atmosphere

MDPI

Supplementary Materials:

Ningbo Jiang *, Matthew L. Riley, Merched Azzi, Praveen Puppala, Hiep Duc and Giovanni Di Virgilio

To support our interpretations in the main text, here we show some additional results from analyses on the all-day (full) dataset, in Figure S1 on general data distributions (box-plots), Figure S2 on the clustering of monitoring stations into two air quality subregions, and in Figures S9 and 10 on visualisation of mean PM10 concentrations and number of poor air quality days by station and subregion and by each year and month, respectively. Figures S5 to S8 show example results for Singleton (in SE subregion) and Muswellbrook (in WNW subregion), from the wavelet analyses applied to the PM10 data at individual stations, confirming the temporal variability patterns identified from two dominant PCs.

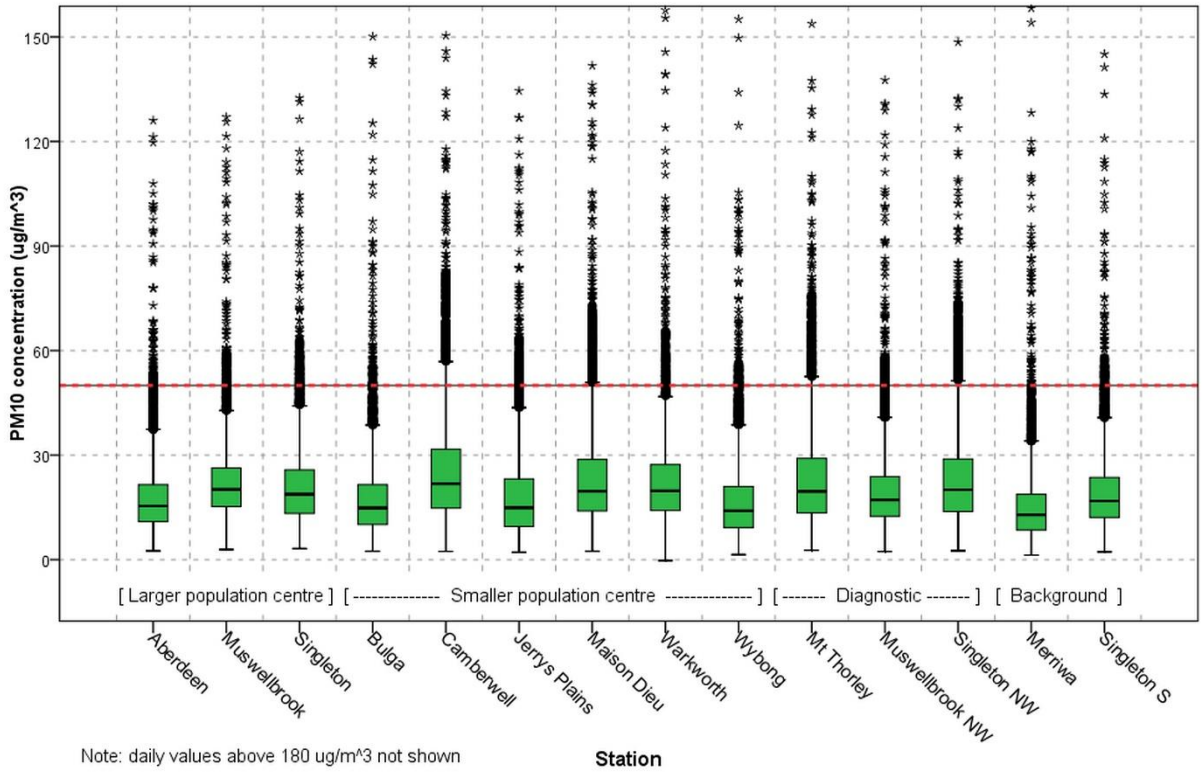

Figure S1. Box plots of daily PM_{10} data by station for all-day dataset in 2012–2022 (including data for exceptional event days). The lower and upper boundaries of the green bar (box) are respectively the 25th and 75th percentile; the horizontal line inside the box represents the median; asterisks represent extreme values, cases with values more than 3 box-lengths from the upper or lower edge of the box; dots denote outliers, cases with values between 1.5 and 3 box-lengths from the upper or lower edge of the box; horizontal lines connected to two ends of the box correspond to the largest or smallest observed values that are not outliers. Red dashed line shows the Australian national standard of 50 $\mu g/m^3$ for daily PM10.

