Supplement: Development and Application of a United States wide correction for $PM_{2.5}$ data collected with the PurpleAir sensor

Karoline K. (Johnson) Barkjohn[1], Brett Gantt[2], & Andrea L. Clements[3]

[1]ORISE fellow hosted by the Office of Research and Development, U.S. Environmental Protection Agency 109 T.W. Alexander Drive Research Triangle Park, NC 27711, johnson.karoline@epa.gov

[2]Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, 109 T.W. Alexander Drive Research Triangle Park, NC 27711, gantt.brett@epa.gov

[3]Office of Research and Development, U.S. Environmental Protection Agency 109 T.W. Alexander Drive Research Triangle Park, NC 27711, clements.andrea@epa.gov

Correspondences to: Karoline K. Barkjohn (johnson.karoline@epa.gov)

1 Performance statistics equations

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{d=1}^{N}(x_d - R_d)^2} \quad \text{Eq. S1}$$

$$\text{MBE} = \frac{1}{N}\sum_{d=1}^{N}(x_d) - \frac{1}{N}\sum_{d=1}^{N}(R_d) \quad \text{Eq. S2}$$

$$\text{MAE} = \frac{1}{N}\sum_{d=1}^{N}\text{abs}(x_d - R_d) \quad \text{Eq. S3}$$

where:

RMSE = root mean square error ($\mu g/m^3$)

MBE= mean bias error ($\mu g/m^3$)

MAE= mean absolute error ($\mu g/m^3$)

M = number of identical sensors operated simultaneously during a field test

N = number of 24-hour periods during which all identical instruments are operating and returning valid averages over the duration of the field test

x_d = valid 24-hour averaged sensor $PM_{2.5}$ concentration for day d µ (g/m^3)

R_d = valid 24-hour averaged FRM or FEM $PM_{2.5}$ concentration for day d ($\mu g/m^3$)

2 Discussion of state by state performance

Uncorrected PurpleAir sensors in this work overestimate $PM_{2.5}$ in every state on average except for Florida (Figure S4); this typical overprediction is indicated by MBE greater than 0 μg m^{-3}. Figure S4 shows the state-by-state performance before and after correction. The MBE after correction is within 3 μg m^{-3} in each state with 24-hr RMSE ≤ 3.7 μg m^{-3}, showing large improvements from the initial dataset. The RMSE is reduced and the MBE moves closer to zero across all states except Alaska and Arizona where RMSE improves but MBE moves slightly further from 0 (AK MBE_{raw}=0.8 μg m^{-3}, $MBE_{corrected}$=-1.0 μg m^{-3}; AZ MBE_{raw}=0.4 μg m^{-3}, $MBE_{corrected}$=-0.5 μg m^{-3}) and Florida where correction slightly increases RMSE and increases underprediction (MBE_{raw}=-0.3 μg m^{-3}, $MBE_{corrected}$=-3.7 μg m^{-3}). It is important to note that the reported performance may not accurately summarize state-wide performance in states with less than a year of data (indicated in Figure S4 by states with white background labels) and in states with a single PurpleAir collocated at a single site. To illustrate, Florida is represented by a single PurpleAir sensor located at a single site for less than a year; the performance in Florida may be influenced by a different aerosol composition in Sarasota, Florida (e.g., higher sea salt contribution) or may be an issue unique to the FEM and PurpleAir running at the site.

Table S1. Details for each AQS site as reported in AQS.

AQS ID	Land use type	Location setting	Measurement scale	Network
02-090-0034	Commercial	Urban and center city	Neighborhood	CSN supplemental, NCORE
04-013-0019	Residential	Suburban	Neighborhood	
04-013-1004	Residential	Suburban	Neighborhood	
04-013-9812	Commercial	Urban and center city	Neighborhood	
06-067-0010	Residential	Urban and center city	Neighborhood	CSN supplemental
06-111-0007	Residential	Suburban	Neighborhood	
06-083-0011	Residential	Urban and center city	Neighborhood	
06-071-0306	Residential	Suburban	Neighborhood	
06-083-1008	Commercial	Urban and center city	Neighborhood	
06-111-2002	Residential	Suburban	Neighborhood	
06-083-2004	Commercial	Urban and center city	Neighborhood	
06-079-2007	Industrial	Rural	Neighborhood	
06-083-2011	Residential	Suburban	Urban scale	
06-111-3001	Residential	Rural	Neighborhood	
06-079-8002	Commercial	Suburban	Neighborhood, urban scale	
06-037-9033	Commercial	Urban and center city	Neighborhood	
08-031-0026	Residential	Urban and center city	Neighborhood	CSN STN, NCORE, proposed NCORE
10-003-2004	Commercial	Urban and center city	Neighborhood	CSN supplemental, NCORE
12-115-0013	Residential	Urban and center city	Neighborhood	
13-089-0002	Residential	Suburban	Neighborhood	CSN STN, IMPROVE, NCORE, proposed NCORE
19-163-0015	Residential	Urban and center city	Neighborhood	CSN supplemental, NCORE, proposed NCORE
19-163-0020	Residential	Urban and center city	Neighborhood	
19-153-0030	Commercial	Urban and center city	Neighborhood	
19-113-0040	Residential	Urban and center city	Neighborhood	
19-103-2001	Residential	Urban and center city	Neighborhood	
20-177-0013	Residential	Urban and center city	Neighborhood	
30-063-0024	Residential	Suburban	Neighborhood	
37-067-0022	Residential	Urban and center city	Neighborhood	CSN supplemental

40-109-1037	Residential	Suburban	Urban scale	CSN supplemental
40-115-9007	Residential	Rural	Neighborhood	
50-021-0002	Commercial	Urban and center city	Neighborhood	
51-087-0014	Residential	Suburban	Neighborhood	CSNSTN, NATTS, NCORE
53-033-0057	Industrial	Suburban	Neighborhood	CSN supplemental
53-061-1007	Residential	Suburban	Neighborhood	CSN supplemental
53-061-0005	Commercial	Suburban	Neighborhood	
55-087-0009	Residential	Suburban	Urban scale	
55-079-0026	Commercial	Urban and center city	Neighborhood, urban scale	
55-133-0027	Residential	Urban and center city	Middle scale, neighborhood	
55-025-0041	Commercial	Urban and center city	Urban scale	

Table S2. Details for each individual sensor summarized as median (min, max).

