

Supplementary Materials

**Daily PM_{2.5} and seasonal-trend decomposition to identify extreme
air pollution events from 2001 to 2020 for continental Australia
using a Random Forest Model**

Table S1. Summary of variables and datasets used for modelling and predicting

Spatial / Spatiotemporal / Temporal	Variable name	Broad type	Narrow type	Variable	Unit(s)	Data source	Time basis	Buffers (m)
Spatiotemporal	pm25	Response	Monitor PM _{2.5}	Fine particulate matter (PM _{2.5})	(µg/m ³)	National Air Pollution Monitor Database (NAPMD) (Centre for Safe Air, 2021; Riley et al., 2020)	daily	N/A
Spatiotemporal	aod_merra_2_[time_basis]	Satellite-based air pollution estimates	Satellite AOD	Total Aerosol Scattering AOT [550 nm]	(dimensionless)	NASA MERRA-2 (Buchard et al., 2017)	daily/ monthly/ yearly	N/A
Spatiotemporal	bc_merra_2_[time_basis]	Satellite-based air pollution estimates	Satellite Black Carbon	Black Carbon Surface Mass Concentration	(kg/m ³)	NASA MERRA-2 (Buchard et al., 2017)	daily/ monthly/ yearly	N/A
Spatiotemporal	dust_merra_2_[time_basis]	Satellite-based air pollution estimates	Satellite PM _{2.5} - Dust	Dust Surface Mass Concentration - PM _{2.5}	(kg/m ³)	NASA MERRA-2 (Buchard et al., 2017)	daily/ monthly/ yearly	N/A
Spatiotemporal	oc_merra_2_[time_basis]	Satellite-based air pollution estimates	Satellite Organic Carbon	Organic Carbon Surface Mass Concentration - PM _{2.5}	(kg/m ³)	NASA MERRA-2 (Buchard et al., 2017)	daily/ monthly/ yearly	N/A
Spatiotemporal	so4_merra_2_[time_basis]	Satellite-based air pollution estimates	Satellite SO ₄	SO ₄ Surface Mass Concentration	(kg/m ³)	NASA MERRA-2 (Buchard et al., 2017)	daily/ monthly/ yearly	N/A
Spatiotemporal	ss_merra_2_[time_basis]	Satellite-based air pollution estimates	Satellite PM _{2.5} - Sea Salt	Sea Salt Surface Mass Concentration - PM _{2.5}	(kg/m ³)	NASA MERRA-2 (Buchard et al., 2017)	daily/ monthly/ yearly	N/A
Spatiotemporal	pm25_merra_2_[time_basis]	Satellite-based air pollution estimates	Satellite-derived PM _{2.5}	Calculated PM _{2.5}	(µg/m ³)	NASA MERRA-2 (Buchard et al., 2016, 2017)	daily/ monthly/ yearly	N/A
Spatial	pos_elevation	Position and Elevation	Elevation above sea level	Elevation	(m)	Satellite-derived (SRTM) Geoscience Australia 1-second smoothed digital elevation model (Gallant et al., 2011)	N/A	N/A
Spatial	pos_distocean	Position and Elevation	Distance to ocean	Distance to ocean	(decimal degrees)	Geoscience Australia Euclidean Distance Coastline 2011	N/A	N/A

Spatial / Spatiotemporal / Temporal	Variable name	Broad type	Narrow type	Variable	Unit(s)	Data source	Time basis	Buffers (m)
						(Geoscience Australia, 2011)		
Spatial	lon, lat	Position and Elevation	Longitude, latitude	Longitude and latitude	(decimal degrees)	Location data	N/A	N/A
Spatial	lc_tree_cover_[buffer]	Land cover	Tree cover	Tree cover percentage of buffer	(%)	MODIS satellite - land cover (Dimiceli et al., 2015)	N/A	50, 100, 200, 300, 400, 500, 700, 1000, 1500, 2000, 3000, 5000, 10000
Spatial	NDVI_[buffer]	Land cover	NDVI	Normalised Difference Vegetation Index	(dimensionless)	MODIS Terra satellite (Didan, 2015)	Monthly	0, 1000, 1500, 2000, 3000, 5000, 10000
Spatial	lc_impasa_[buffer]	Land cover	Impervious surfaces	Impervious surface percentage of buffer	(%)	(Brown de Colstoun et al., 2017)	N/A	50, 100, 200, 300, 400, 500, 700, 1000, 1500, 2000, 3000, 5000, 10000
Spatial	lc_water_modis_[buffer]	Land cover	Water bodies	Ratio of water bodies coverage in buffer	(dimensionless)	MODIS satellite - water cover (Dimiceli et al., 2015)	N/A	50, 100, 200, 300, 400, 500, 700, 1000, 1500, 2000, 3000, 5000, 10000
Spatial	lc_park_[buffer]	Land cover	Parkland	Ratio of parkland in buffer	(dimensionless)	ABS meshblocks (Hanigan, I. & Learnihan, 2017)	N/A	50, 100, 200, 300, 400, 500, 700, 1000, 1500, 2000, 3000, 5000, 10000
Spatial	lu_res_[buffer]	Land use	Residential	Ratio of residential area in buffer	(dimensionless)	ABS meshblocks (Hanigan, I. & Learnihan, 2017)	N/A	50, 100, 200, 300, 400, 500, 700, 1000, 1500, 2000, 3000, 5000, 10000
Spatial	lu_com_[buffer]	Land use	Commercial	Ratio of commercial area in buffer	(dimensionless)	ABS meshblocks (Hanigan, I. & Learnihan, 2017)	N/A	50, 100, 200, 300, 400, 500, 700, 1000, 1500, 2000, 3000, 5000, 10000
Spatial	lu_ind_[buffer]	Land use	Industrial	Ratio of industrial area in buffer	(dimensionless)	ABS meshblocks (Hanigan, I. & Learnihan, 2017)	N/A	50, 100, 200, 300, 400, 500, 700, 1000, 1500, 2000,

Spatial / Spatiotemporal / Temporal	Variable name	Broad type	Narrow type	Variable	Unit(s)	Data source	Time basis	Buffers (m)
								3000, 5000, 10000
Spatial	lu_open_[buffer]	Land use	Open	Ratio of open area in buffer	(dimensionless)	ABS meshblocks (Hanigan, I. & Learnihan, 2017)	N/A	50, 100, 200, 300, 400, 500, 700, 1000, 1500, 2000, 3000, 5000, 10000
Spatiotemporal	em_burned_area_pct_[buffer]	Emission sources	Landscape fires	Total burned vegetation area ratio	(dimensionless)	(Giglio et al., 2018)	monthly	10000, 25000, 50000, 100000, 250000, 500000
Spatiotemporal	em_fire_dens_fireskm2_[buffer][time_basis]	Emission sources	Landscape fires	Active fire density	(#fires/km ²)	(NASA_FIRMS, 2021)	daily/ monthly/ yearly	10000, 25000, 50000, 100000, 250000, 500000
Spatial	em_npi_pm25_dens_kgkm2_[buffer]	Emission sources	Point source emission sites	PM _{2.5} emission mass density in buffer	(kg/km ²)	Australia National Pollutant Inventory Department of Agriculture, Water and the Environment (2022)	N/A	300, 400, 500, 700, 1000, 1500, 2000, 3000, 5000, 10000
Spatial	em_npi_pm25_dens_siteskm2_[buffer]	Emission sources	Point source emission sites	PM _{2.5} emission sites density in buffer	(sites/km ²)	Australia National Pollutant Inventory Department of Agriculture, Water and the Environment (2022)	N/A	300, 400, 500, 700, 1000, 1500, 2000, 3000, 5000, 10000
Spatial	em_minrd_length_[buffer]	Emission sources	Roads	Minor road length in buffer	(m)	PSMA roads (Yuen, 2020)	N/A	50, 100, 200, 300, 400, 500, 700, 1000, 1500, 2000, 3000, 5000, 10000
Spatial	em_majrd_length_[buffer]	Emission sources	Roads	Major road length in buffer	(m)	PSMA roads (Yuen, 2020)	N/A	50, 100, 200, 300, 400, 500, 700, 1000, 1500, 2000, 3000, 5000, 10000
Spatial	em_minrd_invdist	Emission sources	Roads	Inverse distance to nearest minor road	(1/m)	PSMA roads (Yuen, 2020)	N/A	N/A
Spatial	em_majrd_invdist	Emission sources	Roads	Inverse distance to nearest major road	(1/m)	PSMA roads (Yuen, 2020)	N/A	N/A

Spatial / Spatiotemporal / Temporal	Variable name	Broad type	Narrow type	Variable	Unit(s)	Data source	Time basis	Buffers (m)
Spatial	em_rds_int_imputed	Emission sources	Traffic	Traffic count (vehicles per day) on nearest major road, imputed	(vpd)	PSMA roads and Zenith traffic counts (Yuen, 2020)	N/A	N/A
Spatial	em_rds_intinvdist_imputed	Emission sources	Traffic	Product of traffic count on nearest major road and inverse distance to that road, imputed	(vpd/m)	PSMA roads and Zenith traffic counts (Yuen, 2020)	N/A	N/A
Spatial	em_all_rds_int_imputed	Emission sources	Traffic	Traffic count on nearest road, imputed	(vpd)	PSMA roads and Zenith traffic counts (Yuen, 2020)	N/A	N/A
Spatial	em_all_rds_intinvdist_imputed	Emission sources	Traffic	Product of traffic count on nearest road and inverse distance to that road, imputed	(vpd/m)	PSMA roads and Zenith traffic counts (Yuen, 2020)	N/A	N/A
Spatial	em_wood_heat_pct	Emission sources	Wood heaters	Percentage of households using wood as main energy source for heating	(%)	ABS household energy source survey (Australian Bureau of Statistics, 2011)	N/A	N/A
Spatiotemporal	pop_dens_[buffer]	Population	Population density	Population density	(persons /km ²)	ABS population density grids (Australian Bureau of Statistics, 2021)	yearly	50, 100, 200, 300, 400, 500, 700, 1000, 1500, 2000, 3000, 5000, 10000
Spatiotemporal	wthr_maxtmp_[time_basis]_avg	Weather	Meteorological	Maximum temperature, aggregated to monthly and yearly means	(°C)	Australian Water Availability Project (AWAP) Weather grids (Bureau of Meteorology, 2020)	daily/ monthly/ yearly	N/A
Spatiotemporal	wthr_mintmp_[time_basis]_avg	Weather	Meteorological	Minimum temperature, aggregated to	(°C)	Australian Water Availability Project (AWAP) Weather grids	daily/ monthly/ yearly	N/A

Spatial / Spatiotemporal / Temporal	Variable name	Broad type	Narrow type	Variable	Unit(s)	Data source	Time basis	Buffers (m)
				monthly and yearly means		(Bureau of Meteorology, 2020)		
Spatiotemporal	wthr_meantmp_[time_basis]_avg	Weather	Meteorological	Mean temperature, aggregated to monthly and yearly means	(°C)	Australian Water Availability Project (AWAP) Weather grids (Bureau of Meteorology, 2020)	daily/ monthly/ yearly	N/A
Spatiotemporal	wthr_rain_[time_basis]	Weather	Meteorological	Total rainfall, aggregated to monthly and yearly sums	(mm)	Australian Water Availability Project (AWAP) Weather grids (Bureau of Meteorology, 2020)	daily/ monthly/ yearly	N/A
Spatiotemporal	wthr_solarexp_dly	Weather	Meteorological	Solar exposure	(MJ/m ²)	Australian Water Availability Project (AWAP) Weather grids (Bureau of Meteorology, 2020)	daily	N/A
Spatiotemporal	wthr_relhum_dly	Weather	Meteorological	Relative humidity	(%)	Australian Water Availability Project (AWAP) Weather grids (Bureau of Meteorology, 2020)	daily	N/A
Spatiotemporal	wthr_wind_[height]_[time_basis]	Weather	Meteorological	Daily mean wind speed at height = 10m, 100m	(m/s)	ECMWF Reanalysis v5 (ERA5) (Muñoz Sabater, 2019)	daily	N/A
Spatiotemporal	pblh	Weather	Meteorological	Planetary boundary layer height	(m)	ECMWF Reanalysis v5 (ERA5) (Muñoz Sabater, 2019)	daily	N/A
Spatiotemporal	clim_wind_[height]_[time_basis]	Weather	Meteorological	Long-term mean wind speed for calendar month/year at height = 10m, 100m	(m/s)	ECMWF Reanalysis v5 (ERA5) (Muñoz Sabater, 2019)	monthly/ yearly	N/A
Spatiotemporal	wthr_drt_spei	Weather	Drought	Calculated SPEI Index	(dimensionless)	AWAP Weather grids (Beguería et al., 2014; Bureau of Meteorology, 2020;	monthly	N/A