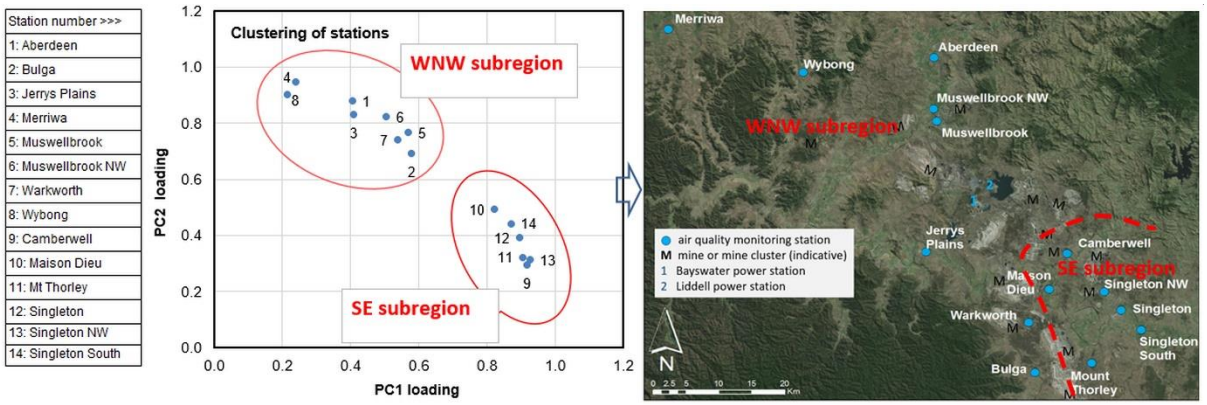

Figure S2. Identification of two air quality clusters/subregions in the Upper Hunter Valley based on Varimax rotated principal component analysis (RPCA) on daily PM10 data for 2012-2022 (including data for exceptional event days). Left panel: key of station number; middle panel: map showing the UHAQMN stations separated into the WNW and SE subregions. White/grey nugget areas on map represent locations of open-cut mining, with "M" indicative of current active mine site or cluster. The Liddell Power Station was decommissioned from April 2022 to April 2023. Base map source: Google.

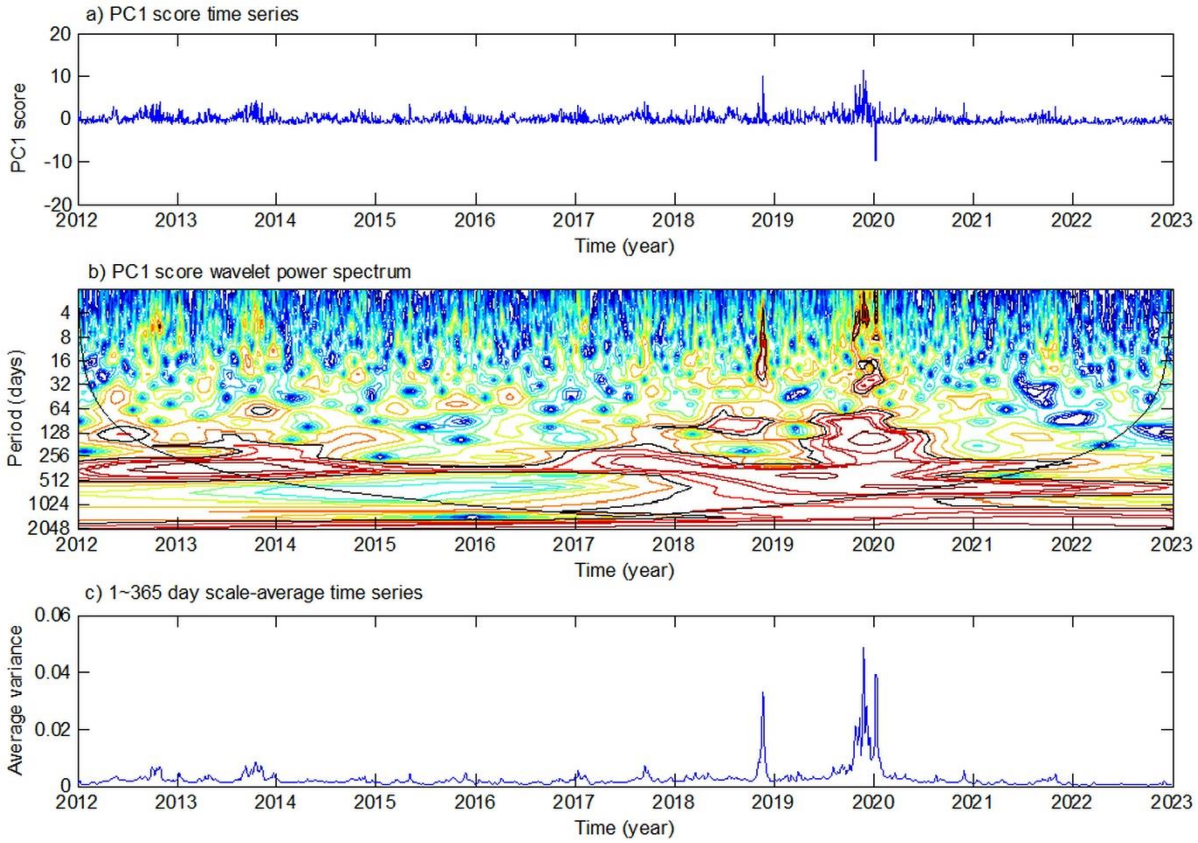

Figure S3. SE subregion temporal variability patterns derived from RPCA of all-day dataset (exceptional event data included). (a) The first principal component (PC1) scores used for the wavelet analysis, where missing data were replaced with overall median for each station. (b) The local normalised wavelet power spectrum of (a) using the Morlet wavelet. The contour lines are at normalised variances of low to high values shown in dark (blue) to bright (light) colours. The thick black contour encloses regions of greater than 95% confidence for a red-noise process with lag-1 coefficient. Regions under the bowl-shape curve on either end indicates the "cone of influence", where edge effects become important. (c) The scale-averaged wavelet power (variance) over the 1-365 days band for PC1 scores.