PA ID	AQS site	Start Date	End Date	# of Days	FEM or FRM	FEM or FRM $PM_{2.5}$ ($\mu g\ m^{-3}$)	PA $PM_{2.5}$ ($\mu g\ m^{-3}$)	PA T (^oC)	PA RH (%)
DE1	10-030-2004	7/27/2019	11/18/2019	205	both	7 (1,17)	9 (1,35)	25 (6,35)	51 (34,75)
AK2	02-090-0034	11/7/2018	1/12/2019	54	FRM	12 (2,32)	10 (1,28)	-4 (-25,2)	53 (37,59)
AK4	02-090-0034	1/15/2019	6/16/2019	189	FRM	4 (0,35)	4 (0,37)	6 (-20,24)	46 (22,71)
AK1	02-090-0034	5/15/2019	9/30/2019	126	FRM	4 (0,60)	4 (0,131)	18 (7,29)	46 (21,76)
AZ4	04-013-0019	11/9/2018	12/31/2019	435	both	7 (2,43)	5 (1,74)	26 (9,44)	25 (5,67)
AZ1	04-013-1004	11/30/2018	2/18/2019	76	FEM	9 (2,26)	14 (1,43)	17 (10,22)	36 (22,65)
AZ3	04-013-9812	11/9/2018	12/31/2019	384	FEM	8 (1,28)	5 (0,50)	25 (9,44)	26 (6,73)
CA10	06-037-9033	4/25/2018	7/25/2019	393	FEM	7 (-1,40)	6 (0,73)	25 (6,39)	21 (4,63)
CA1	06-067-0010	5/20/2018	12/29/2019	161	FRM	6 (1,109)	6 (0,250)	22 (9,36)	45 (23,77)
CA13	06-071-0306	3/22/2018	7/24/2019	619	FEM	8 (1,33)	7 (0,73)	28 (6,42)	17 (4,56)
CA15	06-079-2007	6/2/2018	1/4/2019	183	FEM	7 (-2,47)	9 (1,66)	18 (12,25)	66 (35,100)
CA2	06-079-8002	11/29/2017	10/23/2019	659	FEM	5 (-2,34)	5 (0,69)	19 (6,34)	48 (16,79)
CA16	06-079-8002	5/2/2018	8/14/2018	105	FEM	6 (-1,19)	7 (2,17)	19 (16,26)	60 (42,65)
CA5	06-083-0011	1/31/2019	10/30/2019	246	FEM	7 (1,23)	9 (1,25)	22 (11,29)	56 (26,74)
CA6	06-083-1008	1/31/2019	4/29/2019	86	FEM	4 (1,15)	4 (0,21)	18 (7,23)	54 (34,68)
CA4	06-083-2004	11/15/2018	10/29/2019	282	FEM	4 (-2,30)	7 (1,69)	20 (12,29)	55 (22,78)
CA3	06-083-2011	1/24/2019	10/29/2019	265	FEM	5 (0,14)	8 (1,27)	22 (6,29)	55 (24,77)
CA19	06-111-0007	4/30/2019	10/29/2019	152	FEM	9 (2,15)	12 (1,28)	25 (17,32)	45 (3,60)
CA18	06-111-2002	11/9/2018	10/29/2019	398	FEM	9 (0,23)	10 (0,39)	25 (12,35)	37 (2,69)
CA17	06-111-3001	10/26/2018	10/29/2019	213	FEM	6 (0,24)	7 (0,47)	19 (11,30)	54 (8,72)
CO1	08-031-0026	8/22/2019	11/18/2019	113	both	7 (2,25)	6 (1,45)	18 (-5,32)	33 (18,70)
FL1	12-115-0013	5/31/2019	9/30/2019	119	FEM	6 (3,17)	5 (1,25)	32 (29,35)	60 (49,73)
GA1	13-089-0002	8/2/2019	11/18/2019	184	both	9 (3,18)	15 (5,34)	29 (5,36)	55 (44,77)
IA8	19-103-2001	1/1/2019	1/13/2020	1031	both	7 (1,31)	11 (1,54)	11 (-21,34)	56 (28,80)
IA9	19-103-2001	1/1/2019	1/13/2020	1028	both	7 (1,31)	11 (1,58)	11 (-20,35)	57 (29,100)
IA1	19-113-0040	4/27/2018	1/13/2020	1865	both	7 (-1,28)	12 (0,56)	18 (-21,36)	47 (20,68)
IA2	19-113-0040	5/5/2018	1/13/2020	1827	both	7 (-1,28)	11 (0,50)	17 (-23,35)	51 (26,71)
IA3	19-153-0030	9/26/2017	1/13/2020	2473	both	6 (-1,27)	10 (0,63)	11 (-23,36)	55 (25,100)

IA4	19-153-0030	2/2/2019	1/13/2020	999	both	6 (0,27)	10 (1,53)	16 (-13,36)	56 (30,82)
IA10	19-163-0015	1/1/2019	6/14/2019	481	both	8 (1,29)	16 (2,64)	10 (-26,29)	57 (33,79)
IA11	19-163-0015	1/1/2019	1/13/2020	1005	both	8 (1,36)	13 (1,69)	14 (-27,34)	53 (24,77)
IA12	19-163-0020	5/29/2019	12/31/2019	198	FRM	7 (2,31)	11 (1,64)	22 (-7,34)	59 (42,85)
KS1	20-177-0013	3/13/2019	9/30/2019	124	FEM	9 (2,33)	11 (0,49)	25 (10,34)	53 (30,71)
KS2	20-177-0013	3/13/2019	9/30/2019	111	FEM	8 (3,33)	11 (1,49)	25 (11,33)	53 (30,70)
KS3	20-177-0013	3/13/2019	9/28/2019	71	FEM	9 (3,33)	10 (1,50)	22 (9,33)	51 (31,70)
MT1	30-063-0024	12/3/2019	12/10/2019	8	FEM	10 (5,15)	22 (6,36)	4 (2,6)	54 (42,62)
NC4	37-067-0022	3/25/2018	10/24/2019	700	both	7 (0,20)	13 (1,43)	25 (-1,35)	48 (16,79)
OK1	40-109-1037	8/3/2019	11/18/2019	120	both	10 (4,23)	10 (1,32)	30 (1,38)	51 (29,78)
OK3	40-115-9007	7/10/2019	9/30/2019	70	FEM	8 (1,25)	14 (1,35)	30 (24,33)	64 (53,86)
VT1	50-021-0002	3/30/2019	9/30/2019	146	both	6 (2,18)	8 (1,31)	24 (12,34)	52 (36,71)
VA1	51-087-0014	10/27/2019	12/29/2019	30	FRM	5 (2,20)	10 (2,41)	12 (8,25)	48 (35,65)
WA1	53-033-0057	10/16/2017	10/28/2019	561	FEM	7 (1,35)	8 (1,80)	14 (-1,30)	59 (26,81)
WA2	53-061-0005	12/4/2017	8/6/2018	235	FEM	3 (1,26)	6 (0,89)	12 (1,27)	64 (35,84)
WA3	53-061-1007	10/25/2018	9/25/2019	239	FEM	7 (0,41)	9 (1,81)	12 (-2,23)	67 (36,81)
WI4	55-025-0041	1/19/2019	10/24/2019	185	both	6 (2,32)	9 (1,58)	-142 (-142,9)	47 (35,63)
WI5	55-025-0041	5/5/2019	5/7/2019	3	FEM	5 (3,7)	7 (3,9)	16 (14,18)	55 (41,56)
WI1	55-079-0026	7/24/2019	11/18/2019	115	FEM	7 (2,21)	10 (1,50)	22 (-6,33)	57 (37,82)
WI2	55-087-0009	1/6/2019	10/24/2019	450	FEM	6 (1,30)	9 (1,64)	19 (-18,32)	55 (31,78)
WI3	55-087-0009	3/30/2019	4/4/2019	12	FEM	5 (3,11)	5 (2,23)	5 (1,7)	42 (35,63)
WI6	55-133-0027	1/1/2019	3/27/2019	46	both	10 (4,25)	15 (3,55)	0 (-21,11)	54 (34,72)

Table S3. Summary of comparison between A and B channels and AB channel averages and the FEM or FRM data. Table includes the average of the Channel A $PM_{2.5}$ data (A), mean bias error (MBE), normalized mean bias error (NMBE), slope and intercept (B=A*s+i and AB=FM*s+i), Pearson correlation (r), Spearman correlation (ρ), the number of 24-hr averages (N), and the percent of the dataset removed based on the 24-hr A B exclusion criteria (rem). Highlighted rows were removed because the A and B channels were more than 25% different on average after removing individual 24-hr averages that met the exclusion criteria.

