A Random Forest Approach to Estimate Daily Particulate Matter, Nitrogen Dioxide and Ozone at Fine Spatial Resolution in Sweden

Table S1. Results of the Stage 1 model relating $PM_{2.5}$ and $PM_{2.5-10}$ to co-located PM_{10}: descriptive statistics and cross-validation fit.

Year	with PM_{10}	OOB CV		10-fold CV		Predictions ($\mu g/m^3$)	
		R^2	RMSE ($\mu g/m^3$)	R^2	RMSE ($\mu g/m^3$)	median	25–75 pct.
$PM_{2.5}$							
2005	0.70	0.74	3.4	0.66	3.9	9.8	7.0–14.5
2006	0.70	0.73	4.1	0.60	5.0	10.1	7.0–14.5
2007	0.71	0.68	3.3	0.59	3.7	8.0	5.8–11.0
2008	0.65	0.62	3.1	0.43	3.8	8.0	5.7–11.1
2009	0.60	0.63	3.1	0.31	4.1	6.9	4.7–10.3
2010	0.60	0.74	3.1	0.61	3.7	6.0	4.3–9.1
2011	0.67	0.83	3.1	0.76	3.7	6.8	4.5–10.4
2012	0.52	0.68	2.8	0.38	4.0	6.1	4.0–9.1
2013	0.55	0.67	2.5	0.50	3.0	5.6	3.7–7.8
2014	0.70	0.79	2.7	0.71	3.2	6.6	4.3–10.0
2015	0.67	0.79	2.2	0.69	2.6	5.5	3.6–8.0
2016	0.61	0.67	2.3	0.47	3.0	5.1	3.5–7.4
$PM_{2.5-10}$							
2005	0.89	0.89	3.4	0.85	4.0	5.5	3.0–10.0
2006	0.84	0.85	3.9	0.78	4.7	5.6	3.3–10.8
2007	0.91	0.88	3.4	0.85	3.8	7.0	4.0–12.5
2008	0.92	0.90	3.0	0.83	3.8	6.7	3.7–12.1
2009	0.91	0.90	2.9	0.81	4.0	7.1	4.0–11.8
2010	0.86	0.89	3.0	0.83	3.8	6.0	3.6–10.5
2011	0.82	0.90	3.1	0.84	3.8	7.2	4.3–12.0
2012	0.89	0.91	2.7	0.82	3.8	6.1	3.3–10.2
2013	0.93	0.94	2.5	0.90	3.1	7.5	4.4–12.7
2014	0.82	0.87	2.7	0.79	3.3	6.8	3.8–10.4
2015	0.87	0.90	2.2	0.84	2.8	5.9	3.6–9.9
2016	0.91	0.90	2.3	0.83	3.1	5.7	3.3–9.8

Table S2. Results of the Stage 2 model relating MAIAC AOD to co-located CAMS AOD: descriptive statistics and statistics of model fit in OOB predictions.

Year	MAIAC AOD Observations			with	OOB CV		Predictions	
	% Missing	Median	25–75 pct.	CAMS AOD *	R²	RMSE	Median	25–75 pct.
2005	91	0.10	0.05–0.17	0.70	0.83	0.04	0.26	0.17–0.34
2006	91	0.12	0.06–0.20	0.78	0.88	0.04	0.30	0.21–0.37
2007	91	0.08	0.04–0.15	0.70	0.84	0.04	0.28	0.20–0.35
2008	91	0.10	0.06–0.17	0.68	0.83	0.04	0.27	0.19–0.35
2009	91	0.11	0.06–0.18	0.65	0.83	0.04	0.30	0.21–0.38
2010	92	0.07	0.03–0.14	0.73	0.85	0.04	0.30	0.23–0.35
2011	92	0.11	0.06–0.20	0.70	0.86	0.04	0.29	0.19–0.36
2012	93	0.09	0.04–0.16	0.63	0.83	0.04	0.28	0.20–0.35
2013	91	0.09	0.04–0.17	0.73	0.87	0.04	0.28	0.16–0.36
2014	96	0.10	0.06–0.18	0.64	0.88	0.03	0.27	0.19–0.35
2015	91	0.07	0.03–0.14	0.69	0.85	0.04	0.26	0.16–0.34
2016	91	0.08	0.04–0.16	0.64	0.82	0.04	0.27	0.19–0.35

* CAMS AOD at band = 0.469 nm, hour = 12.00 am

Table S3. Results of the Stage 3 model: statistics of model fit in predictions from "Out-of-bag" (OOB) samples.