Spatial / Spatiotemporal / Temporal	Variable name	Broad type	Narrow type	Variable	Unit(s)	Data source	Time basis	Buffers (m)
						Vicente-Serrano et al., 2010)		
Spatiotemporal	wthr_drt_duration_spei	Weather	Drought	Cumulative count of uninterrupted drought months (SPEI < 0)	(dimensionless)	Australian Water Availability Project (AWAP) Weather grids (Beguería & Vicente-Serrano, 2023; Bureau of Meteorology, 2020)	monthly	N/A
Spatiotemporal	wthr_drt_count_months	Weather	Drought	Number of drought months (SPEI < 0) in a calendar year	(# of months)	Australian Water Availability Project (AWAP) Weather grids (Beguería & Vicente-Serrano, 2023; Bureau of Meteorology, 2020)	yearly	N/A
Spatiotemporal	wthr_drt_max_length	Weather	Drought	Maximum length (in months) of uninterrupted drought months (SPEI < 0) within a calendar year	(# of months)	Australian Water Availability Project (AWAP) Weather grids (Beguería & Vicente-Serrano, 2023; Bureau of Meteorology, 2020)	yearly	N/A
Temporal	date	Temporal	date	Date (YYYY-MM-DD)	N/A	N/A	N/A	N/A
Temporal	Year	Temporal	Year	Year	N/A	N/A	N/A	N/A
Temporal	Month	Temporal	Month	Month number of year	N/A	N/A	N/A	N/A
Temporal	Day	Temporal	Day	Day number of month	N/A	N/A	N/A	N/A
Temporal	Week	Temporal	Week	Week number of year (number of complete 7-day cycles plus one since start of year)	N/A	N/A	N/A	N/A
Temporal	wday	Temporal	wday	Day name	N/A	N/A	N/A	N/A
Temporal	yday	Temporal	yday	Day number of year (Julian day)	N/A	N/A	N/A	N/A
Temporal	dow	Temporal	dow	Day number of week (starting 1 = Sunday)	N/A	N/A	N/A	N/A

Table S2. Groups of variables considered for excluding highly correlated variables

Variable group	Variables considered
wthr_tmp_dly	wthr_mintmp_dly_avg, wthr_meantmp_dly_avg, wthr_maxtmp_dly_avg
wthr_tmp_mnthly	wthr_mintmp_mnthly_avg, wthr_meantmp_mnthly_avg, wthr_maxtmp_mnthly_avg
wthr_tmp_yrly	wthr_mintmp_yrly_avg, wthr_meantmp_yrly_avg, wthr_maxtmp_yrly_avg
pop_dens	pop_dens_{buffer} (buffer = 50..10,000)
wthr_wind_daily	wthr_wind_10m_daily, wthr_wind_100m_daily
clim_wind_month	clim_wind_10m_month, clim_wind_100m_month
clim_wind_annual	clim_wind_10m_annual, clim_wind_100m_annual
lu_res	lu_res_{buffer} (buffer = 50..10,000)
lu_com	lu_com_{buffer} (buffer = 50..10,000)
lu_open	lu_open_{buffer} (buffer = 50..10,000)
lu_ind	lu_ind_{buffer} (buffer = 50..10,000)
lc_park	lc_park_{buffer} (buffer = 50..10,000)
lcimpsa	lcimpsa_{buffer} (buffer = 50..10,000)
lc_tree_cover	lc_tree_cover_{buffer} (buffer = 50..10,000)
lc_water_modis	lc_water_modis_{buffer} (buffer = 50..10,000)
em_fire_dens_fireskm2_daily	em_fire_dens_fireskm2_{buffer}_daily (buffer = 10,000..500,000)
em_fire_dens_fireskm2_monthly	em_fire_dens_fireskm2_{buffer}_monthly (buffer = 10,000..500,000)
em_fire_dens_fireskm2_yearly	em_fire_dens_fireskm2_{buffer}_yearly (buffer = 10,000..500,000)
em_npi_pm25_dens_kgkm2	em_npi_pm25_dens_kgkm2_{buffer} (buffer = 300..10,000)
em_npi_pm25_dens_siteskm2	em_npi_pm25_dens_siteskm2_{buffer} (buffer = 300..10,000)
em_burned_area_pct	em_burned_area_pct_{buffer} (buffer = 10,000..500,000)
em_rds	em_rds_dist, em_majrd_invdist, em_rds_int_imputed, em_rds_intinvdist_imputed, em_minrd_invdist, em_all_rds_int_imputed, em_all_rds_intinvdist_imputed
em_majrd_length	em_majrd_length_{buffer} (buffer = 50..10,000)
em_minrd_length	em_minrd_length_{buffer} (buffer = 50..10,000)
NDVI	NDVI_{buffer} (buffer = 0..10,000)
merra_2_dly	bc_merra_2_dly, dust_merra_2_dly, oc_merra_2_dly, so4_merra_2_dly, ss_merra_2_dly, aod_merra_2_dly, pm25_merra_2_dly
merra_2_mnthly	bc_merra_2_mnthly, dust_merra_2_mnthly, oc_merra_2_mnthly, so4_merra_2_mnthly, ss_merra_2_mnthly, aod_merra_2_mnthly, pm25_merra_2_mnthly
merra_2_yrly	bc_merra_2_yrly, dust_merra_2_yrly, oc_merra_2_yrly, so4_merra_2_yrly, ss_merra_2_yrly, aod_merra_2_yrly, pm25_merra_2_yrly

Table S3. PM_{2.5} imputation model summary

Variable	Default model	Tuned model
num.trees	500	600
mtry	3	4
num_variables	9	9
OOB_RMSE	4.01	3.90
OOB_R_squared	81.1%	82.15
test_RMSE	3.36	3.34
test_R_squared	80%	80%
train_RMSE	2.31	2
train_R_squared	94.74%	95.87%