	Before 24-hr point removal												After 24-hr point removal											
	A B compare							FEM or FRM compare					A B comparison							FEM or FRM compare					
ID	A	MBE	NMBE	s	i	r	ρ	s	i	r	ρ	N	A	MBE	NMBE	s	i	r	ρ	s	i	r	ρ	N	rem
avg	13	-3	3%	0.97	2	0.84	0.94	1.91	0	0.80	0.85	396	11	0	2%	0.99	0	0.98	0.98	1.80	-3	0.88	0.85	388	3%
med	12	0	4%	1.01	0	1.00	0.99	1.84	-3	0.90	0.88	203	12	0	4%	1.01	0	1.00	1.00	1.88	-3	0.92	0.88	194	0%
min	5	-216	-181%	-0.28	-1	-0.10	-0.06	0.03	-10	0.00	0.38	3	5	-3	-49%	0.58	-1	0.76	0.74	0.24	-10	0.38	0.43	3	0%
max	46	40	150%	2.33	31	1.00	1.00	12.54	61	0.99	1.00	2476	23	3	45%	1.26	3	1.00	1.00	3.49	8	0.99	1.00	2473	47%
AK1	10	0	-1%	0.97	0	1.00	0.99	2.24	-4	1.00	0.93	126	10	0	-1%	0.97	0	1.00	0.99	2.24	-4	1.00	0.93	126	0%
AK2	14	3	26%	1.04	3	0.51	0.69	0.62	4	0.68	0.82	71	10	-1	-7%	0.82	1	0.86	0.89	0.59	3	0.85	0.88	54	24%
AK3	8	-7	-57%	0.57	0	0.99	0.99	1.22	-4	0.95	0.93	227	5	-3	-49%	0.66	0	0.99	0.99	1.00	-4	0.95	0.92	205	10%
AK4	7	-2	-21%	0.75	1	0.99	1.00	1.29	-1	0.97	0.94	189	7	-2	-21%	0.75	1	0.99	1.00	1.29	-1	0.97	0.94	189	0%
AZ1	16	0	3%	1.01	0	1.00	0.99	1.77	-2	0.96	0.97	76	16	0	3%	1.01	0	1.00	0.99	1.77	-2	0.96	0.97	76	0%
AZ3	10	1	6%	1.02	0	1.00	1.00	1.68	-6	0.90	0.88	385	10	1	6%	1.02	0	1.00	1.00	1.68	-6	0.90	0.88	384	0%
AZ4	11	1	7%	1.06	0	1.00	1.00	1.78	-6	0.97	0.93	435	11	1	7%	1.06	0	1.00	1.00	1.78	-6	0.97	0.93	435	0%
CA1	13	-1	-10%	0.94	-1	1.00	1.00	2.36	-8	0.98	0.84	161	13	-1	-10%	0.94	-1	1.00	1.00	2.36	-8	0.98	0.84	161	0%
CA10	8	1	10%	1.08	0	1.00	1.00	1.40	-2	0.80	0.73	393	8	1	10%	1.08	0	1.00	1.00	1.40	-2	0.80	0.73	393	0%
CA13	8	1	9%	1.09	0	1.00	1.00	1.51	-5	0.80	0.77	619	8	1	9%	1.09	0	1.00	1.00	1.51	-5	0.80	0.77	619	0%
CA16	12	-1	-11%	0.93	0	1.00	0.99	1.00	4	0.69	0.59	183	12	-1	-11%	0.93	0	1.00	0.99	1.00	4	0.69	0.59	183	0%
CA18	10	0	1%	1.04	0	1.00	1.00	1.86	-3	0.91	0.86	213	10	0	1%	1.04	0	1.00	1.00	1.86	-3	0.91	0.86	213	0%
CA19	12	1	6%	1.05	0	1.00	1.00	1.74	-4	0.88	0.89	398	12	1	6%	1.05	0	1.00	1.00	1.74	-4	0.88	0.89	398	0%
CA2	8	-1	-9%	0.93	0	0.99	0.98	1.88	-2	0.86	0.75	661	8	-1	-9%	0.93	0	0.99	0.99	1.88	-2	0.86	0.76	659	0%
CA20	12	0	4%	1.07	0	1.00	1.00	1.57	-2	0.75	0.71	152	12	0	4%	1.07	0	1.00	1.00	1.57	-2	0.75	0.71	152	0%
CA3	10	2	24%	1.32	0	0.40	0.96	1.63	1	0.49	0.78	269	9	1	14%	1.07	1	0.98	0.97	1.40	1	0.76	0.78	265	1%
CA4	46	40	150%	2.33	31	0.04	0.75	0.78	23	0.02	0.64	339	9	2	23%	1.01	2	0.97	0.91	1.63	0	0.85	0.75	282	17%
CA5	9	0	4%	1.02	0	0.98	0.98	0.93	3	0.65	0.75	246	9	0	4%	1.02	0	0.98	0.98	0.93	3	0.65	0.75	246	0%
CA6	5	-1	-25%	0.93	-1	0.94	0.87	0.83	1	0.69	0.77	88	5	-1	-23%	0.95	-1	0.95	0.90	0.85	1	0.70	0.80	86	2%
CA7	8	-5	-49%	0.29	4	0.45	0.45	0.22	9	0.33	0.38	201	9	-3	-27%	0.58	2	0.76	0.75	0.24	8	0.38	0.43	129	36%