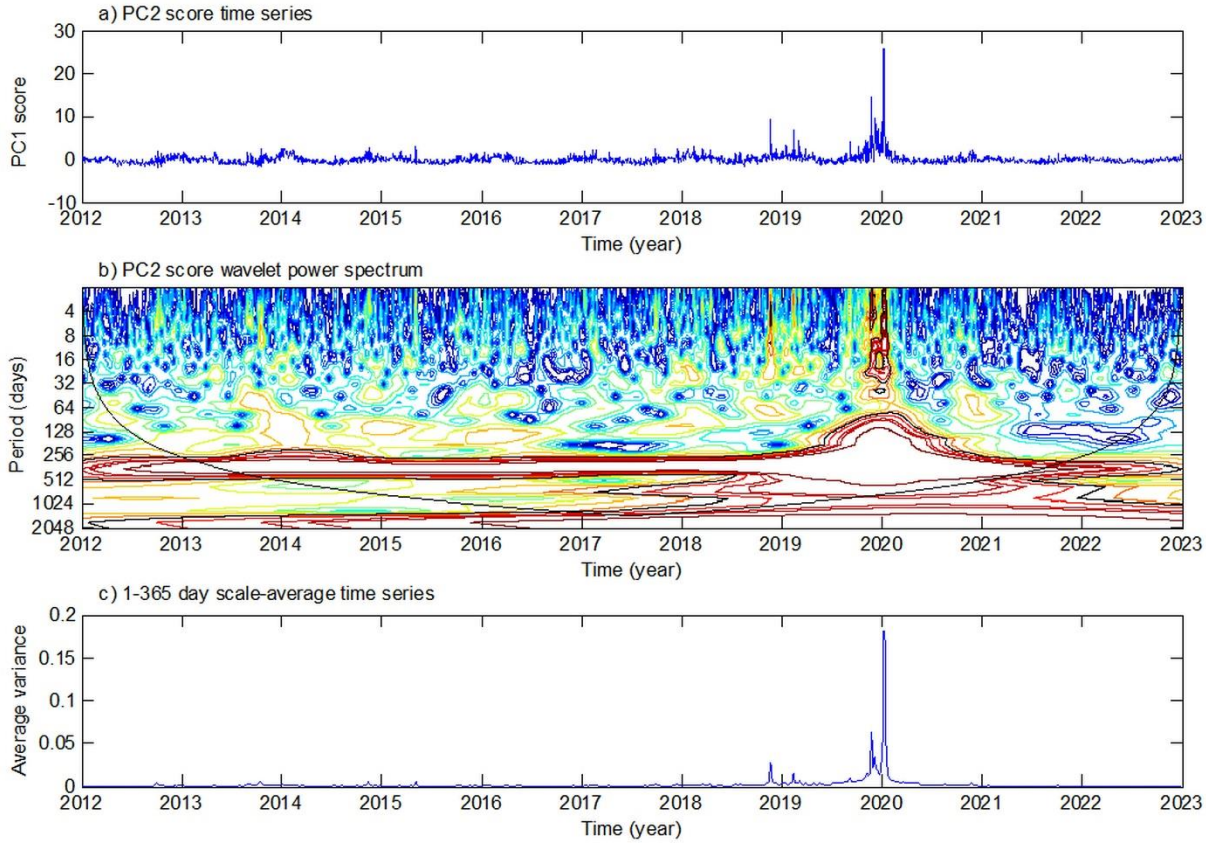

Figure S4. WNW subregion temporal variability patterns derived from all-day dataset (exceptional event data included). (a) The second principal component (PC2) scores used for the wavelet analysis, where missing data were replaced with overall median for each station. (b) The local normalised wavelet power spectrum of (a) using the Morlet wavelet. The contour lines are at normalised variances of low to high values shown in dark (blue) to bright (light) colours. The thick black contour encloses regions of greater than 95% confidence for a red-noise process with lag-1 coefficient. Regions under the bowl-shape curve on either end indicates the "cone of influence", where edge effects become important. (c) The scale-averaged wavelet power (variance) over the 1-365 days band for PC2 scores.

