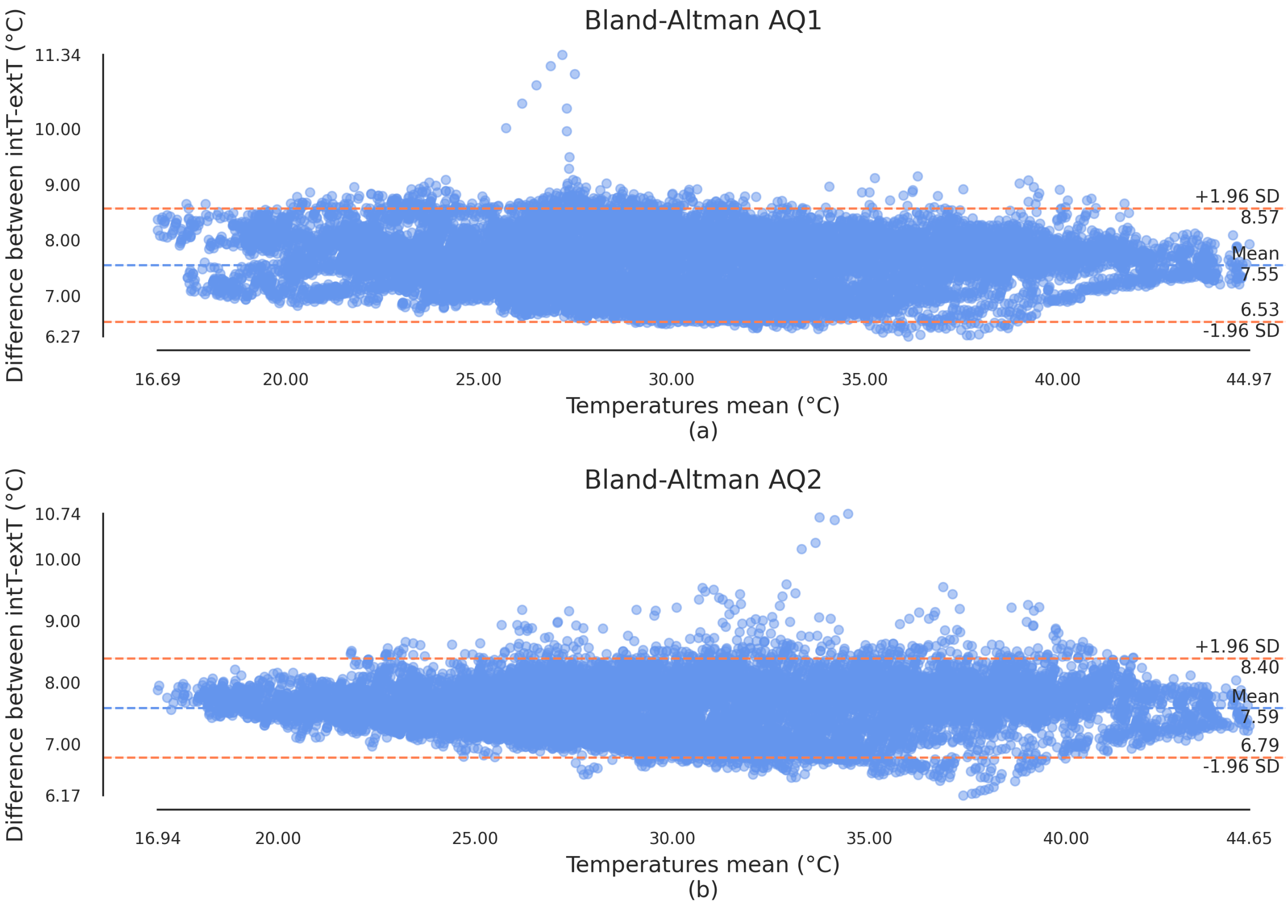

Figure S4. Bland–Altman plots of extT and intT for AQ1 (a) and AQ2 (b). The red lines indicate the limit of agreement (± 1.96 SD) of temperature differences, whilst the blue lines show the average difference between the temperatures.