	PM_{10}		$PM_{2.5}$		$PM_{2.5-10}$		NO_2		O_3	
	R²	RMSE (µg/m³)	R²	RMSE (µg/m³)	R²	RMSE (µg/m³)	R²	RMSE (µg/m³)	R²	RMSE (µg/m³)
Overall	0.64	7.2	0.69	3.1	0.65	4.8	0.74	8.3	0.77	8.7
Spatial *	0.67	2.5	0.86	1.0	0.80	1.3	0.86	3.6	0.79	3.0
Temporal **	0.61	6.9	0.65	3.0	0.59	4.8	0.56	7.6	0.76	8.5
By Year										
2005	0.62	8.2	0.64	4.1	0.63	4.6	0.77	7.9	0.74	9.7
2006	0.62	8.1	0.62	4.5	0.64	4.8	0.75	8.4	0.72	10.8
2007	0.66	7.7	0.67	3.3	0.66	4.9	0.77	7.8	0.68	9.3
2008	0.62	7.5	0.60	2.8	0.63	5.4	0.70	8.4	0.79	8.8
2009	0.60	7.6	0.59	3.2	0.59	5.5	0.74	7.8	0.76	8.6
2010	0.66	6.7	0.67	2.9	0.69	4.6	0.70	9.9	0.77	8.6
2011	0.65	7.3	0.74	3.5	0.66	4.8	0.74	8.5	0.80	8.5
2012	0.64	6.7	0.61	3.0	0.67	4.8	0.73	8.3	0.80	8.2
2013	0.65	7.1	0.62	2.2	0.67	5.6	0.73	8.6	0.82	7.9
2014	0.68	6.4	0.73	3.1	0.69	3.8	0.75	7.8	0.83	7.6
2015	0.62	6.3	0.69	2.6	0.64	3.9	0.73	8.0	0.80	7.3
2016	0.59	6.5	0.64	2.2	0.63	4.7	0.75	7.8	0.77	8.2
By Season										
Jan–Mar	0.63	8.6	0.68	3.8	0.65	5.6	0.74	9.1	0.74	9.0
Apr–Jun	0.62	8.2	0.63	3.2	0.65	5.8	0.74	8.0	0.69	8.6
Jul–Sep	0.66	4.3	0.72	2.2	0.62	2.8	0.74	7.2	0.68	8.7
Oct–Dec	0.61	6.3	0.72	2.9	0.56	4.3	0.72	8.6	0.70	8.4
By Monitor Location										
Urban Traffic	0.64	7.6	0.69	3.1	0.65	5.4	0.62	10.1	0.79	8.1
Urban Background	0.60	6.5	0.71	3.1	0.52	4.1	0.64	6.4	0.76	8.8
Rural	0.64	4.3	0.66	2.9	0.62	2.5	0.72	2.9	0.75	8.8
In Major Cities										
Stockholm	0.81	5.8	0.80	2.4	0.77	5.1	0.53	11.9	0.82	8.1
Göteborg	0.68	5.7	0.70	2.8	0.62	4.6	0.77	8.5	0.80	8.8
Malmö	0.80	4.3	0.84	3.1	0.52	3.1	0.68	7.5	0.89	5.8
Uppsala	0.47	8.8	0.69	2.8	0.44	6.8	0.34	12.2	-	-
Umeå	0.59	7.1	0.60	2.5	0.61	5.8	0.75	9.8	-	-

* Computed comparing long-term average observations versus predictions.

** Computed comparing deviations between daily and long-term average observations versus predictions.

Table S4. Results of the Stage 3 model: statistics of model fit in predictions from 10-fold cross-validation by monitors.