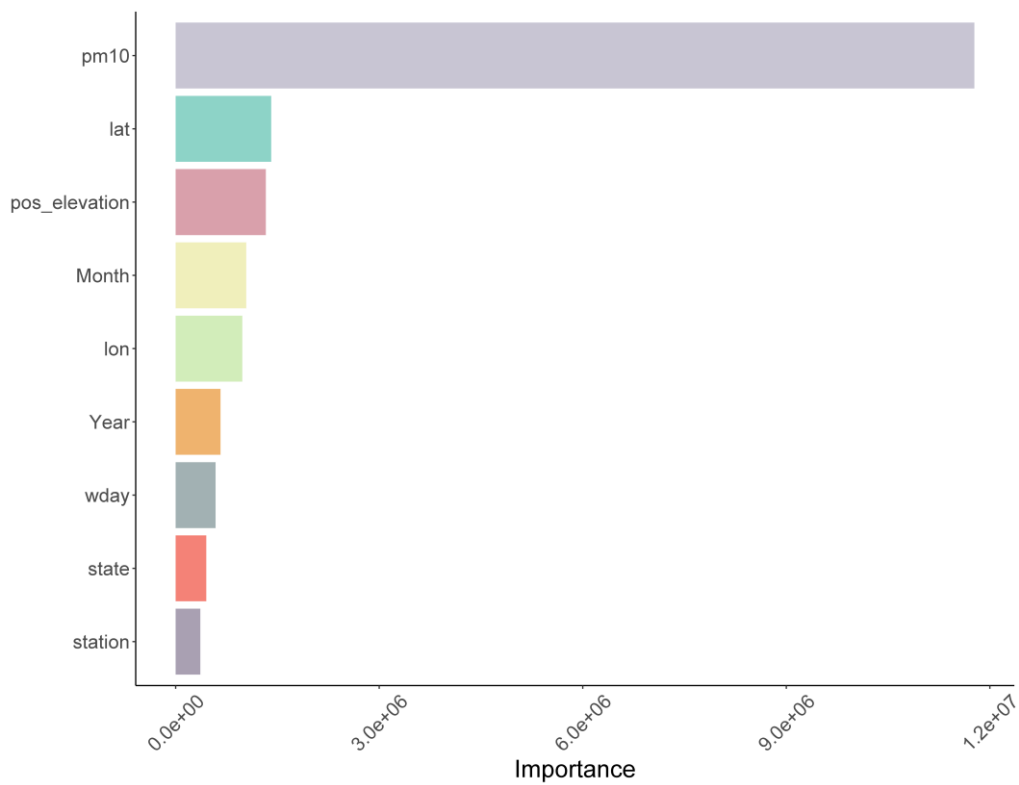


Figure S1. Variable importance of PM_{2.5} imputation model

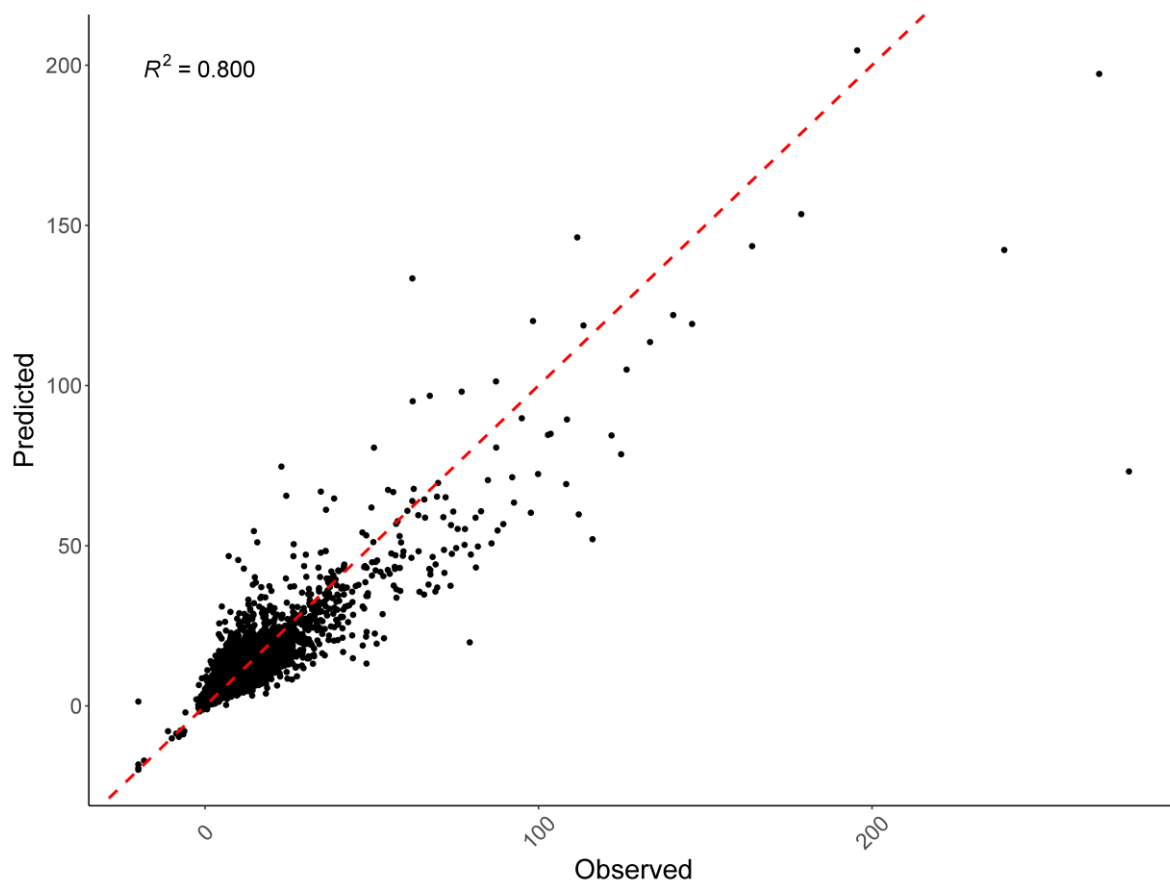


Figure S2. R-squared on testing set (10% left out) of PM_{2.5} imputation model

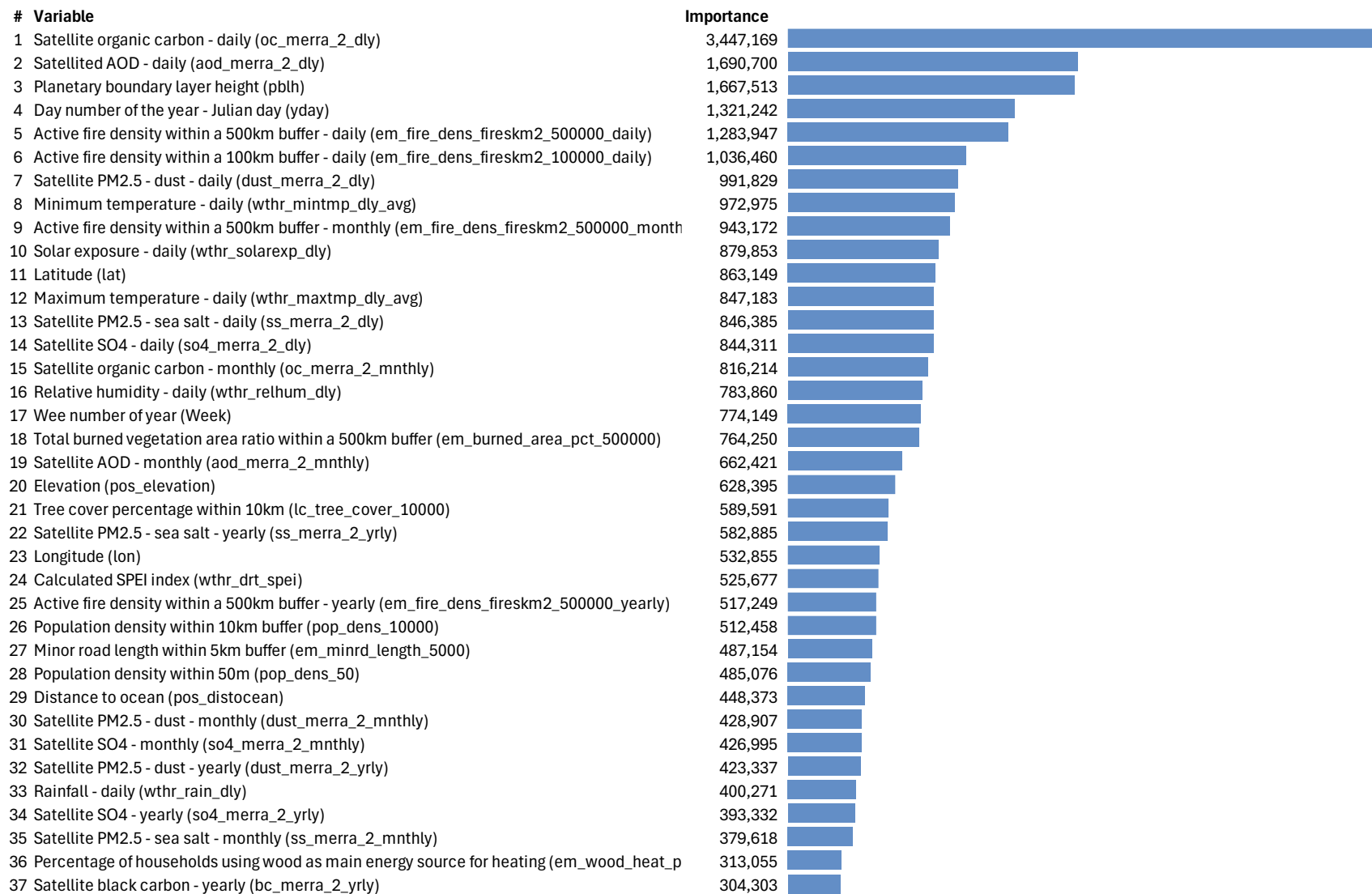


Figure S3. Variable importance of final model in Stage 2: PM_{2.5} modelling (37 variables)

Table S4. Daily PM_{2.5} model performance summaries for steps 1-4

Model	Step 1 (base model)	Step 2 (exclusion of highly correlated variables)	Step 3 (hyperpara meter tuning)	Step 4 (final model)
num.trees	500	500	400	500
mtry	16	10	30	6
num_variables	259	101	101	37
OOB_RMSE	4.82	4.64	4.43	4.45
OOB_R_squared	66.7%	69.1%	71.8%	71.5%
test_RMSE	4.40	4.18	4.10	4.08
test_R_squared	64.3%	68.0%	69.1%	69.5%
train_RMSE	2.42	2.28	2.00	2.12
train_R_squared	93.8%	94.5%	95.5%	95.0%

Table S5. Daily PM_{2.5} model performance for steps 1-4 across spatial folds

Fold	Step 1		Step 2		Step 3		Step 4	
	R- squared	RMSE	R- squared	RMSE	R- squared	RMSE	R- squared	RMSE
1	47.83%	4.71	54.74%	4.37	57.31%	4.26	59.13%	4.14
2	47.86%	5.27	45.08%	5.36	45.73%	5.38	58.43%	4.67
3	45.49%	5.64	49.97%	5.40	53.01%	5.23	54.55%	5.18
4	31.21%	8.71	31.96%	8.65	34.76%	8.47	33.39%	8.56
5	48.91%	3.90	50.20%	3.85	52.94%	3.76	52.60%	3.75
6	47.78%	4.30	49.24%	4.24	51.23%	4.14	56.84%	3.93
7	53.33%	4.22	54.55%	4.16	55.58%	4.11	62.78%	3.77
8	63.38%	6.36	67.14%	6.04	68.75%	5.91	73.08%	5.43
9	61.04%	9.90	64.78%	9.75	70.03%	8.88	71.34%	8.64
10	53.76%	5.58	50.36%	5.76	49.22%	5.75	58.96%	5.23
Average	50.06%	5.86	51.80%	5.76	53.85%	5.59	58.11%	5.33

Table S6. Descriptive statistics of observed (including imputed values) PM_{2.5} monitoring data vs. modelled PM_{2.5} for 2001-2020

State/Territory	PM _{2.5} (including imputed values)		PM _{2.5} (modelled)	
	Median	5th-95th percentile	Median	5th-95th percentile
ACT	5.3	1.7 - 19.5	5.7	2.5 - 17.3
NSW	6.1	1.9 - 15.3	6.4	2.5 - 14.2
NT	6.0	0.9 - 22.9	6.6	1.4 - 20.9
QLD	5.3	1.9 - 12.8	5.5	2.4 - 11.9
SA	6.3	3 - 11.9	6.5	3.7 - 11
TAS	2.8	0 - 17.3	3.0	0.9 - 16.2
VIC	6.0	2.5 - 13.9	6.2	3.3 - 13.1
WA	7.2	4 - 14.1	7.4	4.8 - 13
National	5.5	1.1 - 15.1	5.8	1.5 - 14

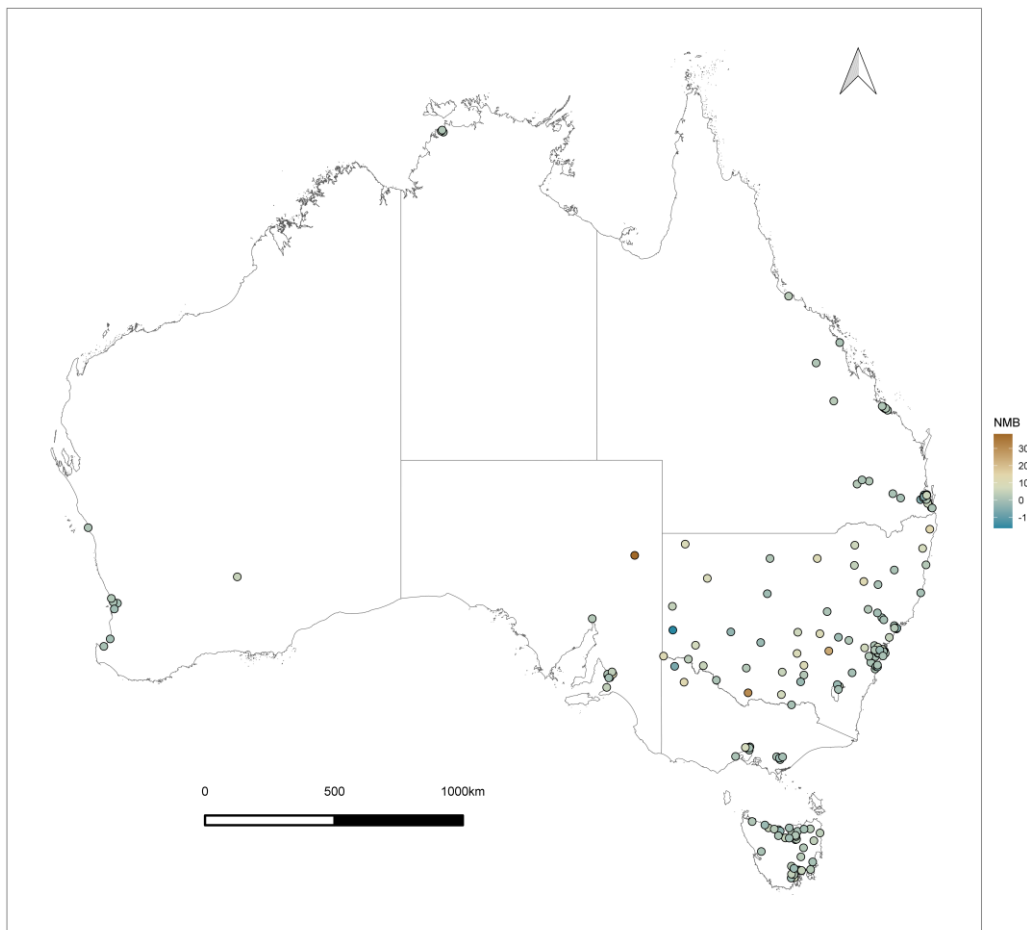


Figure S4. Normalised mean bias (NMB) between predicted and observed PM_{2.5} for each monitoring site

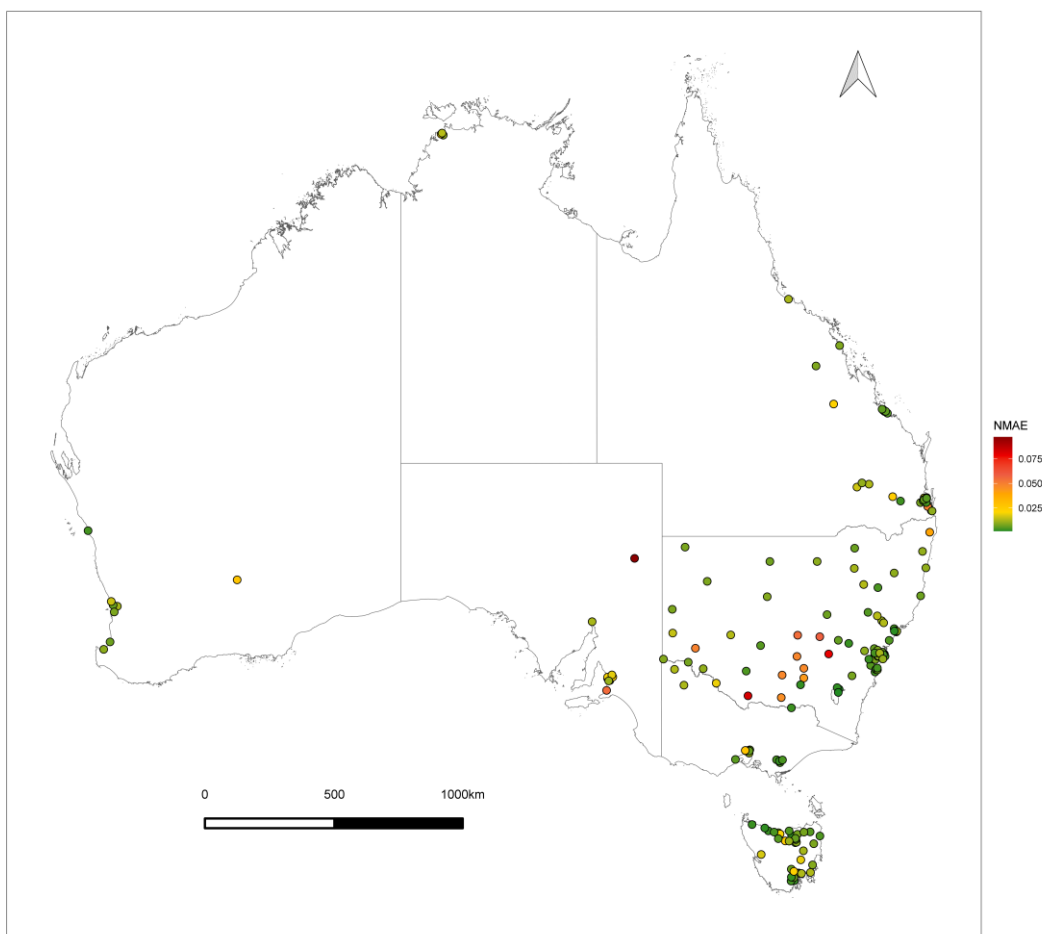


Figure S5. Normalised mean absolute error (NMAE) between predicted and observed PM_{2.5} for each monitoring site