Table S5. Variance inflation factors (VIF) for the main covariates set.

Pollutant	AQ id	Variables	Intercept	O3	NO2	intT	extT	RH
O3	AQ1	O3,intT	47.38	1.44	–	1.44	–	–
		O3,intT,RH	299.06	1.63	–	3.08	–	3.45
		O3,intT,extT,RH	441.18	1.95	–	151.19	147.78	3.55
		O3,NO2,intT,RH	304.88	3.13	2.25	4.00	–	3.61
		O3,NO2,intT,extT,RH	465.46	3.24	2.35	153.54	154.09	3.77
	AQ2	O3,intT	54.05	1.44	–	1.44	–	–
		O3,intT,RH	210.67	1.63	–	2.30	–	2.52
		O3,intT,extT,RH	472.01	1.92	–	237.86	232.71	2.57
		O3,NO2,intT,RH	221.90	2.97	2.24	3.03	–	2.72
		O3,NO2,intT,extT,RH	484.17	3.25	2.24	238.39	232.72	2.77
NO2	AQ1	NO2,intT	36.86	–	1.00	1.00	–	–
		NO2,intT,RH	241.53	–	1.17	3.37	–	3.58
		NO2,intT,extT,RH	434.86	–	1.42	144.55	149.03	3.70
		NO2,O3,intT,RH	304.88	3.13	2.25	4.00	–	3.61
		NO2,O3,intT,extT,RH	465.46	3.24	2.35	153.54	154.09	3.77
	AQ2	NO2,intT	37.46	–	1.01	1.01	–	–
		NO2,intT,RH	149.67	–	1.23	2.40	–	2.71
		NO2,intT,extT,RH	472.57	–	1.32	211.00	212.93	2.74
		NO2,O3,intT,RH	221.90	2.97	2.24	3.03	–	2.72
		NO2,O3,intT,extT,RH	484.17	3.25	2.24	238.39	232.72	2.77

Table S6. Summary of MLR O3 models for AQ1 and AQ2.