CO1	8	1	8%	1.12	0	1.00	1.00	1.97	-6	0.96	0.87	113	8	1	8%	1.12	0	1.00	1.00	1.97	-6	0.96	0.87	113	0%
DE1	11	-1	-7%	0.95	0	1.00	1.00	2.24	-5	0.92	0.92	205	11	-1	-7%	0.95	0	1.00	1.00	2.24	-5	0.92	0.92	205	0%
FL1	7	0	-4%	0.98	0	1.00	1.00	0.99	0	0.69	0.81	119	7	0	-4%	0.98	0	1.00	1.00	0.99	0	0.69	0.81	119	0%
GA1	16	1	8%	0.78	5	0.90	0.85	1.98	-2	0.91	0.88	200	16	1	4%	0.85	3	0.94	0.93	2.10	-3	0.92	0.90	184	8%
IA1	14	1	6%	1.04	0	0.99	0.99	1.87	-2	0.89	0.85	1869	14	1	6%	1.05	0	0.99	0.99	1.87	-2	0.89	0.85	1865	0%
IA10	41	26	92%	1.67	15	0.12	0.97	1.71	12	0.11	0.86	493	18	3	17%	1.24	-1	0.99	0.99	2.25	-4	0.91	0.90	481	2%
IA11	17	3	18%	1.25	-1	0.99	0.99	2.23	-4	0.93	0.9	1005	17	3	18%	1.25	-1	0.99	0.99	2.23	-4	0.93	0.90	1005	0%
IA12	14	0	-1%	1.01	0	1.00	1.00	1.98	-3	0.93	0.91	198	14	0	-1%	1.01	0	1.00	1.00	1.98	-3	0.93	0.91	198	0%
IA2	13	-1	-5%	0.89	1	0.98	0.98	1.81	-1	0.89	0.85	1827	13	-1	-5%	0.89	1	0.98	0.98	1.81	-1	0.89	0.85	1827	0%
IA3	12	1	6%	1.09	0	0.98	0.99	2.14	-3	0.90	0.88	2476	12	1	6%	1.08	0	0.99	0.99	2.14	-3	0.90	0.88	2473	0%
IA4	12	1	13%	0.96	2	0.90	0.96	1.90	-2	0.90	0.88	1013	12	2	16%	1.10	1	0.96	0.97	1.88	-2	0.91	0.88	999	1%
IA8	14	1	4%	1.00	1	1.00	1.00	2.01	-4	0.93	0.9	1031	14	1	4%	1.00	1	1.00	1.00	2.01	-4	0.93	0.90	1031	0%
IA9	15	1	10%	1.12	0	0.99	0.99	1.97	-3	0.93	0.9	1028	15	1	10%	1.12	0	0.99	0.99	1.97	-3	0.93	0.90	1028	0%
KS1	12	0	1%	1.02	0	1.00	1.00	1.72	-3	0.95	0.94	124	12	0	1%	1.02	0	1.00	1.00	1.72	-3	0.95	0.94	124	0%
KS2	13	2	14%	0.99	2	0.57	0.97	1.67	-2	0.77	0.91	113	12	1	5%	1.04	0	1.00	1.00	1.74	-4	0.95	0.94	111	2%
KS3	12	0	-3%	0.98	0	1.00	1.00	1.79	-4	0.95	0.92	71	12	0	-3%	0.98	0	1.00	1.00	1.79	-4	0.95	0.92	71	0%
MT1	23	0	0%	0.99	0	1.00	1.00	3.22	-10	0.98	0.98	8	23	0	0%	0.99	0	1.00	1.00	3.22	-10	0.98	0.98	8	0%
NC4	15	1	6%	1.04	0	0.82	0.95	2.10	-3	0.85	0.92	715	14	1	4%	1.00	1	0.96	0.97	2.06	-2	0.90	0.92	700	2%
OK1	11	2	22%	1.19	1	0.90	0.87	1.36	-3	0.86	0.82	127	11	2	18%	1.26	-1	0.97	0.96	1.38	-3	0.86	0.81	120	6%
OK3	15	1	5%	1.06	0	1.00	1.00	1.19	5	0.77	0.8	70	15	1	5%	1.06	0	1.00	1.00	1.19	5	0.77	0.80	70	0%
VA1	12	0	-3%	0.97	0	1.00	1.00	2.32	-3	0.97	0.86	30	12	0	-3%	0.97	0	1.00	1.00	2.32	-3	0.97	0.86	30	0%
VT1	10	0	5%	1.03	0	1.00	1.00	2.15	-5	0.96	0.95	146	10	0	5%	1.03	0	1.00	1.00	2.15	-5	0.96	0.95	146	0%
WA1	11	0	4%	1.07	0	1.00	1.00	1.90	-4	0.92	0.85	561	11	0	4%	1.07	0	1.00	1.00	1.90	-4	0.92	0.85	561	0%
WA2	12	-216	-181%	0.00	12	0.02	0.97	12.54	61	0.09	0.8	246	12	0	-2%	0.99	0	1.00	0.99	3.49	-4	0.84	0.81	235	4%
WA3	15	-39	-112%	0.00	15	-0.03	0.94	0.03	35	0.00	0.87	243	16	-1	-5%	0.95	0	1.00	0.99	2.25	-3	0.97	0.93	239	2%
WA5	9	8	143%	-0.28	10	-0.10	-0.06	0.58	1	0.76	0.63	338	5	2	45%	0.65	3	0.89	0.74	0.71	1	0.46	0.43	179	47%
WI1	11	0	3%	1.00	0	1.00	1.00	2.11	-5	0.93	0.94	115	11	0	3%	1.00	0	1.00	1.00	2.11	-5	0.93	0.94	115	0%
WI2	13	2	19%	1.03	2	0.54	0.98	2.03	-3	0.74	0.93	454	12	1	11%	1.09	0	1.00	1.00	2.09	-4	0.95	0.95	450	1%
WI3	9	0	-5%	0.94	0	1.00	1.00	2.64	-8	0.99	0.99	12	9	0	-5%	0.94	0	1.00	1.00	2.64	-8	0.99	0.99	12	0%
WI4	11	0	-1%	0.95	0	1.00	1.00	2.24	-4	0.92	0.88	185	11	0	-1%	0.95	0	1.00	1.00	2.24	-4	0.92	0.88	185	0%
WI5	6	0	3%	1.02	0	1.00	1.00	2.04	-4	0.99	1	3	6	0	3%	1.02	0	1.00	1.00	2.04	-4	0.99	1.00	3	0%
WI6	17	0	-2%	0.96	0	0.99	0.99	2.37	-8	0.92	0.88	46	17	0	-2%	0.96	0	0.99	0.99	2.37	-8	0.92	0.88	46	0%

Table S4. Correction equation forms considered and the adjusted R^2. The best performing model from each increasing complexity (as indicated with *) was validated using withholding in the next sections (Sections 4.3.2 and 4.4).

Name	Eqn	R^2_{adj}
Cf_1	$PA[cf\_1]=PM_{2.5}*s_1+b$	0.780*
Cf_atm	$PA[cf\_atm]=PM_{2.5}*s_1+b$	0.765
Bins	$PM_{2.5} = (s_1*B>0.3) + (s_2*B>0.5) + (s_3*B>1.0) + (s_4*B>2.5) + (s_5*B>5.0) + (s_6*B>10.0) + i$	0.769
RH	$PA[cf\_1] = s_1*PM_{2.5} + s_2*RH + i$	0.831*
+T	$PA[cf\_1] = s_1*PM_{2.5} + s_2*T + i$	0.792
+D	$PA[cf\_1] = s_1*PM_{2.5} + s_2*D + i$	0.788
Nonlinear RH	$PA = s_1*PM_{2.5} + s_2\frac{RH^2}{(1-RH)}*PM_{2.5} + s_3*\frac{RH^2}{(1-RH)} + i$	0.782
+RH+T	$PA = s_1*PM_{2.5} + s_2*RH + s_3*T + i$	0.832*
+RH+D	$PA = s_1*PM_{2.5} + s_2*RH + s_3*D + i$	0.832
+D+T	$PA = s_1*PM_{2.5} + s_2*T + s_3*D + i$	0.827
+RH+T+D	$PA = s_1*PM_{2.5} + s_2*RH + s_3*T + s_4*D + i$	0.832
PM*RH	$PA = s_1*PM_{2.5} + s_2*RH + s_3*RH*PM_{2.5} + i$	0.836*
PM*RH*T	$PA = s_1*PM_{2.5} + s_2*RH + s_3*T + s_4*PM_{2.5}*RH + s_5*PM_{2.5}*T + s_6*RH*T + s_7*PM_{2.5}*RH*T + i$	0.838*

Table S5. Summary of the performance parameters for the full dataset correction, and corrections built and tested using leave one state out (LOSO) and leaving out 12 random weeks, leave out by date (LOBD). Spearman correlation, root mean squared error (RMSE), and mean bias error (MBE) are summarized as median (min, max).