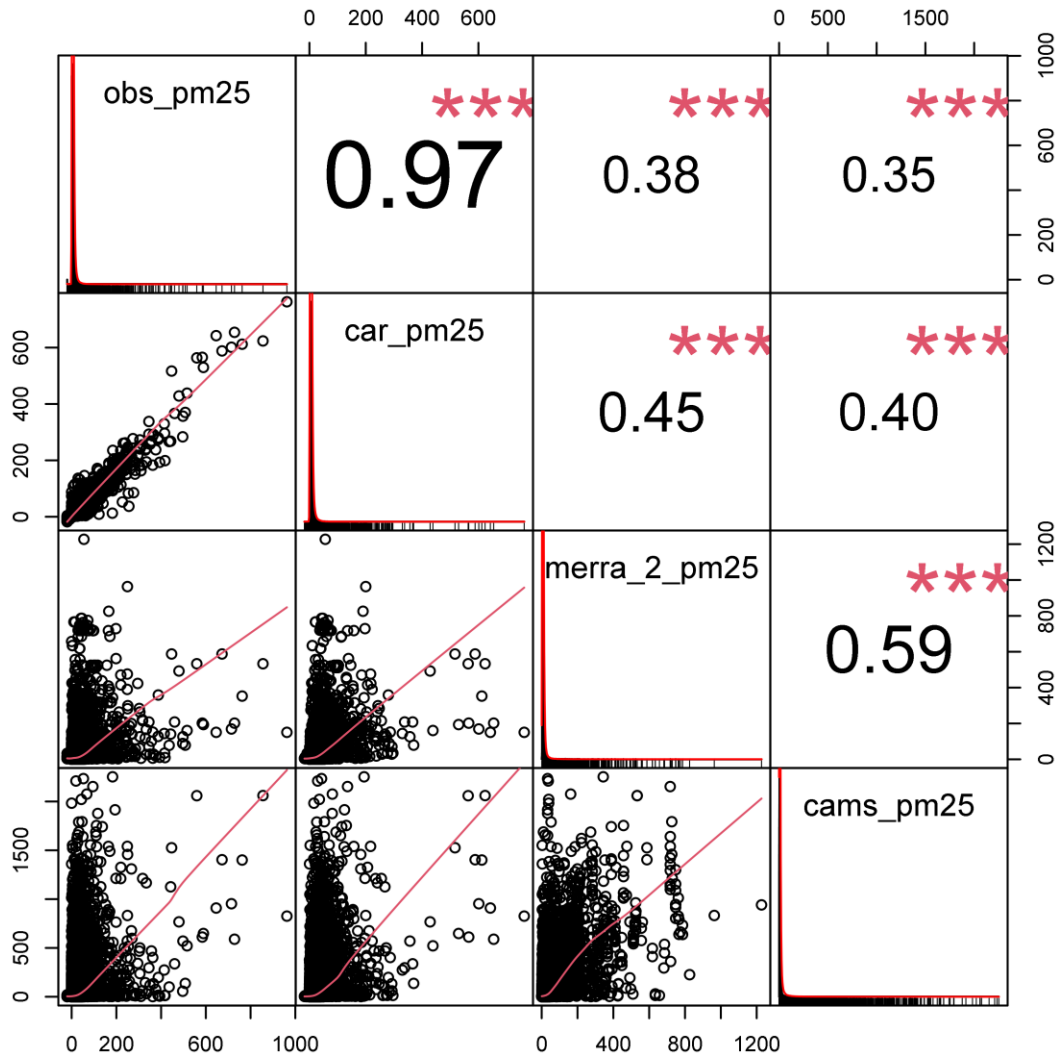


Figure S6. Correlation plots with other PM_{2.5} models – daily. Obs_pm25 = observed PM_{2.5}, car_pm25 = predicted PM_{2.5} from this study, merra_2_pm25 = PM_{2.5} from Merra-2 (Buchard et al., 2017), cams_pm25: PM_{2.5} from CAMS (Inness et al., 2019). ***: $p < 0.001$, **: p between 0.001 - 0.01, *: p between 0.01 - 0.05, •: p between 0.05 - 0.10.

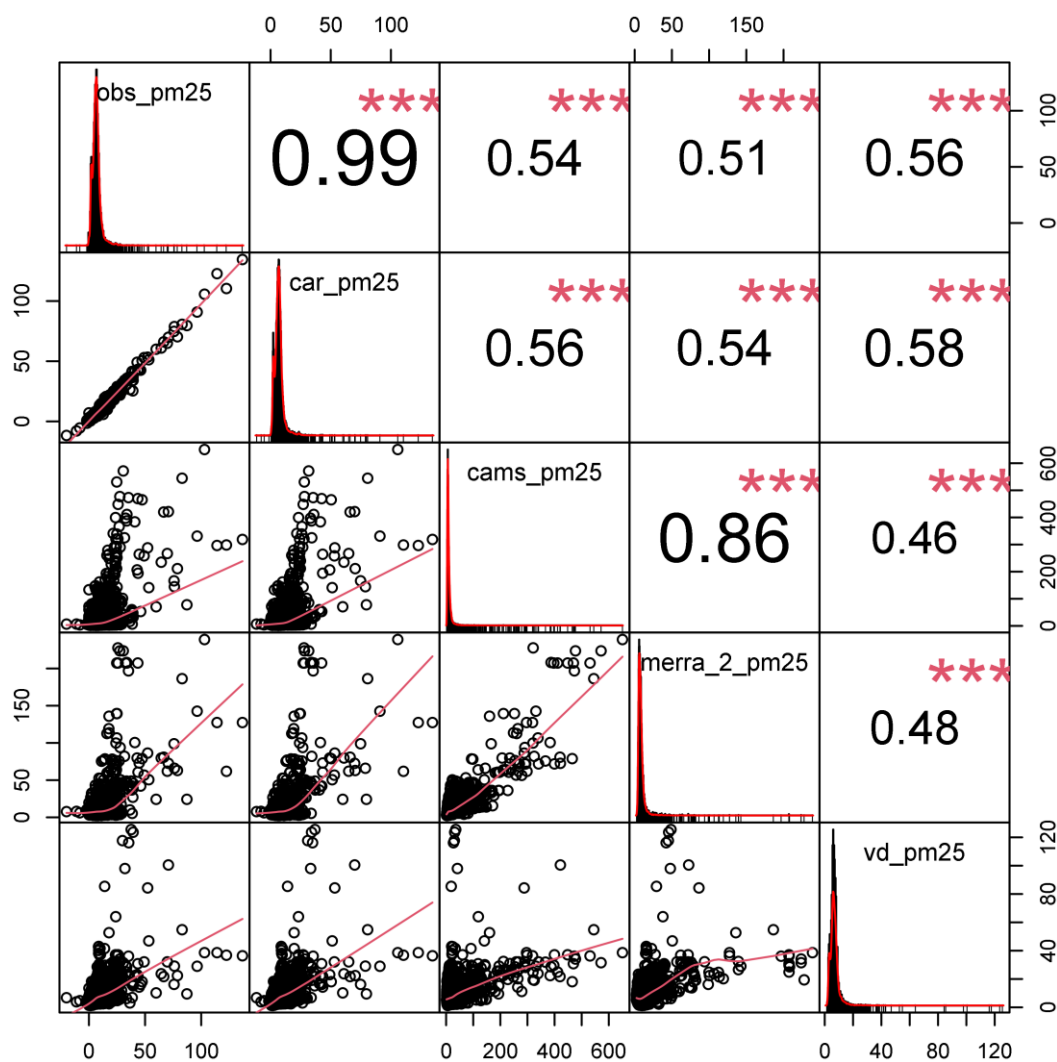


Figure S7. Correlation plots with other PM_{2.5} models - monthly. NOTE: Except for Van Donkelaar (2021), mean monthly PM_{2.5} was calculated when at least 80% of daily values were available. Obs_pm25 = observed PM_{2.5}, car_pm25 = predicted PM_{2.5} from this study, merra_2_pm25 = PM_{2.5} from Merra-2 (Buchard et al., 2017), cams_pm25: PM_{2.5} from CAMS (Inness et al., 2019), vd_pm25: PM_{2.5} from Van Donkelaar (2021). ***: $p < 0.001$, **: p between 0.001 - 0.01, *: p between 0.01 - 0.05, •: p between 0.05 - 0.10.

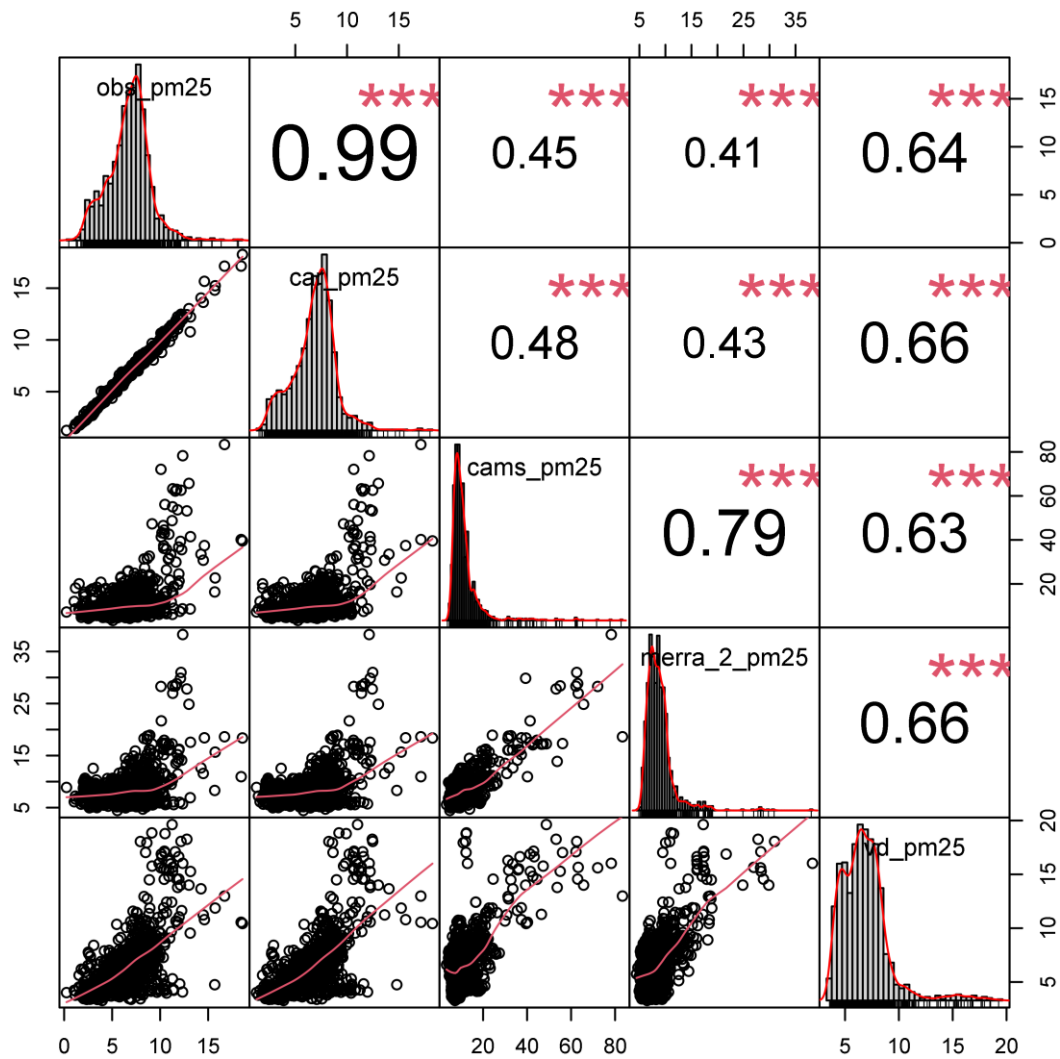


Figure S8. Correlation plots with other PM_{2.5} models - annual. NOTE: Except for Van Donkelaar (2021), mean annual PM_{2.5} was calculated when at least 80% of daily values were available. Obs_pm25 = observed PM_{2.5}, car_pm25 = predicted PM_{2.5} from this study, merra_2_pm25 = PM_{2.5} from Merra-2 (Buchard et al., 2017), cams_pm25: PM_{2.5} from CAMS (Inness et al., 2019), vd_pm25: PM_{2.5} from Van Donkelaar (2021). ***: p < 0.001, **: p between 0.001 - 0.01, *: p between 0.01 - 0.05, •: p between 0.05 - 0.10.