			Coefficient					Stat.					
Pollutant	AQ id	Variables	$\beta0$	$\beta1$	$\beta2$	$\beta3$	$\beta4$	R^2	RMSE	AIC	$AdjR^2$	MAE	MBE
O3	AQ1	O3	–146.10	0.40	–	–	–	0.81	16.92	1.09E+05	0.81	13.42	–0.28
		intT	–109.92	5.46	–	–	–	0.61	24.33	1.19E+05	0.62	19.56	0.21
		RH	165.02	–1.71	–	–	–	0.57	25.61	1.21E+05	0.57	20.35	–0.17
		NO2	97.15	–0.31	–	–	–	0.27	33.63	1.27E+05	0.27	27.42	–0.46
		O3,intT	–186.58	0.30	2.88	–	–	0.94	9.92	9.52E+04	0.94	7.39	–0.06
		O3,RH	–61.47	0.32	–0.71	–	–	0.88	13.61	1.04E+05	0.88	10.69	–0.26
		NO2,O3	–162.73	0.05	0.43	–	–	0.82	16.74	1.09E+05	0.82	13.24	–0.23
		RH,intT	–12.71	–0.71	3.70	–	–	0.65	23.07	1.17E+05	0.65	18.21	0.09
		NO2,intT	–84.20	–0.29	5.36	–	–	0.86	14.52	1.05E+05	0.86	11.35	–0.11
		NO2,RH	170.74	–0.21	–1.52	–	–	0.69	21.88	1.17E+05	0.69	17.50	–0.39
		O3,RH,intT	–193.19	0.30	0.04	2.97	–	0.94	9.93	9.52E+04	0.94	7.39	–0.06
		NO2,O3,intT	–159.76	–0.10	0.23	3.44	–	0.95	8.69	9.18E+04	0.95	6.39	–0.12
		NO2,O3,RH	–64.32	0.01	0.32	–0.71	–	0.88	13.60	1.04E+05	0.88	10.69	–0.26
		NO2,RH,intT	–87.77	–0.29	0.03	5.42	–	0.86	14.52	1.05E+05	0.86	11.36	–0.11
		NO2,O3,RH,intT	–180.76	–0.11	0.23	0.15	3.79	0.95	8.62	9.14E+04	0.95	6.30	–0.10
	AQ2	O3	–140.53	0.31	–	–	–	0.77	17.58	9.93E+04	0.77	14.11	0.14
		intT	–90.68	5.04	–	–	–	0.64	22.12	1.04E+05	0.64	18.07	–0.24
		RH	148.13	–1.37	–	–	–	0.55	24.62	1.07E+05	0.55	19.48	–0.29
		NO2	101.33	–0.34	–	–	–	0.31	30.43	1.12E+05	0.31	24.74	0.23
		O3,intT	–174.35	0.23	2.86	–	–	0.91	10.86	8.83E+04	0.91	8.18	–0.05
		O3,RH	–59.90	0.24	–0.61	–	–	0.84	14.77	9.51E+04	0.84	11.96	–0.03
		NO2,O3	–139.86	–0.00	0.31	–	–	0.77	17.58	9.93E+04	0.77	14.11	0.14
		RH,intT	–9.43	–0.60	3.52	–	–	0.69	20.50	1.03E+05	0.69	16.33	–0.30
		NO2,intT	–62.83	–0.30	4.79	–	–	0.89	12.22	9.07E+04	0.89	9.52	–0.08
		NO2,RH	150.72	–0.21	–1.15	–	–	0.65	21.64	1.04E+05	0.65	17.38	–0.13
		O3,RH,intT	–155.80	0.22	–0.12	2.62	–	0.91	10.76	8.81E+04	0.91	8.18	–0.07
		NO2,O3,intT	–129.74	–0.16	0.14	3.51	–	0.94	8.59	8.30E+04	0.95	6.51	–0.04
		NO2,O3,RH	–53.73	–0.02	0.24	–0.62	–	0.84	14.76	9.51E+04	0.84	11.96	–0.03
		NO2,RH,intT	–60.82	–0.30	–0.02	4.76	–	0.89	12.21	9.07E+04	0.89	9.52	–0.09
		NO2,O3,RH,intT	–133.43	–0.16	0.14	0.03	3.58	0.94	8.58	8.29E+04	0.95	6.50	–0.03

Table S7. Summary of MLR NO2 models for AQ1 and AQ2.