Correction Model	Dataset	R^2 (Train dataset)	Spearman correlation (Test dataset)	RMSE (μg m^{-3}) (Test dataset)	MBE (μg m^{-3}) (Test dataset)	MAE (μg m^{-3}) (Test dataset)
0: Raw (cf_atm)	All		0.86	7.4	4.2	5.2
1: Linear	All	0.78	0.86	2.8	0	2.0
2: +RH	All	0.83	0.90	2.5	0	1.7
3: +RH+T	All	0.83	0.90	2.5	0	1.7
4: PM*RH	All	0.84	0.90	2.5	0	1.7
5: PM*RH*T	All	0.84	0.90	2.4	0	1.7
0: Raw	LOSO		0.91 (0.71,0.98)	5.9 (3.2,12.9)	3.3 (-0.3,11.3)	4.2 (2.1,11.3)
1: Linear	LOSO	0.78 (0.73,0.81)	0.91 (0.71,0.98)	2.6 (1,4)	0 (-2.4,3.2)	1.8 (0.9,3.3)
2: +RH	LOSO	0.83 (0.79,0.85)	0.91 (0.78,0.98)	2 (1,3.8)	-0.2 (-3,2.8)	1.6 (0.9,3)
3: +RH+T	LOSO	0.83 (0.78,0.85)	0.92 (0.78,0.98)	2 (1,3.8)	-0.3 (-2.9,2.7)	1.6 (0.9,2.9)
4: PM*RH	LOSO	0.84 (0.79,0.86)	0.92 (0.79,0.98)	2 (1,3.7)	-0.3 (-2.8,2.6)	1.6 (0.8,2.8)
5: PM*RH*T	LOSO	0.84 (0.79,0.86)	0.92 (0.78,0.98)	2 (1,3.9)	-0.3 (-2.8,2.5)	1.6 (0.8,2.9)
0: Raw	LOBD		0.85 (0.58,0.94)	7.4 (4.2,13)	4.2 (1.1,9.2)	4.2 (2.1,11.3)
1: Linear	LOBD	0.78 (0.75,0.8)	0.85 (0.58,0.94)	2.8 (1.9,6.3)	0 (-1.4,2.3)	1.8 (0.9,3.3)
2: +RH	LOBD	0.83 (0.81,0.85)	0.89 (0.75,0.95)	2.5 (1.8,5.6)	0 (-1,1.3)	1.6 (0.9,3)
3: +RH+T	LOBD	0.83 (0.81,0.85)	0.89 (0.75,0.95)	2.5 (1.7,5.6)	0 (-0.9,1.3)	1.6 (0.9,2.9)
4: PM*RH	LOBD	0.84 (0.82,0.85)	0.90 (0.76,0.95)	2.5 (1.7,5.1)	0 (-0.9,1.4)	1.6 (0.8,2.8)
5: PM*RH*T	LOBD	0.84 (0.82,0.85)	0.90 (0.76,0.95)	2.5 (1.7,4.9)	0 (-0.9,1.3)	1.6 (0.8,2.9)

Table S6. Significant difference between performance statistics on the test datasets (as summarized in Table A5) summarized as significantly "improve", "worsen", or no change "—". A Shapiro test was first used to determine whether the metrics were likely normally or non-normally distributed. Most of the LOBD metrics were non-normally distributed while most of the LOSO statistics were normally distributed. Wilcoxon signed rank tests were used for the non-normally distributed comparisons while t-test (paired, 2 tailed) were run for the normally distributed metrics. We compare the raw to linear corrected metrics (0Raw: 1lin) and see significant improvement across most, next we compare the linear to additive RH model (1Lin: 2+RH) and see significant improvement across most metrics. Next, we compare the additive RH model to the additive RH and T model (2+RH:3+RH+T), it does not significantly change the LOSO metrics and improves some and worsens some of the LOBD metrics and so we determine that it is not a better model. Then we compare the additive RH model to the multiplicative RH model (2+RH:4:PM*RH) finding no improvement in most of the LOSO metrics suggesting it is not a significantly better model. Lastly, we do the same with the additive RH model and PM*RH*T model which shows no LOSO metric improvement. In the end the additive RH model appears to provide the most significant improvement.

		0Raw: 1lin	1Lin: 2+RH	2+RH:3+RH+T	2+RH:4:PM*RH	2+RH:5PM*RH*T
LOBD	MAE	improve	improve	worsen	improve	improve
	MBE	improve	improve	improve	--	--
	RMSE	improve	improve	worsen	improve	improve
	spearman	--	improve	improve	improve	improve
LOSO	MAE	improve	--	--	--	--
	MBE	improve	--	--	--	--
	RMSE	improve	improve	--	--	--
	spearman	--	--	--	improve	--

Table S7. Summary of correction parameters built and tested using leave one site out (LOSO) and leaving out 12 random weeks, leave out by date (LOBD). The results of the multiple model builds are reported as median (min, max) where i is the intercept and s_1-s_7 are the coefficients for the increasingly complex models as defined in the 5 proposed correction equations (Section **Error! Reference source not found.**).

Model	type	i	s_1	s_2	s_3	s_4	s_5	s_6	s_7
1: Linear	LOSO	-3 (-3,-3)	1.91 (1.73,1.99)						
2: +RH	LOSO	-11 (-11,-10)	1.91 (1.78,1.97)	0.16 (0.14,0.16)					
3: +RH+T	LOSO	-10 (-12,-9)	1.91 (1.77,1.98)	0.16 (0.14,0.16)	-0.02 (-0.05,0.01)				
4: PM*RH	LOSO	-7 (-7,-6)	1.43 (1.26,1.53)	0.09 (0.05,0.1)	0.01 (0.01,0.01)				
5: PM*RH*T	LOSO	-2 (-4,0)	0.86 (0.51,1.27)	0 (-0.04,0.03)	-0.25 (-0.33,-0.14)	0.019 (0.014, 0.024)	0.028 (0.011, 0.041)	0.0042 (0.0029, 0.0061)	-5e-04 (-8e-04, -3e-04)
1: Linear	LOBD	-3 (-3,-3)	1.92 (1.85,1.96)						
2: +RH	LOBD	-11 (-11,-10)	1.91 (1.85,1.95)	0.16 (0.15,0.17)					
3: +RH+T	LOBD	-10 (-11,-10)	1.91 (1.85,1.94)	0.16 (0.15,0.16)	-0.02 (-0.04,0.00)				
4: PM*RH	LOBD	-7 (-8,-5)	1.45 (1.05,1.56)	0.09 (0.04,0.1)	0.01 (0.01,0.02)				
5: PM*RH*T	LOBD	-2 (-6,2)	0.87 (0.36,1.36)	0 (-0.05,0.06)	-0.25 (-0.35,-0.02)	0.019 (0.012, 0.026)	0.028 (-0.001, 0.043)	0.0042 (4e-04, 0.0057)	-5e-04 (-7e-04, 0)

Table S8. Performance by quality assurance methods and corrections. Quality assurance (QA) criteria include excluding 24-hr averages where <90% of measurements are available (completeness), comparison of the A and B channels where data is excluded when the A and B channels are different by both 5 µg m^{-3} and 60% (AB), and the removal of 3 sensors that had poor agreement in the A and B channel after excluding 24-hr problematic points (problem sensors, details in section 4.1). Performance is compared for the individual channels (i.e. A, B) and also as the average of the A and B channels (AB-shaded rows). Statistics include the percent mean bias error (PMBE=MBE/average FM) linear regression slope (s) and intercept (i), and Lin's concordance correlation coefficient (CCC)