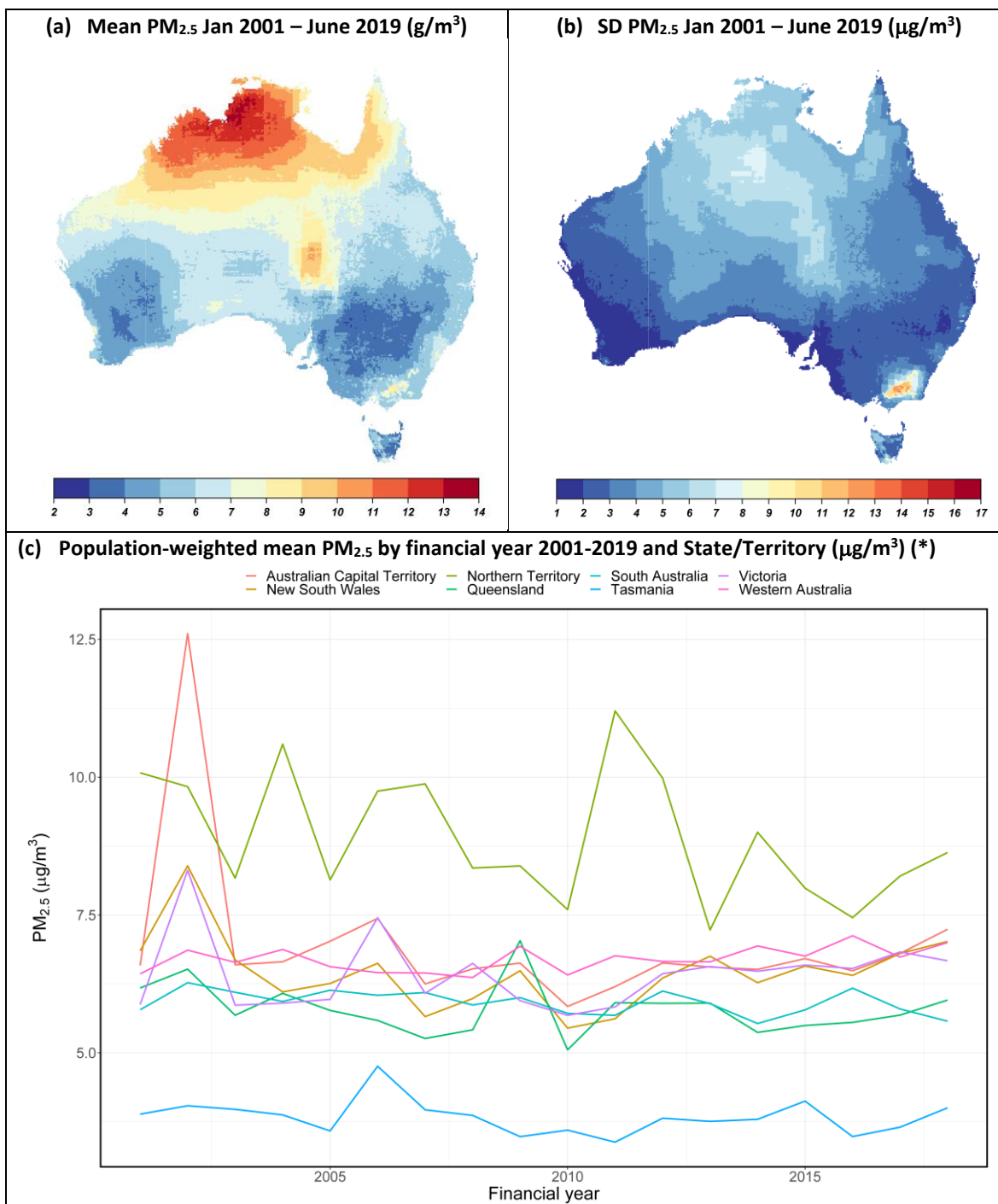


Figure S9. $PM_{2.5}$ prediction results ($\mu g/m^3$): (A) mean $PM_{2.5}$ concentrations Jan 2001 – June 2019, (B) standard deviation (SD) of $PM_{2.5}$ concentrations Jan 2001 – June 2019, (C) Population-weighted mean $PM_{2.5}$ concentration by financial year 2001-2019 and State/Territory. (*) A financial year starts on July 1 and ends on June 30 (i.e., the 2001 financial year runs from 1 July 2001 to 30 June 2002).

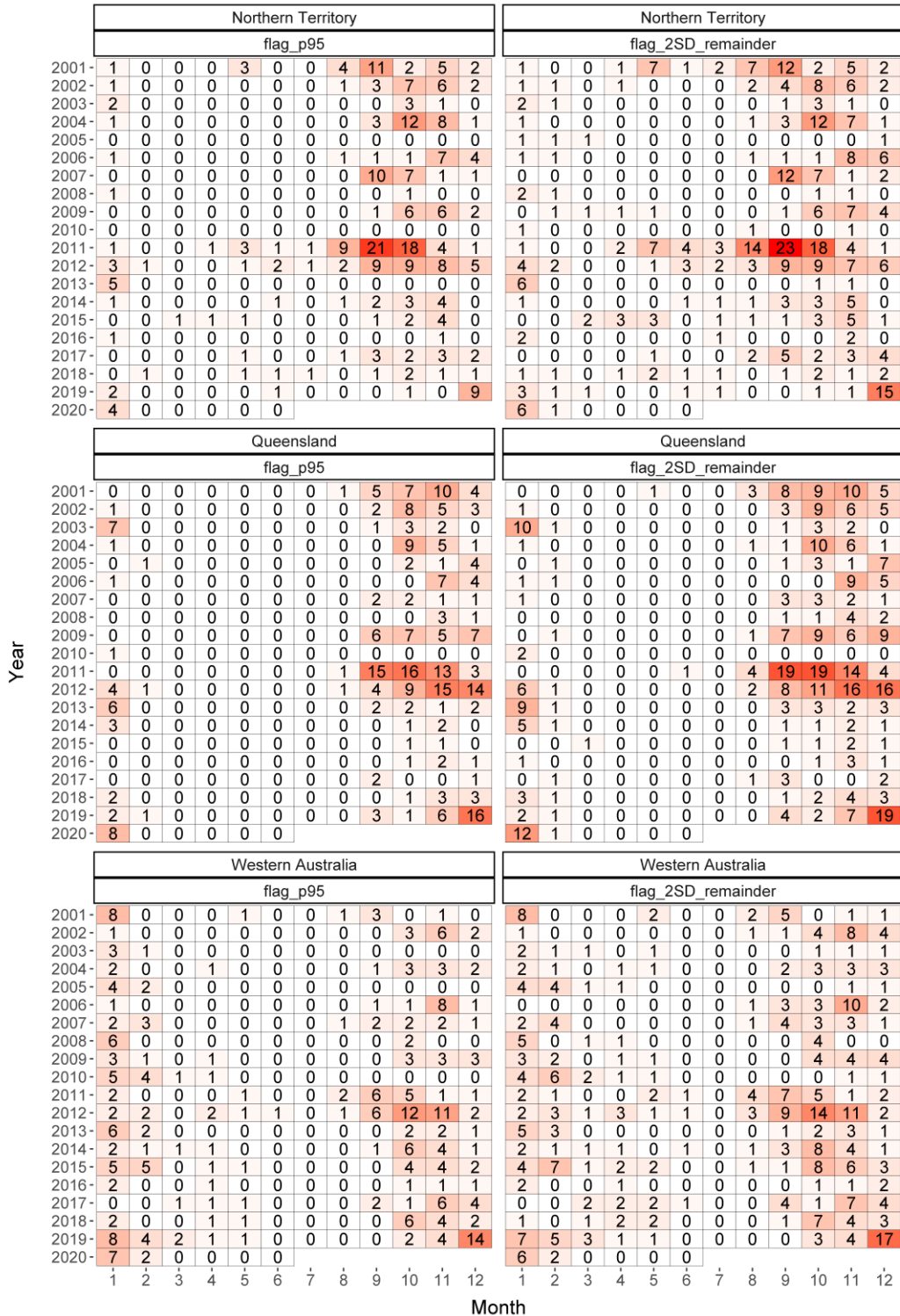


Figure S10. Mean # of extreme pollution days by year and month calculated for the Northern Territory, Queensland, and Western Australia using flag_p95 and flag_2SD_remainder

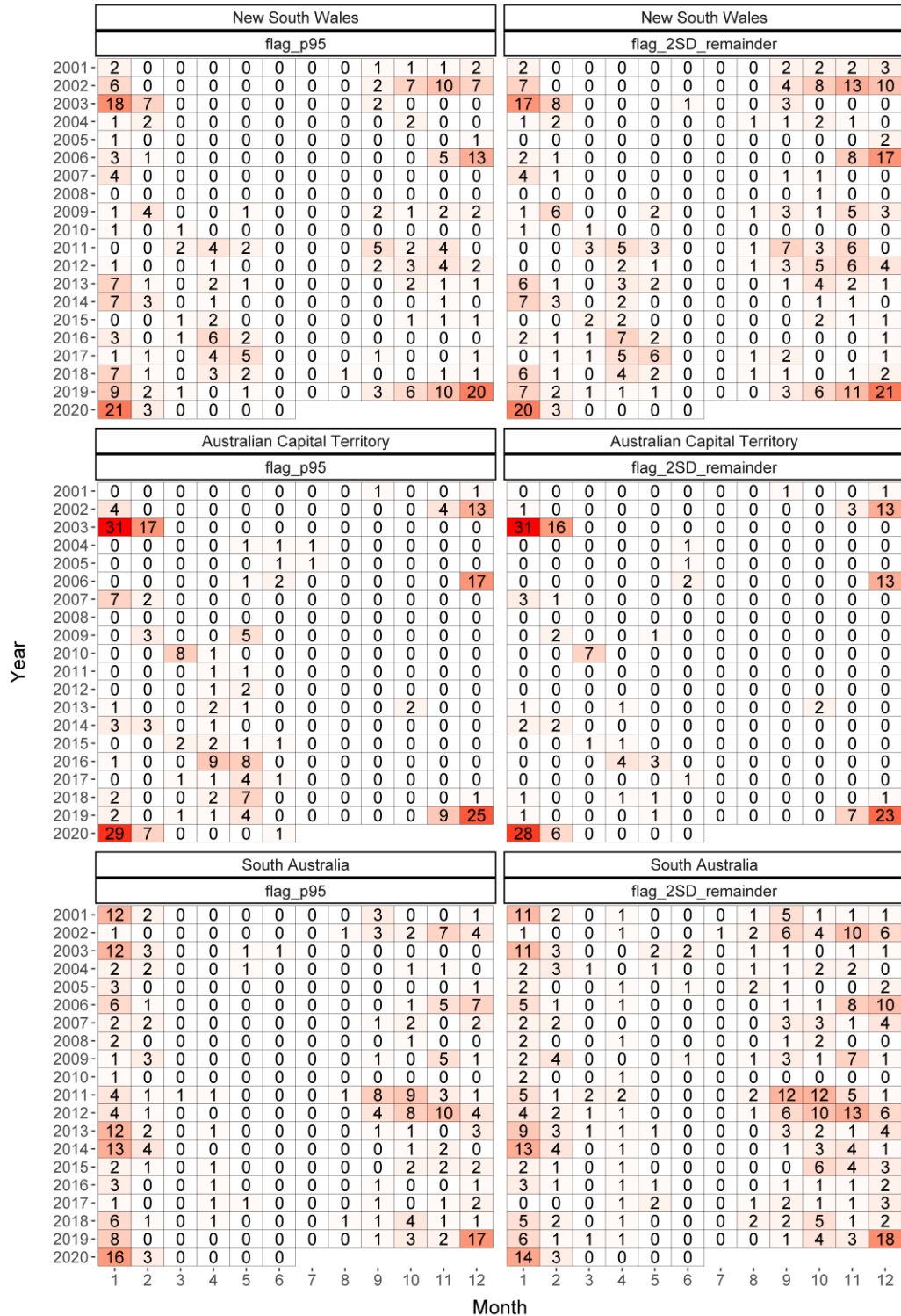


Figure S11. Mean # of extreme pollution days by year and month calculated for New South Wales, the Australian Capital Territory, and South Australia using flag_p95 and flag_2SD_remainder