	PM_{10}		$PM_{2.5}$		$PM_{2.5-10}$		NO_2		O_3	
	R^2	RMSE (µg/m³)	R^2	RMSE (µg/m³)	R^2	RMSE (µg/m³)	R^2	RMSE (µg/m³)	R^2	RMSE (µg/m³)
Overall	0.50	8.5	0.59	3.6	0.45	6.1	0.37	12.9	0.60	11.6
*Spatial**	0.29	3.6	0.67	1.5	0.29	2.5	0.31	7.9	0.05	6.4
*Temporal***	0.53	7.5	0.59	3.3	0.46	5.5	0.37	9.1	0.67	9.8
By Year										
2005	0.46	9.8	0.54	4.7	0.39	5.9	0.46	12.0	0.62	11.8
2006	0.46	9.7	0.50	5.1	0.42	6.2	0.43	12.8	0.54	13.8
2007	0.48	9.5	0.56	3.8	0.41	6.5	0.44	12.1	0.56	10.9
2008	0.44	9.0	0.43	3.3	0.43	6.6	0.37	12.2	0.64	11.4
2009	0.43	9.0	0.39	3.9	0.38	6.8	0.38	12.1	0.61	11.0
2010	0.52	7.9	0.57	3.3	0.50	5.8	0.34	14.7	0.52	12.3
2011	0.55	8.4	0.67	3.9	0.48	6.0	0.32	13.7	0.62	11.8
2012	0.50	7.8	0.46	3.5	0.50	5.9	0.34	12.9	0.65	10.8
2013	0.53	8.2	0.49	2.5	0.52	6.8	0.35	13.3	0.60	11.6
2014	0.54	7.6	0.67	3.4	0.42	5.2	0.34	12.6	0.65	10.8
2015	0.47	7.5	0.59	3.0	0.40	5.0	0.37	12.3	0.58	10.5
2016	0.45	7.6	0.53	2.5	0.42	5.8	0.36	12.5	0.59	10.9
By Season										
Jan-Mar	0.52	9.8	0.60	4.3	0.48	6.9	0.36	14.1	0.56	11.7
Apr-Jun	0.45	9.9	0.50	3.7	0.43	7.4	0.36	12.6	0.33	12.5
Jul-Sep	0.43	5.6	0.57	2.7	0.30	3.8	0.38	11.3	0.44	11.5
Oct-Dec	0.44	7.6	0.63	3.3	0.33	5.3	0.36	13.0	0.52	10.6
By Monitor Location										
Urban Traffic	0.54	8.6	0.59	3.5	0.50	6.5	0.37	13.0	0.59	11.2
Urban Background	0.50	7.2	0.65	3.5	0.38	4.6	0.29	9.0	0.61	11.1
Rural	0.34	5.9	0.54	3.4	0.20	3.6	0.62	3.4	0.60	11.0
In Major Cities										
Stockholm	0.47	9.5	0.70	3.0	0.56	7.1	0.14	16.1	0.30	15.8
Göteborg	0.50	7.1	0.57	3.3	0.35	6.0	0.38	14.1	0.70	10.8
Malmö	0.68	5.5	0.69	4.3	0.26	3.8	0.35	10.7	0.85	7.0
Uppsala	0.46	8.9	0.69	2.8	0.41	7.0	0.03	14.8	-	-
Umeå	0.41	8.5	0.37	3.1	0.44	7.0	0.27	16.8	-	-

Table S5. Distribution of PM_{10} and NO_2 concentrations (µg/m³) from Random Forest (RF) and Dispersion Model (DM), Stockholm, 2015.

Model	Mean	SD	Percentiles					
			5th	25th	50th	75th	95th	
PM_{10}								
RF	12.0	1.5	10.3	10.9	11.6	13.2	21.6	0.22
DM	11.8	2.7	9.7	10.1	11.0	12.5	16.2	
NO_2								
RF	4.6	4.6	2.2	2.4	2.6	4.1	14.6	0.75
DM	7.7	4.8	2.8	4.3	6.4	9.5	16.9	