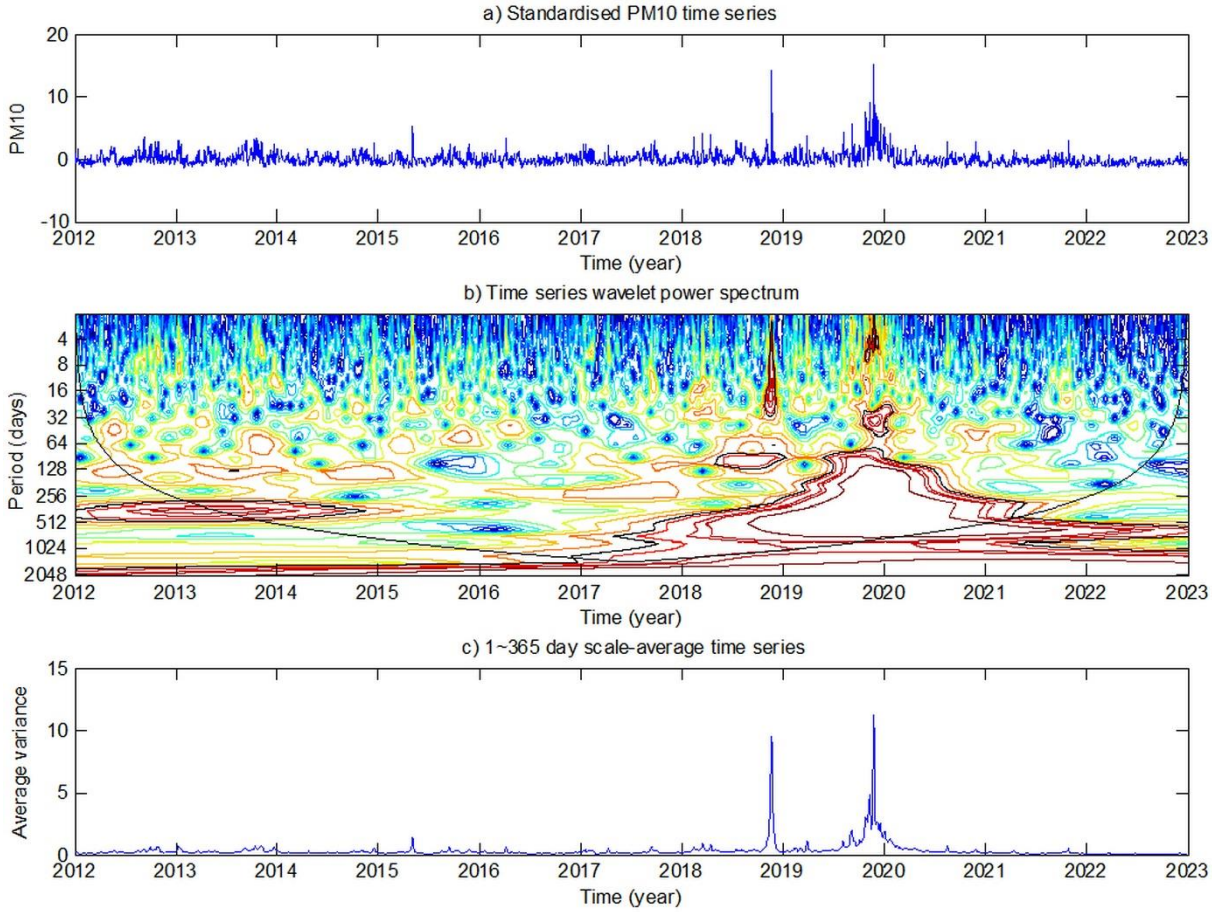

Figure S5. Singleton PM10 temporal variability patterns derived from the all-day data (exceptional event day measurements included). (a) Standardised PM10 time series used for the wavelet analysis, where missing data were replaced with the dataset median for the station. (b) The local normalised wavelet power spectrum of (a) using the Morlet wavelet. The contour lines are at normalised variances of low to high values shown in dark (blue) to bright (light) colours. The thick black contour encloses regions of greater than 95% confidence for a red-noise process with lag-1 coefficient. Regions under the bowl-shape curve on either end indicates the "cone of influence", where edge effects become important. (c) The scale-averaged wavelet power (variance) over the 1-365 days band for the standardised PM10 time series.