Pollutant	AQ id	Variables	Coefficient					Stat					
			$\beta0$	$\beta1$	$\beta2$	$\beta3$	$\beta4$	R^2	RMSE	AIC	$AdjR^2$	MAE	MBE
NO2	AQ1	NO2	13.03	0.15	–	–	–	0.34	14.22	1.05E+05	0.34	10.91	0.09
		intT	70.28	–1.36	–	–	–	0.19	15.72	1.07E+05	0.19	12.14	–0.16
		RH	2.19	0.42	–	–	–	0.18	15.81	1.08E+05	0.18	12.25	–0.07
		O3	111.52	–0.16	–	–	–	0.66	10.22	9.61E+04	0.66	7.74	–0.01
		NO2,intT	57.18	0.15	–1.30	–	–	0.52	12.14	1.01E+05	0.52	9.17	0.01
		NO2,RH	–1.44	0.13	0.30	–	–	0.42	13.28	1.03E+05	0.42	9.99	0.08
		NO2,O3	101.09	0.03	–0.14	–	–	0.67	10.09	9.57E+04	0.67	7.64	0.01
		RH,intT	47.73	0.17	–0.95	–	–	0.20	15.59	1.07E+05	0.20	12.02	–0.13
		O3,intT	111.23	–0.16	0.02	–	–	0.66	10.22	9.61E+04	0.66	7.74	–0.01
		O3,RH	127.09	–0.18	–0.13	–	–	0.67	10.10	9.57E+04	0.67	7.67	–0.01
		NO2,RH,intT	89.29	0.16	–0.24	–1.90	–	0.53	11.96	1.00E+05	0.53	9.09	–0.02
		NO2,O3,intT	100.92	0.04	–0.13	–0.19	–	0.67	10.06	9.57E+04	0.67	7.61	0.01
		NO2,O3,RH	116.25	0.02	–0.16	–0.11	–	0.67	10.01	9.55E+04	0.67	7.61	0.01
		O3,RH,intT	150.78	–0.17	–0.27	–0.54	–	0.68	9.97	9.53E+04	0.67	7.57	–0.05
		NO2,O3,RH,intT	144.78	0.05	–0.14	–0.32	–0.93	0.69	9.68	9.45E+04	0.69	7.36	–0.03
	AQ2	NO2	11.81	0.16	–	–	–	0.38	12.86	9.22E+04	0.38	9.83	–0.09
		intT	66.88	–1.31	–	–	–	0.21	14.45	9.49E+04	0.21	10.91	0.07
		RH	4.94	0.35	–	–	–	0.19	14.67	9.54E+04	0.19	11.09	0.08
		O3	113.00	–0.13	–	–	–	0.64	9.79	8.58E+04	0.64	7.35	–0.04
		NO2,intT	52.45	0.16	–1.19	–	–	0.55	10.91	8.85E+04	0.55	8.19	–0.01
		NO2,RH	3.19	0.14	0.20	–	–	0.43	12.28	9.12E+04	0.43	9.25	–0.03
		NO2,O3	97.76	0.05	–0.11	–	–	0.66	9.53	8.52E+04	0.66	7.15	–0.06
		RH,intT	47.54	0.14	–0.95	–	–	0.23	14.28	9.47E+04	0.23	10.73	0.08
		O3,intT	114.00	–0.13	–0.08	–	–	0.64	9.78	8.58E+04	0.64	7.34	–0.04
		O3,RH	123.40	–0.14	–0.08	–	–	0.64	9.72	8.56E+04	0.64	7.32	–0.07
		NO2,RH,intT	77.38	0.17	–0.20	–1.67	–	0.57	10.71	8.79E+04	0.57	8.07	–0.04
		NO2,O3,intT	96.78	0.06	–0.10	–0.34	–	0.67	9.41	8.50E+04	0.67	7.04	–0.04
		NO2,O3,RH	107.47	0.04	–0.12	–0.07	–	0.66	9.48	8.51E+04	0.66	7.14	–0.08
		O3,RH,intT	137.80	–0.13	–0.15	–0.39	–	0.65	9.63	8.53E+04	0.65	7.26	–0.06
		NO2,O3,RH,intT	126.78	0.08	–0.10	–0.23	–0.87	0.69	9.07	8.40E+04	0.69	6.83	–0.08