QA criteria	Correction	Channels	RMSE (µg m^{-3})	MAE (µg m^{-3})	MBE (µg m^{-3})	PMBE (%)	s	i	R^2	Pearson r	Spearman ρ	CCC	N	Mean FM (µg m^{-3})	Mean PA (µg m^{-3})
None	None	AB cf_atm	155	17	16	194%	0.99	16	0	0.04	0.8	0	56541	8	24
	None	AB cf_1	99	11	10	121%	1.24	8	0.01	0.08	0.83	0.01	56541	8	18
	US	A	87	7	5	60%	0.87	6	0	0.06	0.85	0.01	55299	8	13
	US	B	161	12	8	102%	0.28	14	0	0.01	0.7	0	55299	8	16
	US	AB	92	9	7	81%	0.57	10	0	0.04	0.8	0.01	55299	8	15
	LRAPA	AB	49	5	0	2%	0.62	3	0.01	0.08	0.83	0.02	56541	8	8
	AQ&U	AB	77	9	8	104%	0.97	9	0.01	0.08	0.83	0.01	56541	8	17
Completeness	None	AB cf_atm	64	7	6	80%	1.49	2	0.02	0.13	0.82	0.02	20721	8	14
	None	AB cf_1	43	6	5	65%	1.27	3	0.03	0.16	0.82	0.04	20721	8	13
	US	A	30	3	1	11%	0.9	2	0.03	0.16	0.86	0.06	20554	8	9
	US	B	69	4	2	20%	0.86	3	0	0.07	0.84	0.01	20554	8	9
	US	AB	38	3	1	16%	0.88	2	0.02	0.13	0.86	0.04	20554	8	9
AB	None	AB cf_atm	9	6	5	64%	1.8	-1	0.71	0.84	0.84	0.56	52108	8	13
	None	AB cf_1	9	6	5	64%	1.57	1	0.66	0.81	0.85	0.56	52108	8	13
	US	A	4	2	1	8%	1.09	0	0.71	0.84	0.87	0.81	50866	8	9
	US	B	3	2	0	-2%	1.03	0	0.74	0.86	0.86	0.84	50866	8	8
	US	AB	4	2	0	3%	1.06	0	0.73	0.86	0.87	0.84	50866	8	8
AB, completeness	None	AB cf_atm	9	6	5	67%	1.56	1	0.66	0.81	0.86	0.56	46598	8	13
	None	AB cf_1	9	6	5	67%	1.78	-1	0.71	0.84	0.85	0.56	46598	8	13
	US	A	3	2	1	6%	0.97	1	0.72	0.85	0.88	0.84	45374	8	9
	US	B	3	2	0	-3%	0.91	1	0.75	0.86	0.88	0.86	45374	8	8
	US	AB	3	2	0	2%	0.94	1	0.74	0.86	0.88	0.86	45374	8	8
AB, completeness, problem sensors	None	AB cf_atm	8	5	4	53%	1.69	-1	0.73	0.86	0.83	0.61	20293	8	12
	None	AB cf_1	7	5	4	48%	1.45	0	0.7	0.84	0.83	0.64	20293	8	12
	US	A	3	2	0	-1%	0.89	1	0.74	0.86	0.86	0.86	20126	8	8
	US	B	3	2	0	-5%	0.89	0	0.78	0.88	0.87	0.88	20126	8	7
	US	AB	3	2	0	-3%	0.89	1	0.77	0.88	0.87	0.88	20126	8	8
	LRAPA	AB	4	3	-3	-34%	0.72	-1	0.7	0.84	0.83	0.72	20293	8	5
	AQ&U	AB	6	4	4	49%	1.13	3	0.7	0.84	0.83	0.67	20293	8	12

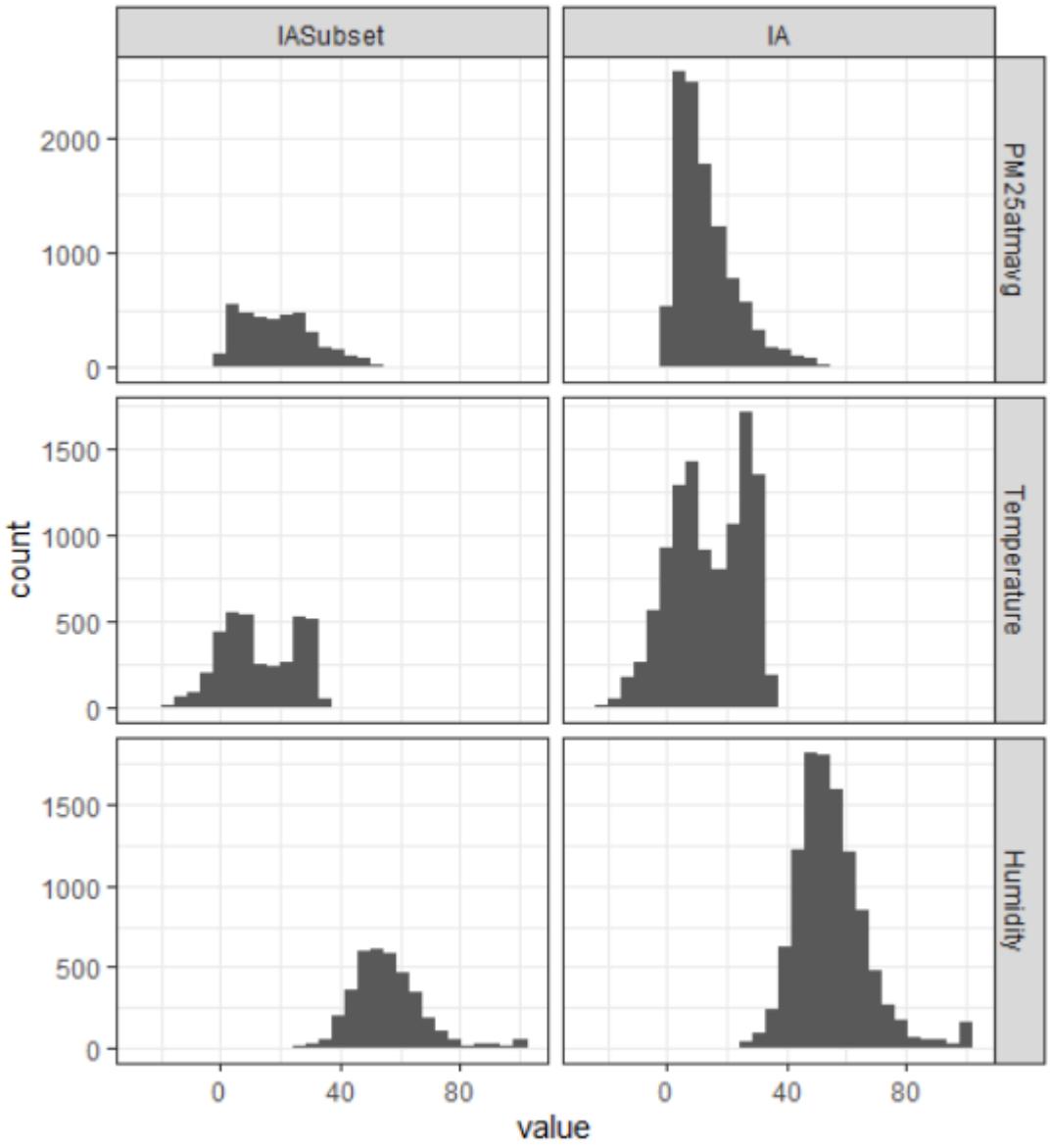

Figure S1. Distribution of subset of Iowa data used in the analysis and the full set of data from Iowa.