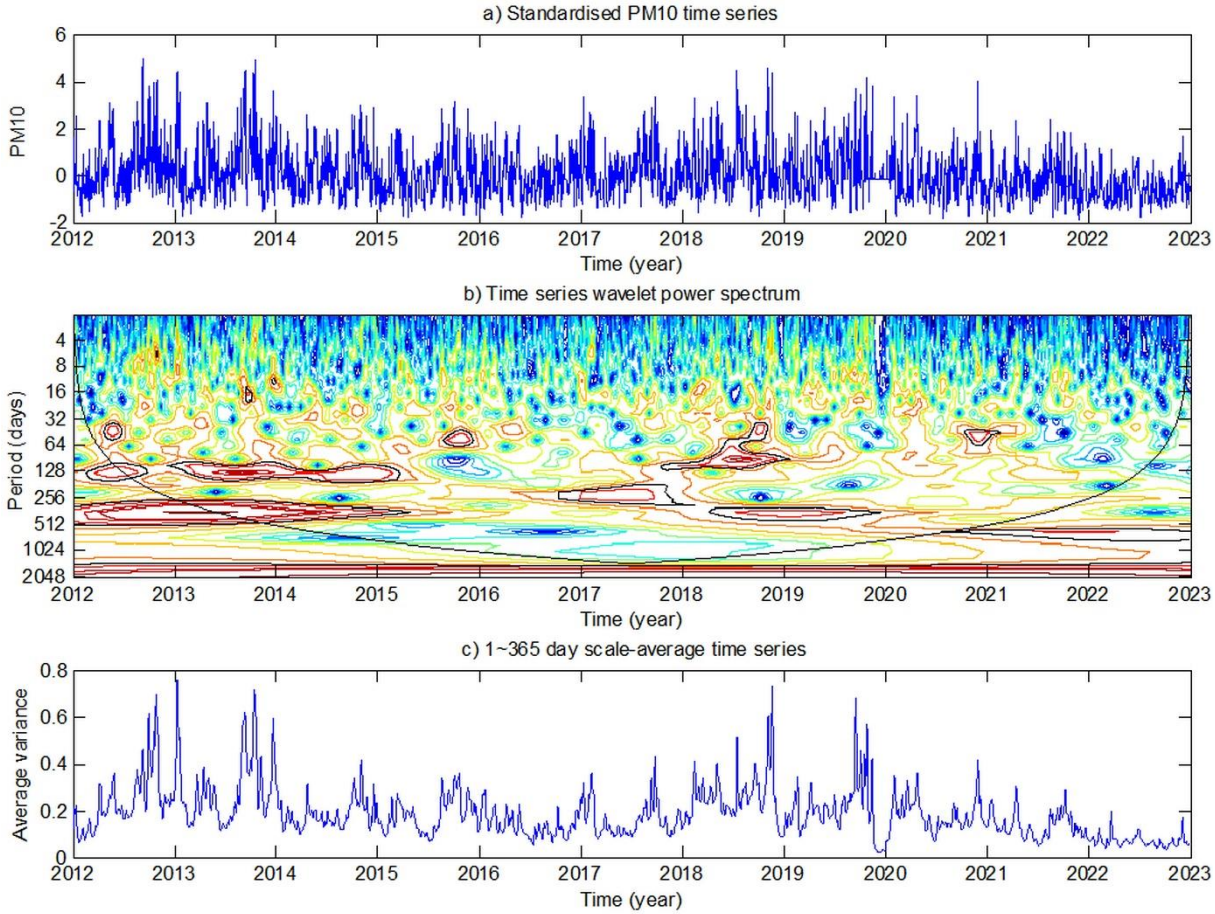

Figure S6. Singleton PM10 temporal variability patterns derived from the normal-day data (exceptional event day measurements excluded). (a) Standardised PM10 time series used for the wavelet analysis, where missing data and records for exceptional event days were replaced with the dataset median for the station. (b) The local normalised wavelet power spectrum of (a) using the Morlet wavelet. The contour lines are at normalised variances of low to high values shown in dark (blue) to bright (light) colours. The thick black contour encloses regions of greater than 95% confidence for a red-noise process with lag-1 coefficient. Regions under the bowl-shape curve on either end indicates the "cone of influence", where edge effects become important. (c) The scale-averaged wavelet power (variance) over the 1-365 days band for the standardised PM10 time series.