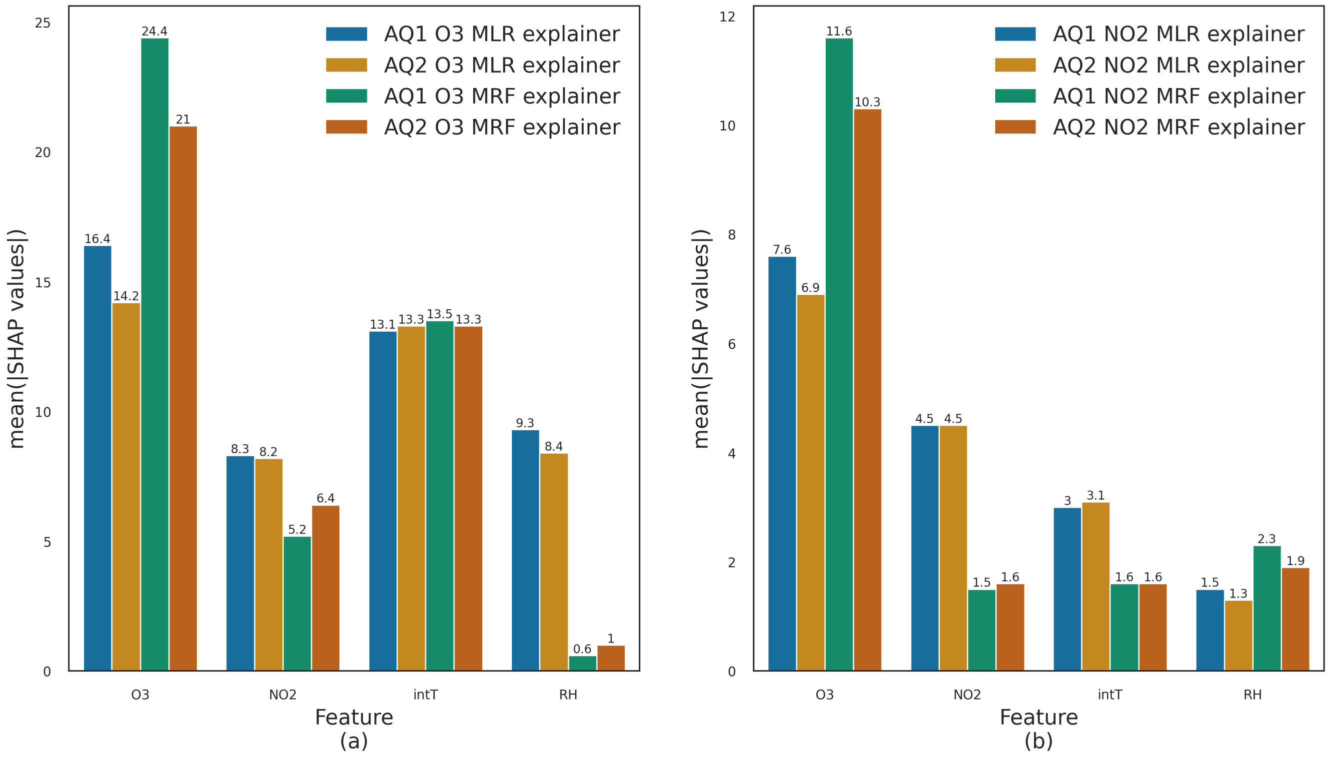

Figure S8. Mean absolute SHAP values of all features for MLR and MRF models for O3 and NO2 concentrations at AQ1 and AQ2 stations. The effect of intT, RH, O3 and NO2 raw data on the overall output of the training model is also shown.