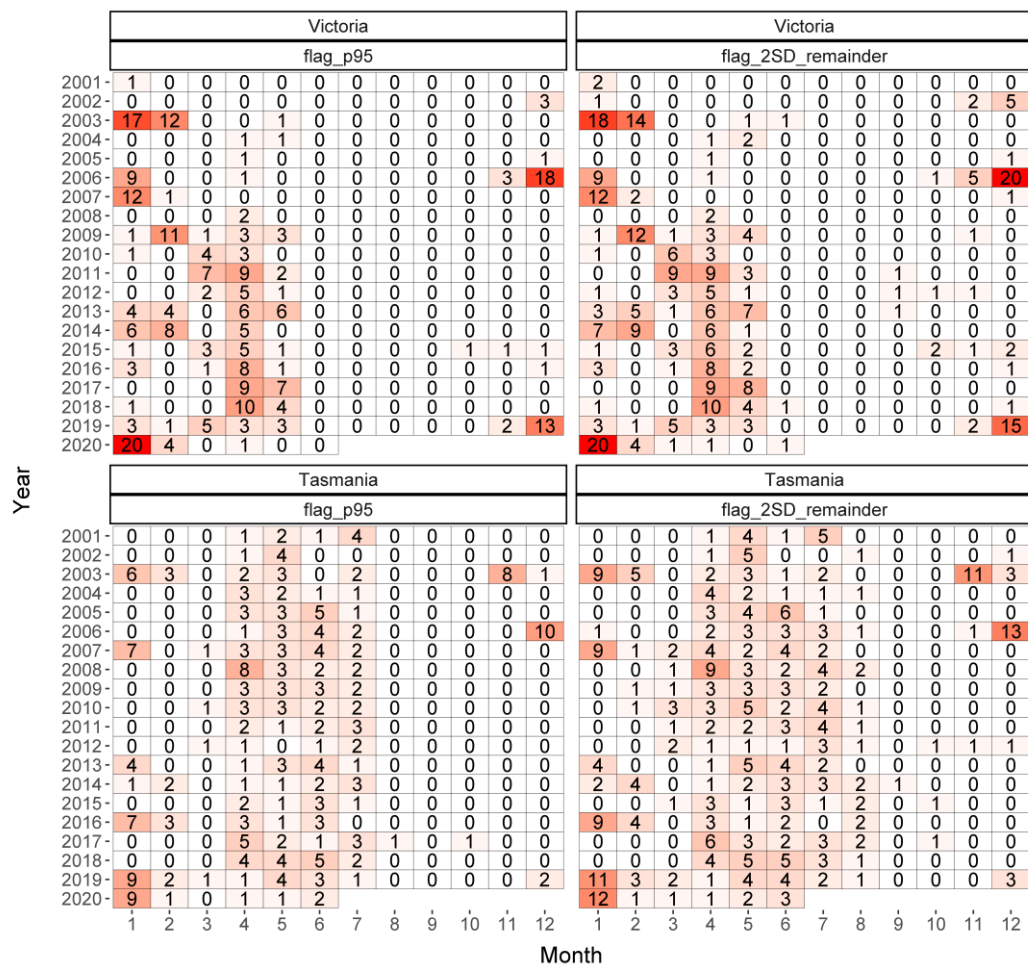


Figure S12. Mean # of extreme pollution days by year and month calculated for Victoria and Tasmania using flag_p95 and flag_2SD_remainder

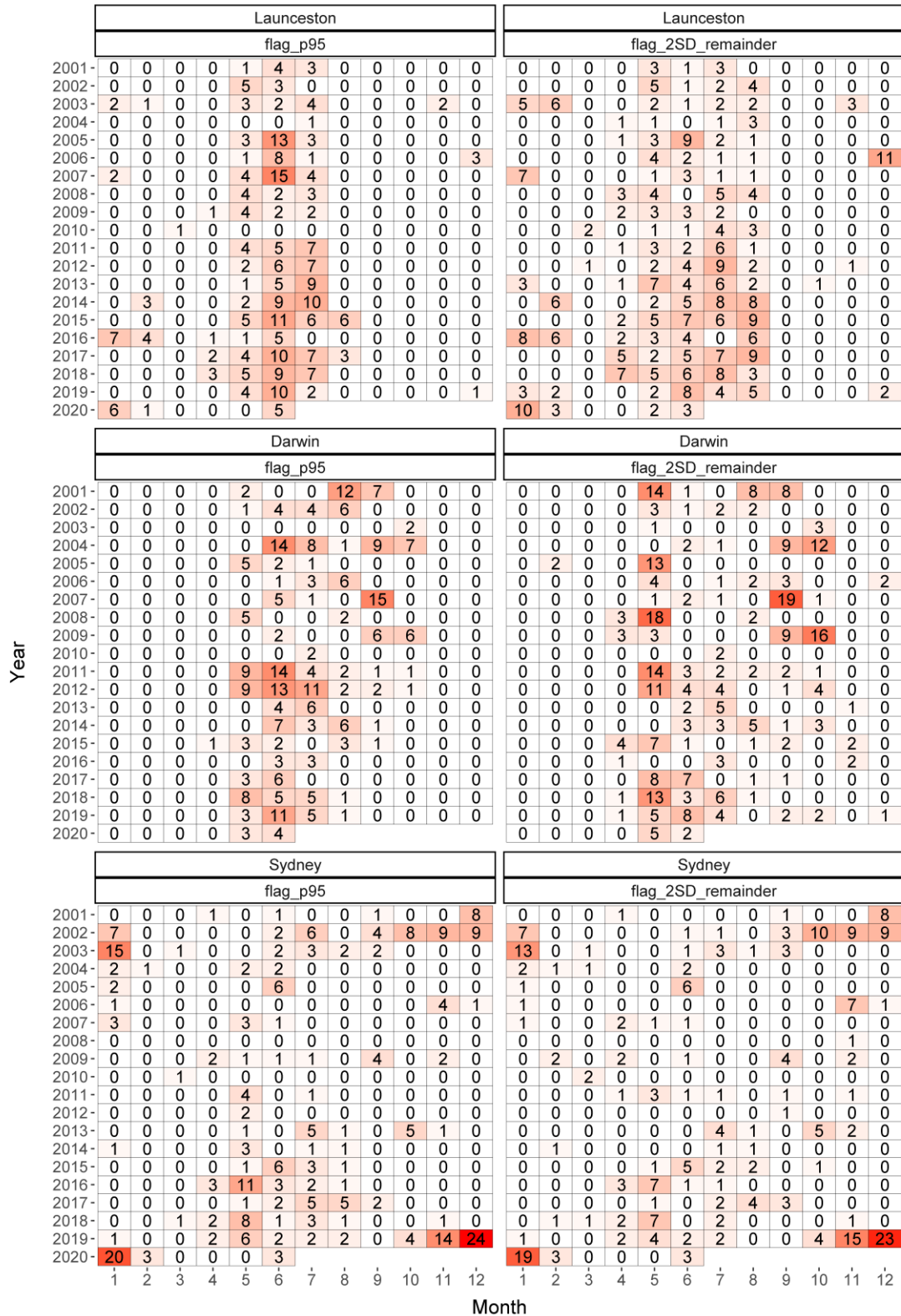


Figure S13. # of extreme pollution days by year and month calculated for Launceston, Darwin and Sydney using flag_p95 and flag_2SD_remainder

Table S7. # of extreme pollution days and mean PM_{2.5} extreme pollution component (µg/m³) for Launceston using flag_p95 and flag_2SD_remainder

year	flag_p95			flag_2SD_remainder		
	# extreme pollution days	Mean PM _{2.5} extreme pollution component on extreme pollution days (µg/m ³)	Mean PM _{2.5} extreme pollution component on all days (µg/m ³)	# extreme pollution days	Mean PM _{2.5} extreme pollution component on extreme pollution days (µg/m ³)	Mean PM _{2.5} extreme pollution component on all days (µg/m ³)
2001	8	4.2	0.09	7	6.4	0.12
2002	8	5.2	0.11	12	6.8	0.22
2003	14	8.7	0.34	21	8.6	0.50
2004	1	3.7	0.01	6	6.4	0.10
2005	19	5.8	0.30	16	7.6	0.33
2006	13	7.6	0.27	19	9.1	0.47
2007	25	4.3	0.30	13	7.7	0.27
2008	9	5.2	0.13	16	6.5	0.28
2009	9	6.1	0.15	10	7.0	0.19
2010	1	15.1	0.04	11	7.5	0.23
2011	16	5.8	0.25	13	7.3	0.26
2012	15	6.5	0.27	19	7.6	0.39
2013	15	7.1	0.29	24	7.8	0.51
2014	24	6.6	0.43	29	7.9	0.63
2015	28	6.9	0.53	29	7.9	0.62
2016	18	30.5	1.50	29	22.1	1.75
2017	26	6.1	0.43	28	7.7	0.59
2018	24	6.5	0.43	29	7.2	0.57
2019	17	6.4	0.30	26	7.7	0.54
2020 (*)	12	21.0	1.38	18	16.7	1.66

(*) PM_{2.5} values were only calculated until June 30, 2020.

Table S8. # of extreme pollution days and mean PM_{2.5} extreme pollution component (µg/m³) for Darwin using flag_p95 and flag_2SD_remainder

year	flag_p95			flag_2SD_remainder		
	# extreme pollution days	Mean PM _{2.5} extreme pollution component on extreme pollution days (µg/m ³)	Mean PM _{2.5} extreme pollution component on all days (µg/m ³)	# extreme pollution days	Mean PM _{2.5} extreme pollution component on extreme pollution days (µg/m ³)	Mean PM _{2.5} extreme pollution component on all days (µg/m ³)
2001	21	9.9	0.57	31	10.5	0.89
2002	15	6.2	0.25	8	10.2	0.22
2003	2	18.8	0.10	4	13.6	0.15
2004	39	7.7	0.82	24	10.2	0.67
2005	8	10.2	0.22	15	10.9	0.45
2006	10	5.4	0.15	12	8.0	0.26
2007	21	10.6	0.61	24	10.7	0.70
2008	7	10.1	0.19	23	9.9	0.62
2009	14	14.6	0.56	31	11.9	1.01
2010	2	8.9	0.05	2	8.9	0.05
2011	31	8.4	0.71	24	10.6	0.70
2012	38	8.8	0.91	24	12.3	0.81
2013	10	8.7	0.24	8	9.9	0.22
2014	17	8.6	0.40	15	10.0	0.41
2015	10	12.3	0.34	17	11.3	0.53
2016	6	8.7	0.14	6	9.9	0.16
2017	9	10.5	0.26	17	9.6	0.45
2018	19	14.6	0.76	24	13.8	0.90
2019	20	11.9	0.65	23	12.4	0.78
2020 (*)	7	11.7	0.45	7	12.9	0.50

(*) PM_{2.5} values were only calculated until June 30, 2020.

Table S9. # of extreme pollution days and mean PM_{2.5} extreme pollution component (µg/m³) for Sydney using flag_p95 and flag_2SD_remainder

year	flag_p95			flag_2SD_remainder		
	# extreme pollution days	Mean PM _{2.5} extreme pollution component on extreme pollution days (µg/m ³)	Mean PM _{2.5} extreme pollution component on all days (µg/m ³)	# extreme pollution days	Mean PM _{2.5} extreme pollution component on extreme pollution days (µg/m ³)	Mean PM _{2.5} extreme pollution component on all days (µg/m ³)
2001	11	29.0	0.87	10	31.2	0.85
2002	45	13.3	1.64	40	14.3	1.57
2003	25	12.6	0.86	22	13.5	0.81
2004	7	8.8	0.17	6	9.6	0.16
2005	8	9.9	0.22	7	10.3	0.20
2006	6	14.3	0.23	9	12.0	0.30
2007	7	6.5	0.13	5	7.7	0.10
2008	0	0.0	0.00	1	7.5	0.02
2009	11	29.2	0.88	11	29.5	0.89
2010	1	8.9	0.02	2	8.1	0.04
2011	5	8.1	0.11	8	8.1	0.18
2012	2	6.2	0.03	1	6.9	0.02
2013	13	11.3	0.40	12	12.0	0.40
2014	6	6.5	0.11	3	8.2	0.07
2015	11	7.8	0.24	11	8.2	0.25
2016	20	10.3	0.56	12	13.2	0.43
2017	15	9.0	0.37	10	10.4	0.28
2018	17	10.7	0.50	14	11.8	0.45
2019	57	22.9	3.57	53	24.3	3.52
2020 (*)	26	18.6	2.65	25	19.0	2.61

(*) PM_{2.5} values were only calculated until June 30, 2020.