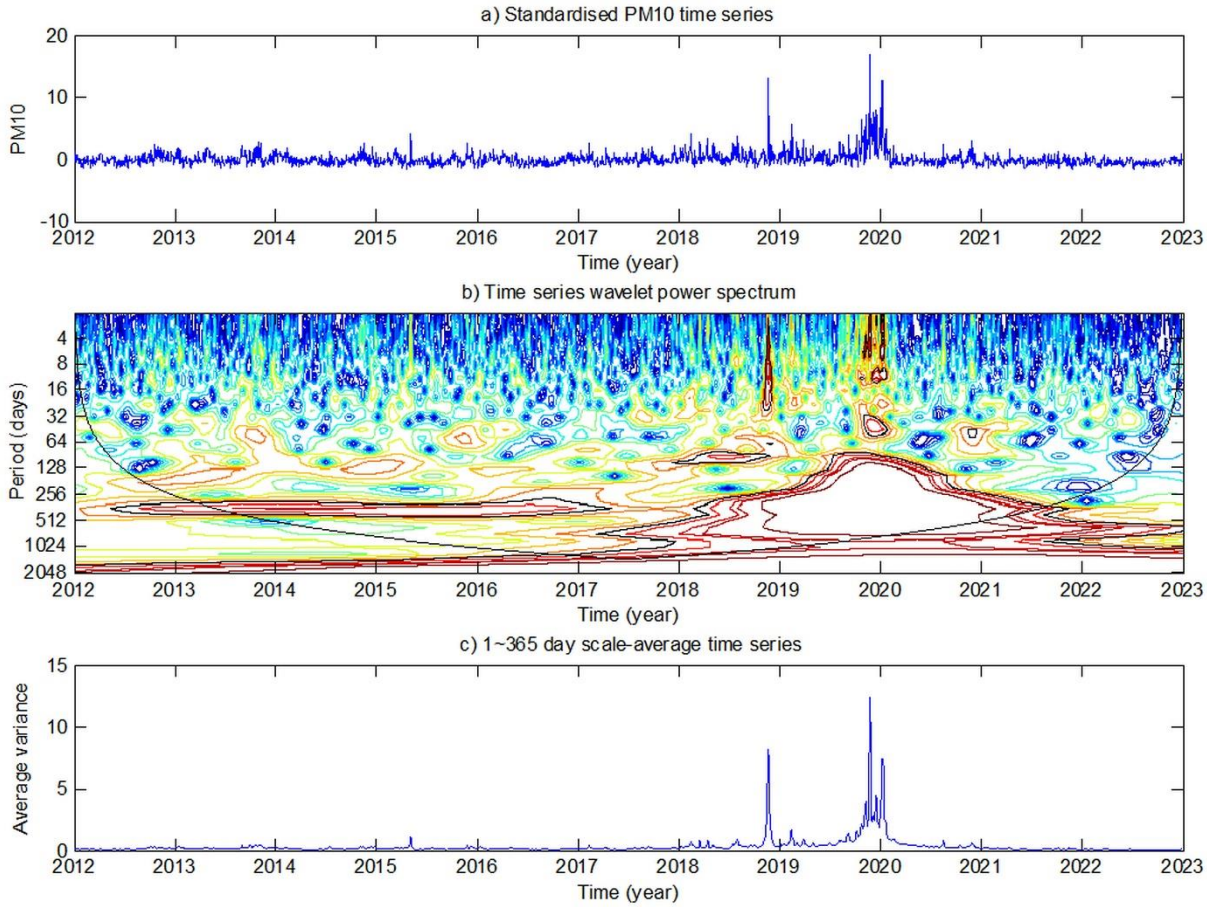

Figure S7. Muswelbrook PM10 temporal variability patterns derived from the all-day data (exceptional event day measurements included). (a) Standardised PM10 time series used for the wavelet analysis, where missing data were replaced with the dataset median for the station. (b) The local normalised wavelet power spectrum of (a) using the Morlet wavelet. The contour lines are at normalised variances of low to high values shown in dark (blue) to bright (light) colours. The thick black contour encloses regions of greater than 95% confidence for a red-noise process with lag-1 coefficient. Regions under the bowl-shape curve on either end indicates the "cone of influence", where edge effects become important. (c) The scale-averaged wavelet power (variance) over the 1-365 days band for the standardised PM10 time series.