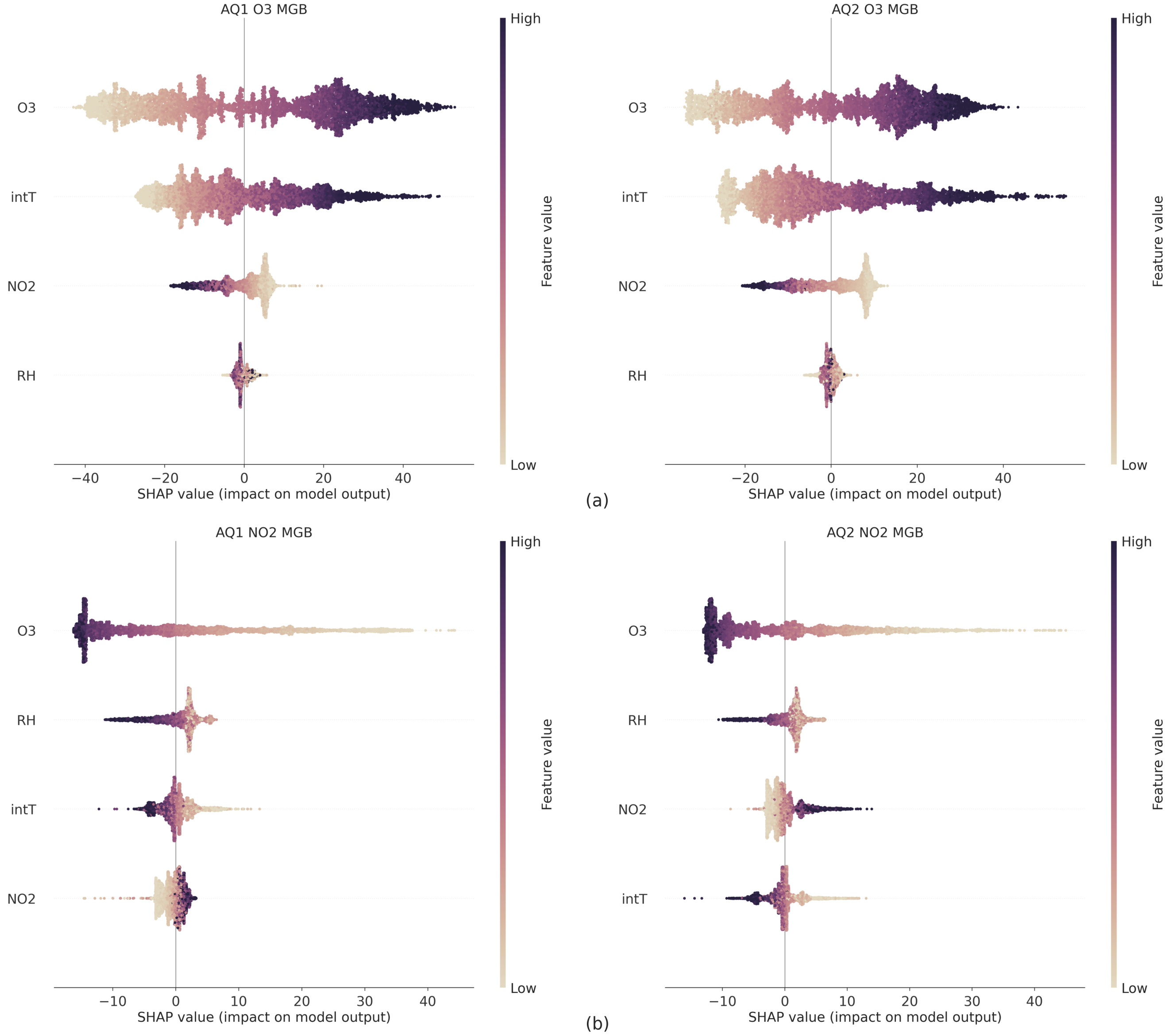

Figure S9. Beeswarm MGB

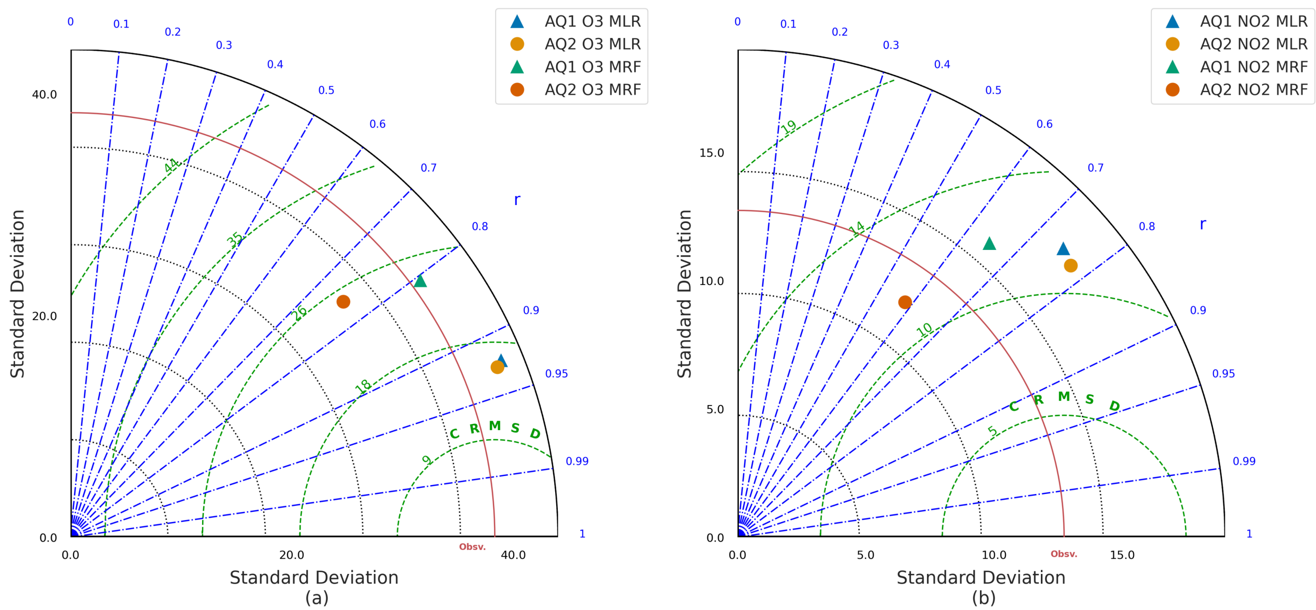

Figure S10. Taylor diagrams of the pre–deployment MLR and MRF calibrated models assessed against the ARPAT reference station for O3 (a) and NO2 (b) concentrations. The Taylor diagram consists of a polar plot in which the radial distance from the origin represents the standard deviation of predictions, and the angle represents the correlation between predictions and observations. Models that agree well with observations lie closer to the red line. CRMSD is a relative measure of model fit, providing a normalized measure of the deviation from the actual values. This deviation is normalized by the standard deviation of the reference values, which allows a fair comparison of models' performance.

Table S11. Seasonal analysis of MFR validation. Min–Max ($\mu g\,m^{-3}$) represents the minimum and maximum concentrations measured by the reference station, while intT is the average internal temperature measured by the AQ stations.