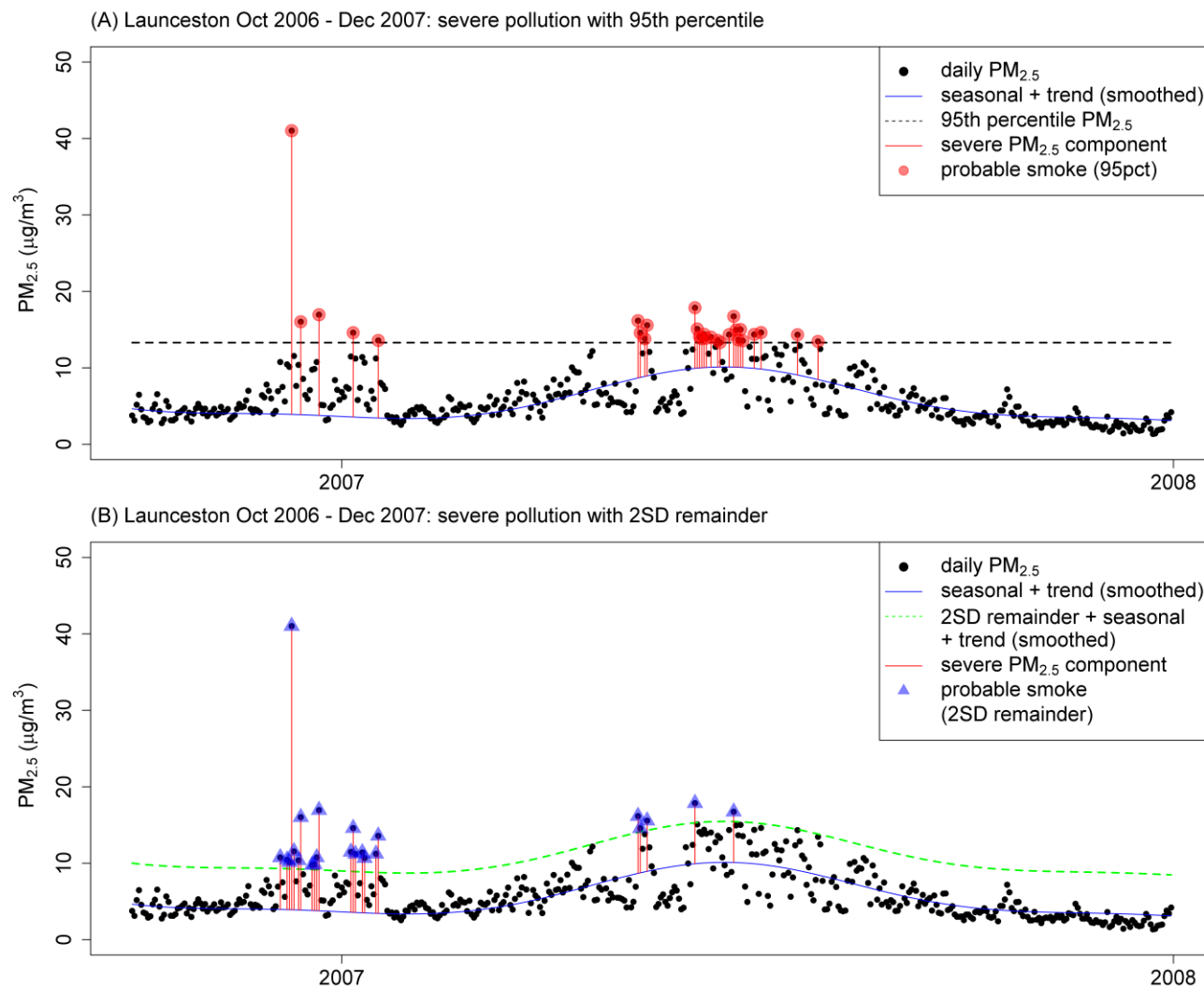


Figure S14. Example of extreme air pollution days between October 2006 – December 2007 in Launceston identified with: (a) 95th percentile and (b) 2SD remainder flags.
NOTE: For illustration purposes “seasonal + trend” and “2SD remainder + seasonal + trend” time series have been smoothed, and probable smoke days flagged with these.

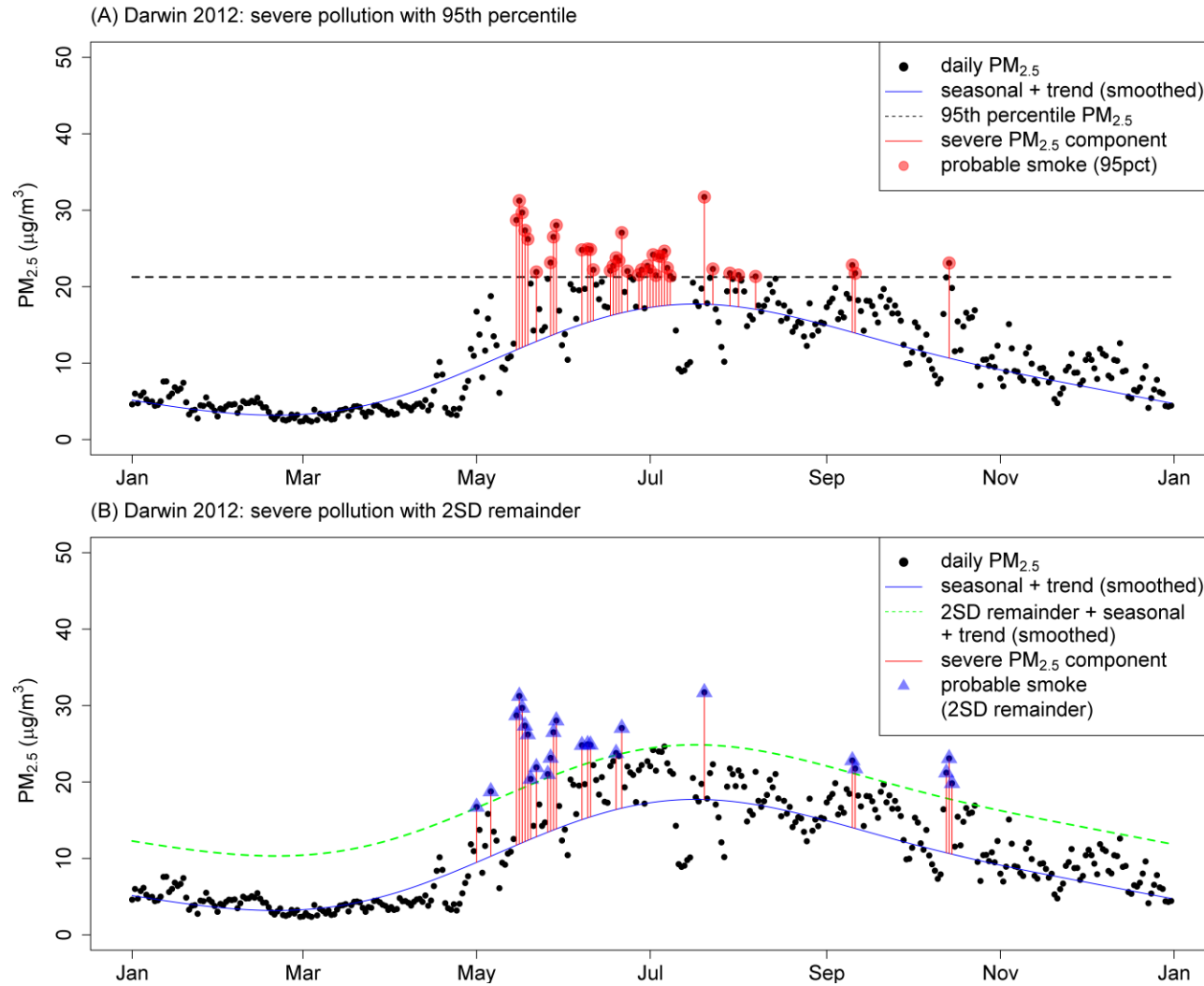


Figure S15. Example of extreme air pollution days between January 2012 – December 2012 in Darwin identified with: (a) 95th percentile and (b) 2SD remainder flags. NOTE: For illustration purposes “seasonal + trend” and “2SD remainder + seasonal + trend” time series have been smoothed, and probable smoke days flagged with these.

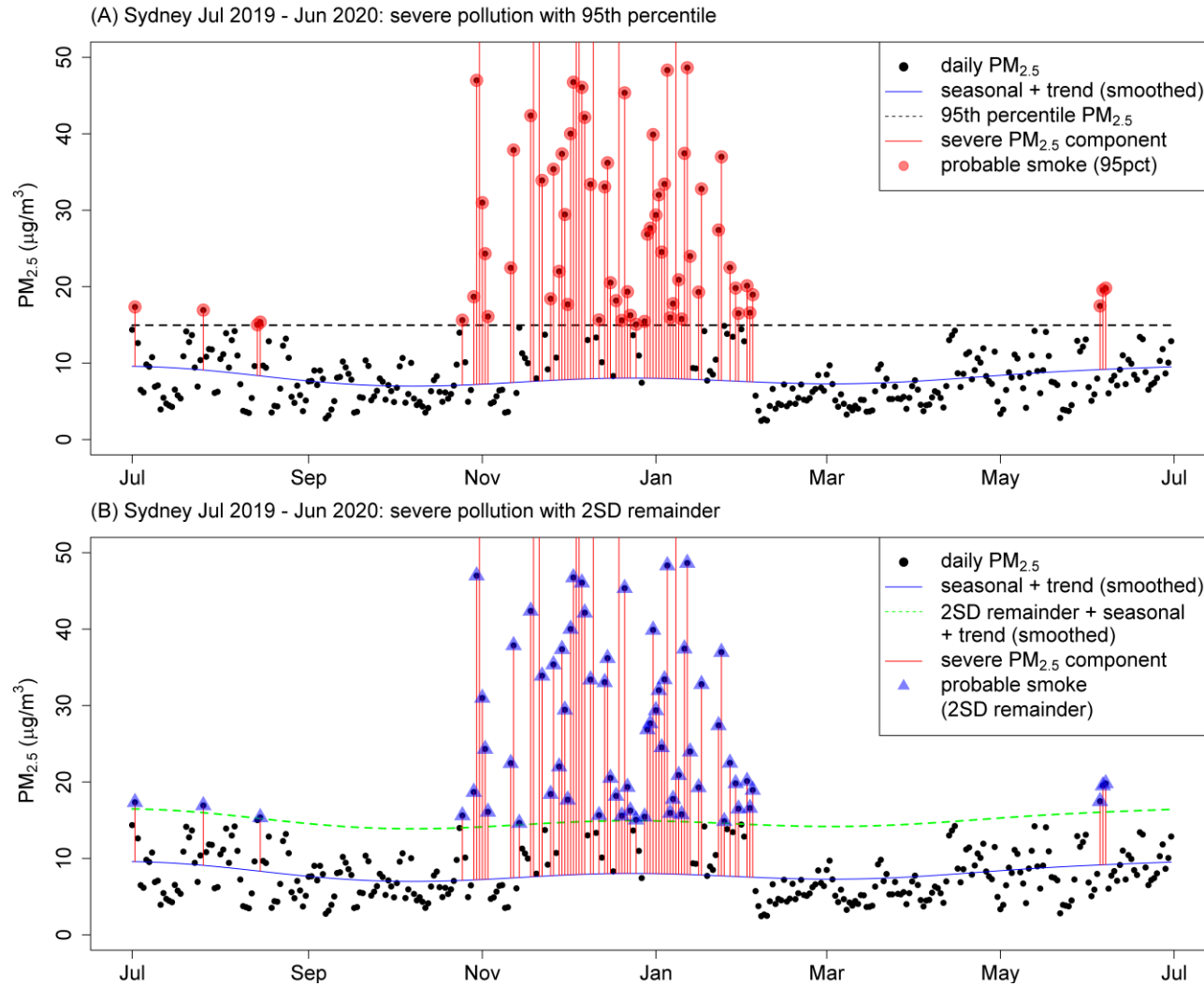


Figure S16. Example of extreme air pollution days between July 2019 – June 20 in Sydney identified with: (a) 95th percentile and (b) 2SD remainder flags. NOTE: For illustration purposes “seasonal + trend” and “2SD remainder + seasonal + trend” time series have been smoothed, and probable smoke days flagged with these.