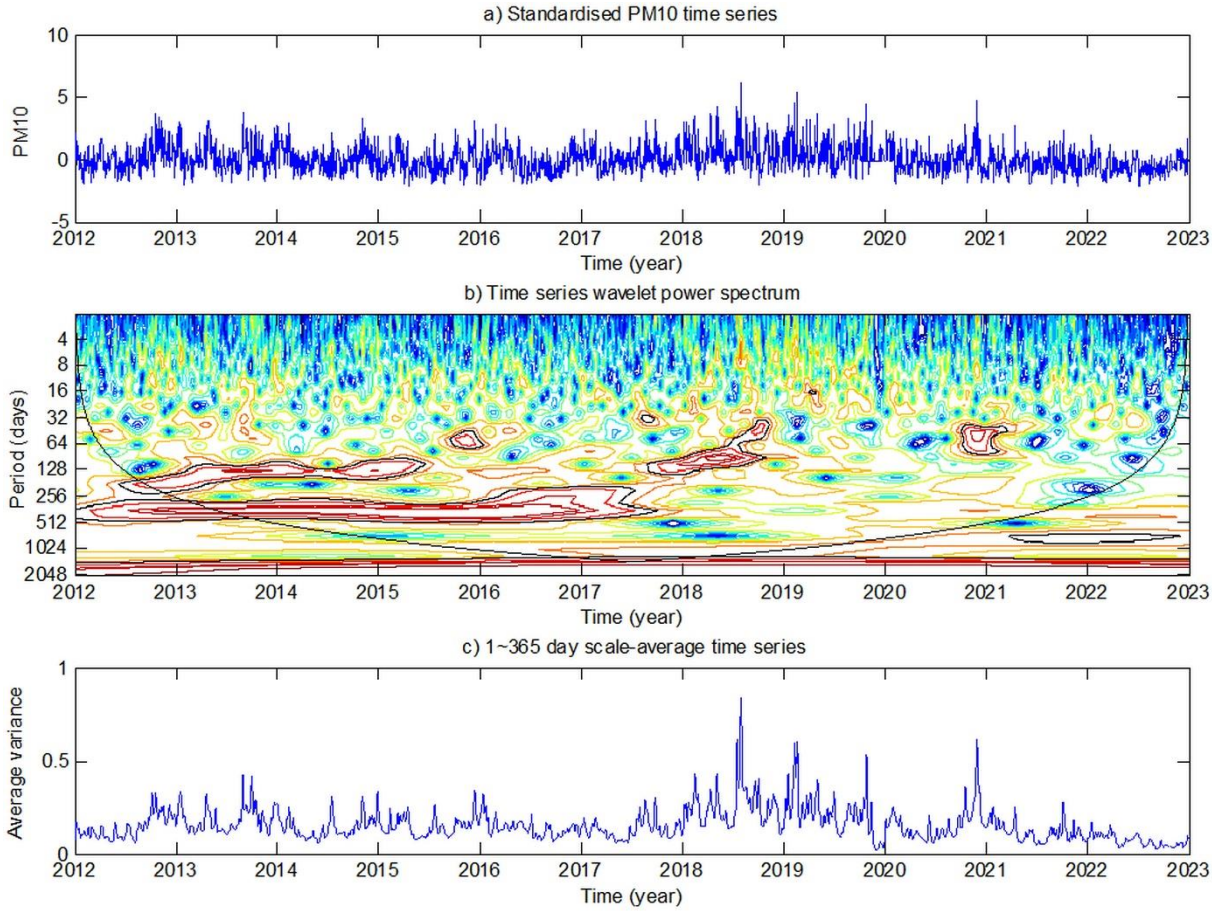

Figure S8. Muswellbrook PM10 temporal variability patterns derived from the normal-day data (exceptional event day measurements excluded). (a) Standardised PM10 time series used for the wavelet analysis, where missing data and measurements for exceptional event days were replaced with the dataset median for the station. (b) The local normalised wavelet power spectrum of (a) using the Morlet wavelet. The contour lines are at normalised variances of low to high values shown in dark (blue) to bright (light) colours. The thick black contour encloses regions of greater than 95% confidence for a red-noise process with lag-1 coefficient. Regions under the bowl-shape curve on either end indicates the "cone of influence", where edge effects become important. (c) The scale-averaged wavelet power (variance) over the 1-365 days band for the standardised PM10 time series.

a. WNW subregion

		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	All months
Mean PM10 levels (µg/m^3)	Merriwa	24.6	19.8	16.4	14.9	11.4	8.0	8.8	10.5	13.4	16.7	22.4	23.6	15.8
	Wybong	25.6	21.1	17.9	16.5	12.4	8.7	9.4	11.5	14.2	18.6	24.1	25.6	17.0
	Bulga	23.1	19.0	17.2	16.2	14.2	10.1	11.9	14.5	16.8	20.8	24.8	24.8	17.8
	Aberdeen	23.2	20.5	18.1	17.8	15.3	11.3	12.1	14.2	16.2	18.9	23.5	22.4	17.8
	Jerrys Plains	25.3	21.6	18.9	17.0	13.5	9.0	10.2	12.5	16.9	21.4	26.5	28.1	18.3
	Muswellbrook NW	25.4	22.0	19.4	19.3	17.0	12.8	14.1	16.5	18.4	22.1	25.9	25.3	19.9
	Muswellbrook	26.0	22.6	20.4	21.4	21.1	17.7	19.0	20.5	20.9	23.5	27.7	26.3	22.3
	Warkworth	28.7	24.0	21.8	19.9	18.0	13.9	15.2	18.0	22.0	26.4	29.5	30.4	22.3
Poor air quality days (count)	Merriwa	13	2	3	1	1	0	1	0	1	7	16	25	70
	Aberdeen	10	5	4	1	1	0	0	3	1	8	19	19	71
	Bulga	12	4	3	2	1	0	1	0	1	9	21	21	75
	Wybong	18	7	3	1	1	0	1	1	2	7	18	21	80
	Muswellbrook_NW	13	6	4	2	1	0	1	4	1	10	23	23	88
	Muswellbrook	15	6	3	2	3	0	2	5	2	10	23	24	95
	Jerrys_Plains	18	6	4	2	1	0	0	2	3	9	27	24	96
	Warkworth	19	10	7	2	1	0	0	3	3	15	25	27	112