						Stat.			
Year	Season	Pollutant	AQ id	min–max	intT	r	nRMSE	MAE	MBE
2018	Summer	O3	AQ1	6 – 166	34.65	0.90	10.20	12.58	–4.91
			AQ2	6 – 166	34.20	0.84	15.59	20.68	16.52
		NO2	AQ1	1 – 47	34.62	0.64	42.26	16.00	15.87
			AQ2	1 – 47	34.16	0.51	18.93	7.27	5.41
	Autumn	O3	AQ1	2 – 146	28.06	0.84	21.49	24.43	22.53
			AQ2	2 – 146	25.53	0.76	35.59	46.81	46.57
		NO2	AQ1	1 – 62	28.07	0.65	21.83	9.58	6.71
			AQ2	1 – 62	25.54	0.43	18.02	8.85	1.04
	Winter	O3	AQ1	2 – 65	16.60	0.88	52.57	22.13	21.12
			AQ2	2 – 72	14.53	0.83	67.43	39.32	39.17
		NO2	AQ1	3 – 88	16.59	0.71	17.59	11.65	3.42
			AQ2	2 – 88	14.53	0.57	20.64	13.40	–2.82
2019	Spring	O3	AQ1	2 – 132	22.41	0.75	29.11	32.63	31.92
			AQ2	2 – 132	21.14	0.68	35.91	40.38	39.57
		NO2	AQ1	2 – 63	22.32	0.67	12.73	5.67	–0.54
			AQ2	2 – 63	21.09	0.34	17.18	8.29	2.31
	Summer	O3	AQ1	7 – 185	34.98	0.82	13.42	19.44	12.35
			AQ2	7 – 185	32.29	0.75	20.71	31.31	28.19
		NO2	AQ1	0 – 47	34.94	0.50	22.19	8.19	6.81
			AQ2	0 – 47	32.25	0.34	18.00	7.28	5.05