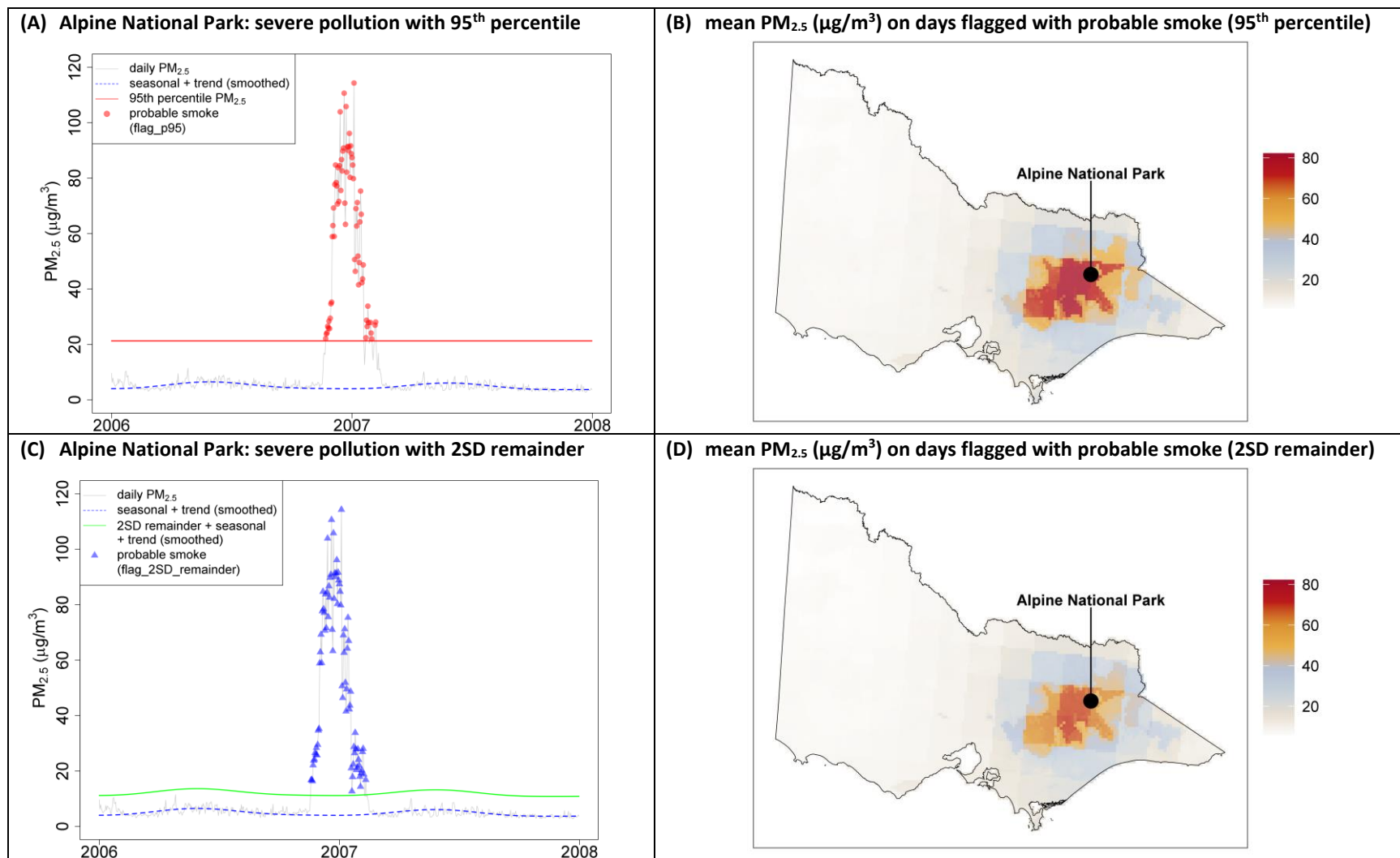


Figure S17. Temporal and spatial example of extreme air pollution during the 2006-07 Eastern Victoria Great Divide bushfires in Victoria: (a) daily $PM_{2.5}$ (95th percentile flag) at Alpine National Park, (b) daily $PM_{2.5}$ (2SD remainder flag) at Alpine National Park, (c) mean $PM_{2.5}$ on days flagged with probable smoke (95th percentile flag), and (d) mean $PM_{2.5}$ on days flagged as probable smoke (2SD remainder flag). NOTE: For illustration purposes “seasonal + trend” and “2SD remainder + seasonal + trend” time series have been smoothed, and probable smoke days flagged with these.

References

- Australian Bureau of Statistics. (2011). Environmental Issues: Energy Use and Conservation - 4602.0.55.001 (Dataset). ABS.
<http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/4602.0.55.001Mar2011?OpenDocument>
- Australian Bureau of Statistics. (2021). Regional population. <https://www.abs.gov.au/statistics/people/population/regional-population/latest-release>
- Beguería, S., & Vicente-Serrano, S. M. (2023). SPEI: Calculation of the Standardized Precipitation-Evapotranspiration Index. <https://spei.csic.es>
- Beguería, S., Vicente-Serrano, S. M., Reig, F., & Latorre, B. (2014). Standardized precipitation evapotranspiration index (SPEI) revisited: parameter fitting, evapotranspiration models, tools, datasets and drought monitoring. *International Journal of Climatology*, 34(10), 3001–3023. <https://doi.org/10.1002/JOC.3887>
- Brown de Colstoun, E. C., C. Huang, P. Wang, J. C. Tilton, B. Tan, J. Phillips, S. Niemczura, P.-Y. Ling, and R. E. Wolfe. 2017. Global Man-made Impervious Surface (GMIS) Dataset From Landsat. Palisades, New York: NASA Socioeconomic Data and Applications Center (SEDAC). <https://doi.org/10.7927/H4P55KKF>. Downloaded from CARDAT https://cloud.cardat.org/index.php/apps/files/?dir=/Environment_General/Landsat_Global_High_Resolution_Urban_Data/Landsat_GMIS_v1_Australia_2010
- Buchard, V., da Silva, A. M., Randles, C. A., Colarco, P., Ferrare, R., Hair, J., Hostetler, C., Tackett, J., & Winker, D. (2016). Evaluation of the surface PM_{2.5} in Version 1 of the NASA MERRA Aerosol Reanalysis over the United States. *Atmospheric Environment*, 125, 100–111. <https://doi.org/10.1016/j.atmosenv.2015.11.004>
- Buchard, V., Randles, C. A., da Silva, A. M., Darmenov, A., Colarco, P. R., Govindaraju, R., Ferrare, R., Hair, J., Beyersdorf, A. J., Ziemba, L. D., & Yu, H. (2017). The MERRA-2 Aerosol Reanalysis, 1980 Onward. Part II: Evaluation and Case Studies. *Journal of Climate*, 30(17), 6851–6872. <https://doi.org/10.1175/JCLI-D-16-0613.1>
- Bureau of Meteorology. (2020). Gridded climatology data. http://www.bom.gov.au/climate/averages/climatology/gridded-data-info/gridded_datasets_summary.shtml
- Centre for Safe Air. (2021). National Air Pollution Monitoring Database, derived from regulatory monitor data from NSW DPIE, Vic EPA, Qld DES, SA EPA, WA DEWR, Tas EPA, NT EPA, and ACT Health. CAR. <https://doi.org/10.17605/OSF.IO/JXD98>
- Department of Agriculture, Water and the Environment (2022): National Pollutant Inventory 1998-2020. Australian Government. Data downloaded from CARDAT. (Dataset). https://cloud.cardat.org/index.php/apps/files/?dir=/Environment_General/NPI/NPI_1998_2020
- Didan, K. (2015): MOD13A3 MODIS/Terra vegetation Indices Monthly L3 Global 1km SIN Grid V006. NASA EOSDIS Land Processes DAAC. (Dataset). <https://doi.org/10.5067/MODIS/MOD13A3.006>. Retrieved from Centre for Air pollution,

energy and health Research https://cloud.cardat.org/index.php/apps/files/?dir=/Environment_General/NDVI_MODIS

DiMiceli, C., Carroll, M., Sohlberg, R., Kim, D., Kelly, M., Townshend, J. (2015): MOD44B MODIS/Terra Vegetation Continuous Fields Yearly L3 Global 250m SIN Grid V006. NASA EOSDIS Land Processes DAAC. (Dataset). <https://doi.org/10.5067/MODIS/MOD44B.006>. Downloaded from CARDAT https://cloud.cardat.org/index.php/apps/files/?dir=/Environment_General/Land_Cover_MODIS/tree_and_water_coverage_2017_2018

Geoscience Australia. (2011). Euclidean Distance to the Australian Coastline. <https://pid.geoscience.gov.au/dataset/ga/71990>, <https://researchdata.edu.au/euclidean-distance-australian-coastline>

Gallant, J., Wilson, N., Dowling, T., Read, A., Inskip, C. (2011): SRTM-derived 1 Second Digital Elevation Models Version 1.0. Record 1. Geoscience Australia, Canberra. (Dataset). <http://pid.geoscience.gov.au/dataset/ga/72759>. Downloaded from CARDAT https://cloud.cardat.org/index.php/apps/files/?dir=/Environment_General/Digital_Elevation_Model/SRTM_1sec_Smooth_DEM_National

Giglio, L., Boschetti, L., Roy, D. P., Humber, M. L., & Justice, C. O. (2018). The Collection 6 MODIS burned area mapping algorithm and product. *Remote Sensing of Environment*, 217, 72–85. <https://doi.org/10.1016/J.RSE.2018.08.005>

Hanigan, I. & Learnihan, V. derived from ABS data (2017): Australian Statistical Geography Standard (ASGS) Volume 1 - Main Structure and Greater Capital City Statistical Areas, July 2016, cat no. 1270.0.55.001. ABS. Downloaded from CARDAT (dataset). https://cloud.cardat.org/index.php/apps/files/?dir=/Environment_General/ABS_data/ABS_meshblocks

Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.-M., Dominguez, J. J., Engelen, R., Eskes, H., Flemming, J., Huijnen, V., Jones, L., Kipling, Z., Massart, S., Parrington, M., Peuch, V.-H., Razinger, M., Remy, S., Schulz, M., & Suttie, M. (2019). The CAMS reanalysis of atmospheric composition. *Atmos. Chem. Phys.*, 19(6), 3515–3556. <https://doi.org/10.5194/acp-19-3515-2019>

Muñoz Sabater, J. (2019). ERA5-Land hourly data from 1950 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). <https://doi.org/10.24381/cds.e2161bac>

NASA_FIRMS. (2021). MODIS Collection 61 NRT Hotspot / Active Fire Detections MCD14DL distributed from NASA FIRMS. Available on-line <https://earthdata.nasa.gov/firms>. doi:10.5067/FIRMS/MODIS/MCD14DL.NRT.0061

Riley, M., Kirkwood, J., Jiang, N., Ross, G., & Scorgie, Y. (2020). Air quality monitoring in NSW: From long term trend monitoring to integrated urban services. *Air Quality and Climate Change*, 54(1), 44–51. <https://search.informit.org/doi/10.3316/informit.078202598997117>

Van Donkelaar, A., Hammer, M. S., Bindle, L., Brauer, M., Brook, J. R., Garay, M. J., Hsu, N. C., Kalashnikova, O. V., Kahn, R. A., Lee, C., Levy, R. C., Lyapustin, A., Sayer, A. M., & Martin, R. V. (2021). Monthly Global Estimates of Fine Particulate Matter and Their Uncertainty.

Environmental Science and Technology, 55(22), 15287–15300.
https://doi.org/10.1021/ACS.EST.1C05309/ASSET/IMAGES/LARGE/ES1C05309_0007.JPEG

Vicente-Serrano, S. M., Beguería, S., & López-Moreno, J. I. (2010). A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index. *Journal of Climate*, 23(7), 1696–1718. <https://doi.org/10.1175/2009JCLI2909.1>

Yuen, C. (2020): Traffic load intensity derived from Zenith 2018 traffic counts and PSMA 2019 road networks. CAR. (Dataset). https://cloud.car-dat.org/index.php/apps/files/?dir=/ResearchProjects_CAR/Traffic_Load/Traffic_Load_Australia_2018