b. SE subregion

		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	All months
Mean PM10 levels (µg/m3)	Singleton S	21.3	17.5	16.9	17.7	18.9	14.5	17.0	20.3	20.3	22.8	24.7	23.0	19.6
	Singleton	22.0	18.5	17.7	19.5	22.2	17.3	20.2	22.9	22.2	23.9	24.8	22.4	21.1
	Singleton NW	23.4	19.4	19.4	21.5	25.2	18.3	21.9	25.9	25.7	26.9	27.9	25.0	23.4
	Maison Dieu	26.5	21.0	20.4	21.1	22.1	15.8	19.4	24.0	25.3	28.3	29.9	27.9	23.5
	Mt Thorley	22.4	19.2	19.3	21.6	25.4	17.8	21.8	27.0	27.1	28.2	28.1	25.3	23.6
	Camberwell	25.4	20.9	20.7	22.7	26.5	20.0	25.4	29.6	29.0	30.4	31.8	27.6	25.9
Poor air quality days (count)	Singleton_S	7	4	4	3	2	0	2	3	4	13	18	17	77
	Singleton	12	1	3	3	1	0	2	3	10	14	23	17	89
	Maison_Dieu	19	9	5	5	3	0	6	11	21	40	33	25	177
	Singleton_NW	15	7	6	9	12	0	5	16	23	36	30	23	182
	Mt_Thorley	10	9	5	11	16	1	13	25	31	38	29	23	211
	Camberwell	23	12	9	12	12	2	20	32	40	47	41	29	279

Figure S9. Monthly mean PM10 levels and total number of poor air quality days (with PM10 levels > 50 µg/m^3) for stations in the (a) WNW and (b) SE subregions. Rows are sorted by the "All months" column (multi-year station means or count of poor air quality days). Data: daily PM10 measurements in 2012–2022 (including exceptional events). Colour scale: green—relatively low value; yellow—near medium value; red—relatively high value.

(a) WNW subregion		2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	All years
Mean PM10 levels (µg/m3)	Merriwa	14.2	14.9	15.2	13.2	13.5	14.2	19.2	27.9	18.2	11.7	11.2	15.8
	Wybong	15.4	15.5	16.9	14.8	15.3	16.6	21.6	28.5	18.2	12.6	11.7	17.0
	Bulga	18.7	19.2	17.7	15.0	16.1	17.2	21.3	28.6	18.1	12.9	10.6	17.8
	Aberdeen	17.0	17.3	17.9	15.2	15.6	17.6	22.3	29.5	17.8	12.9	12.3	17.8
	Jerrys Plains	10.8	18.6	18.2	15.5	16.8	18.0	24.3	32.1	20.5	13.6	13.3	18.3
	Muswellbrook NW	19.1	18.9	19.2	16.7	16.6	18.5	25.0	33.7	21.0	15.6	14.3	19.9
	Muswellbrook	21.8	22.6	21.4	19.1	19.2	21.7	27.2	34.4	22.5	18.2	16.6	22.3
	Warkworth	21.1	21.4	20.6	18.2	18.6	21.8	26.4	33.4	23.7	20.8	19.3	22.3
Poor air quality days (count)	Merriwa	1	0	3	1	0	0	6	47	12	0	0	70
	Aberdeen	0	0	2	1	0	2	7	51	8	0	0	71
	Bulga	2	7	3	2	0	0	8	45	8	0	0	75
	Wybong	1	2	3	1	1	3	9	47	13	0	0	80
	Muswellbrook_NW	1	1	1	2	0	1	10	57	14	0	1	88
	Muswellbrook	1	3	1	2	0	2	13	58	15	0	0	95
	Jerrys_Plains	0	6	6	1	0	1	11	54	17	0	0	96
	Warkworth	0	8	3	3	0	1	16	59	17	4	1	112

(b) SE subregion		2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	All years
Mean PM10 levels (µg/m3)	Singleton S	19.0	20.2	18.3	16.9	18.0	19.4	23.0	30.7	19.8	16.5	14.0	19.6
	Singleton	22.3	23.3	21.0	19.3	19.3	20.8	24.0	30.1	20.5	17.5	14.5	21.1
	Singleton NW	25.9	25.9	22.7	20.9	21.9	22.7	26.9	34.6	22.2	18.8	15.2	23.4
	Maison Dieu	25.8	25.8	22.7	20.4	20.4	23.1	27.9	38.0	22.3	17.6	14.0	23.5
	Mt Thorley	24.8	24.7	21.5	19.8	22.8	25.4	29.1	36.4	22.0	19.0	14.2	23.6
	Camberwell	26.4	27.8	24.6	22.0	24.5	27.4	31.1	39.9	24.3	20.6	16.0	25.9
Poor air quality days (count)	Singleton_S	2	5	0	2	0	2	9	44	11	2	0	77
	Singleton	6	12	1	3	1	5	10	40	10	1	0	89
	Maison_Dieu	20	28	6	5	0	9	25	66	17	1	0	177
	Singleton_NW	29	28	6	4	4	12	22	62	14	1	0	182
	Mt_Thorley	28	26	3	7	5	21	34	69	13	5	0	211
	Camberwell	23	36	12	11	11	33	44	87	18	4	0	279

Figure S10. Annual mean PM10 levels and total number of poor air quality days (with PM_{10} levels > 50 µg/m^3) for stations in the (a) WNW and (b) SE subregions. Rows are sorted by the "All years" column (multi-year station means or total number of poor air quality days). Data: daily PM10 measurements in 2012–2022 (including exceptional event days). Colour scale: green—relatively low value; yellow—near medium value; red—relatively high value.