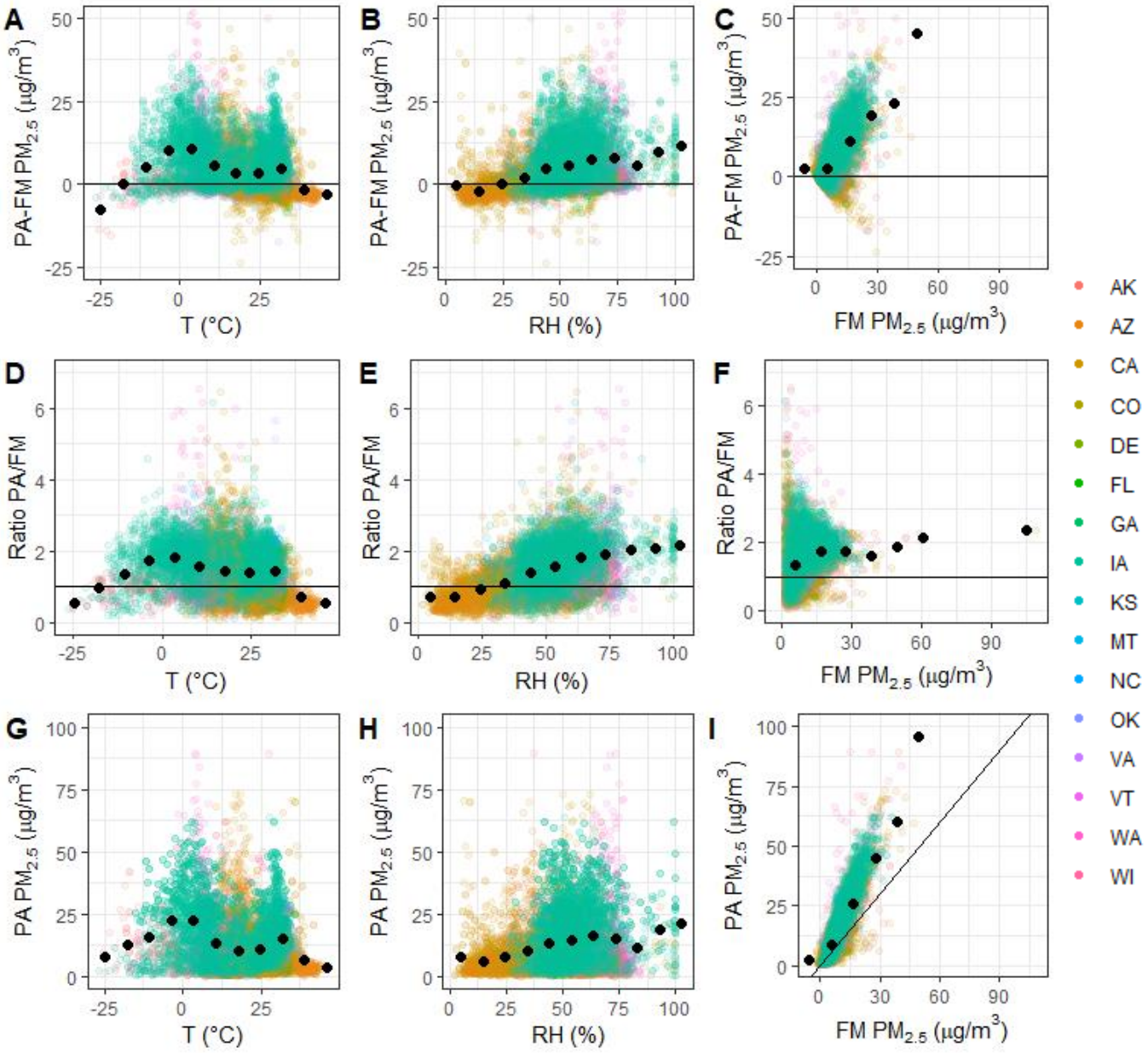

Figure S2. Error and ratio between uncorrected cf_1 PurpleAir (PA) and federal equivalent or reference method measurements (FM) along with raw PurpleAir $PM_{2.5}$ as influenced by temperature, relative humidity, and FEM or FRM $PM_{2.5}$ concentration. Black points indicate averages in 10 bins. Note that there are limited points in the high concentration bins.

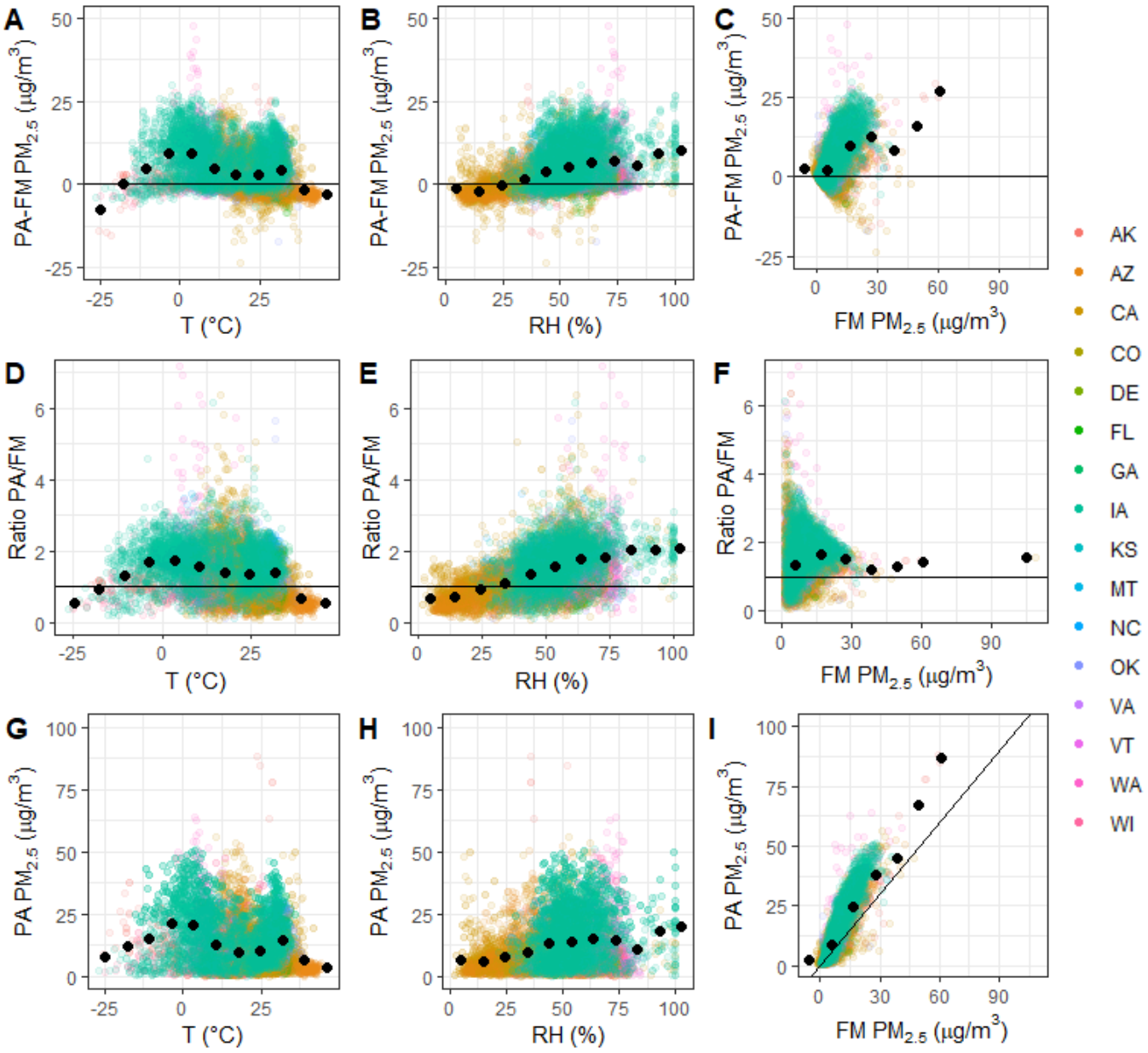

Figure S3. Error and ratio between raw cf_atm PurpleAir (PA) and federal equivalent or reference method measurements (FM) along with raw PurpleAir $PM_{2.5}$ as influenced by temperature, relative humidity, and FEM or FRM $PM_{2.5}$ concentration. Black points indicate averages in 10 bins. Note that there are limited points in high concentration bins leading to greater uncertainty.

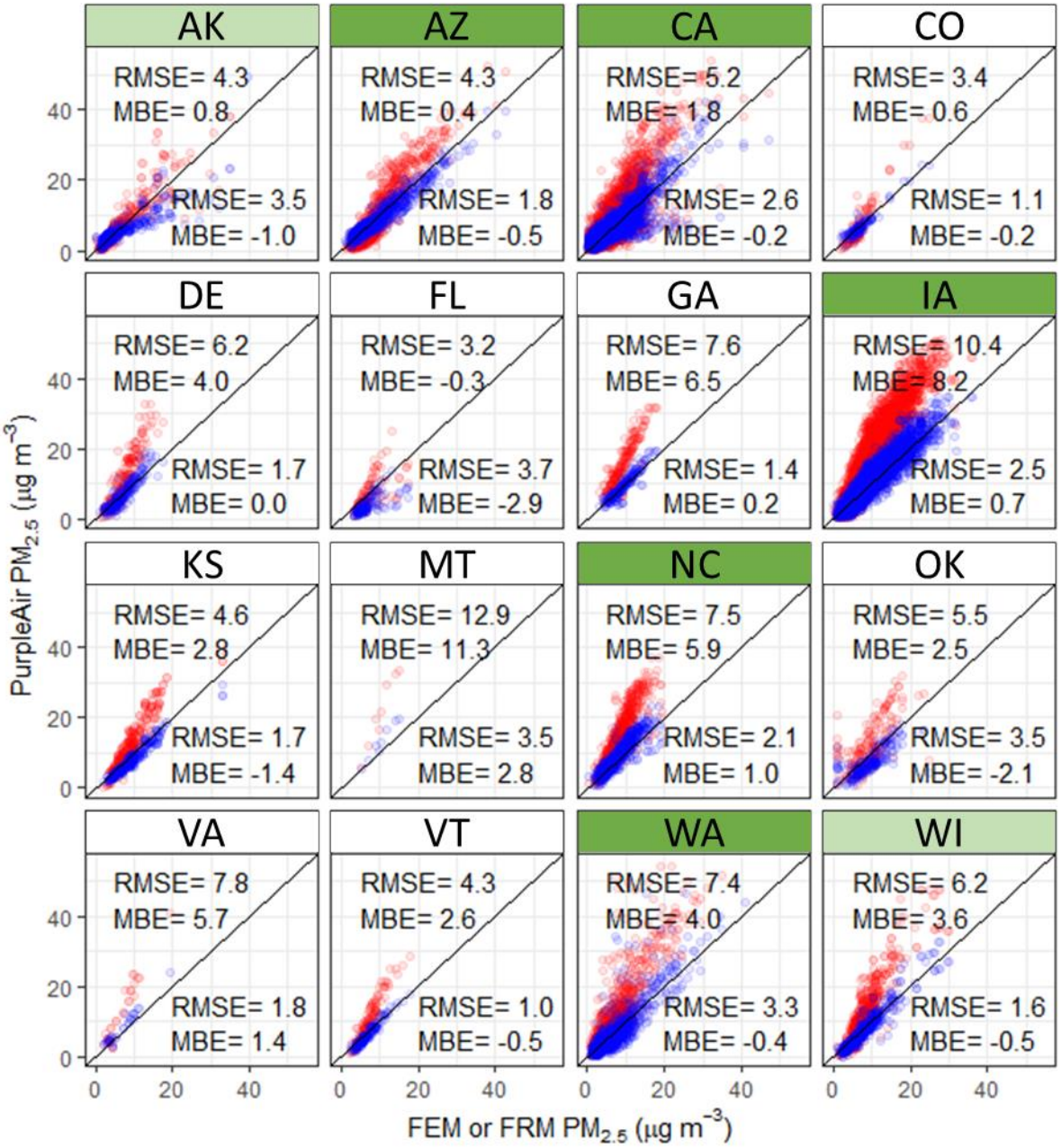

Figure S4. Raw [cf_atm] PurpleAir $PM_{2.5}$ as shown on PurpleAir.com (red) and final corrected data (blue) with the RMSE and MBE in µg m^{-3} reported in the top left corner for the raw data and the lower right for the corrected data. States with at least 1 year of data have green plot labels with light green labels indicating almost 1 year (>10 months).

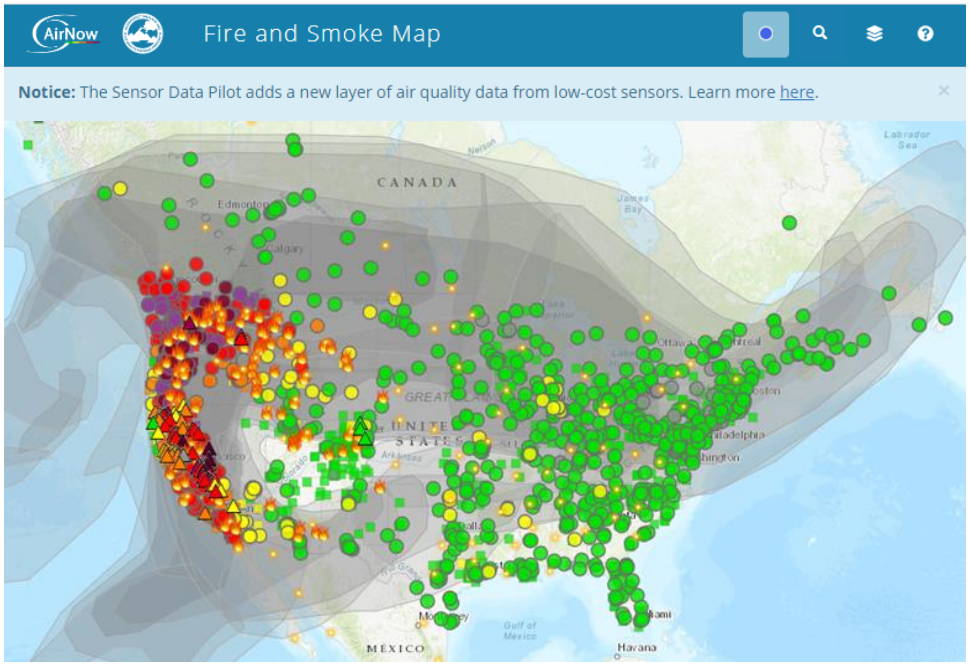

Figure S5. AirNow Fire and Smoke Map (fire.airnow.gov)