Article

Visualising Daily PM10 Pollution Data in an Open-Cut Mining Valley of New South Wales, Australia—Part I: Identification of Spatial and Temporal Variation Patterns

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

Article

Visualising Daily PM10 Pollution Data in an Open-Cut Mining Valley of New South Wales, Australia—Part I: Identification of Spatial and Temporal Variation Patterns

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

New South Wales Department of Climate Change, Energy, the Environment and Water, Sydney, Australia
* Correspondence: ningbo.jiang@environment.nsw.gov.au

Abstract: The Upper Hunter Valley is a major coal mining area containing approximately 40% of the currently identified total coal reserves in New South Wales (NSW), Australia. Due to the ongoing increase of mining activities, PM10 (air-borne particle with an aerodynamic diameter less than 10 micrometres) pollution has become a major air quality concern in local communities. This paper summarises the spatial and temporal variability modes of PM10 pollution in the region, based on long-term multi-site monitoring data and the application of the rotated principal component analysis (RPCA) and wavelet analysis techniques. RPCA identified two distinct air quality clusters/subregions in the valley, one in the west/northwest, the other in the southeast. Wavelet analysis revealed the annual cycle to be most persistent temporal mode of PM10 variability in both subregions, with intermittent signals observed at time scales of around 120, 30~90 and under 30 days. How these variation modes are related to the effects of local PM10 emissions and the influence of meteorology at different time scales deserves further attention in future work. The findings will be used in air quality reporting and forecasting in NSW. The methodology and results can also be useful for air quality research in similar regions elsewhere.

Keywords: PM10 pollution; spatial-temporal variability mode; rotated principal component analysis (RPCA); wavelet analysis; open-cut mining; valley environment

1. Intorduction

The Upper Hunter Valley is located within the northern end of the Hunter region, around 200 km north of Sydney and 50 km northwest of Newcastle in the State of New South Wales (NSW), Australia (Figure 1; Olguin and Everingham, 2015). On a broad scale the valley is oriented northwest-southeast (NW-SE), approximately 30 km wide and with the terrain elevation of around 300~380 m from the lower (Singleton South) end to the upper (Merriwa) end (Connor et al., 2008; OEH, 2017a). It is a major coal mining area in NSW, with significant coal mines located between Aberdeen in the north to Bulga in the south, containing approximately 40 percent of the currently identified total coal reserves in the state (ABARES, 2023). It also has a significant agriculture industry (dairy, beef, horse breeding, and viticulture) and two large coal-fired electricity generation plants (DPI, 2018; Olguin and Everingham, 2015), with the Liddell Power Station fully decommissioned during April 2022 to April 2023. There are multiple medium-to-small towns in the region. The largest population centres are Singleton (population ~ 24,577) and Muswellbrook (population ~ 16,357) (ABS, 2023), with many small settlements and isolated rural residences scattered throughout the valley.
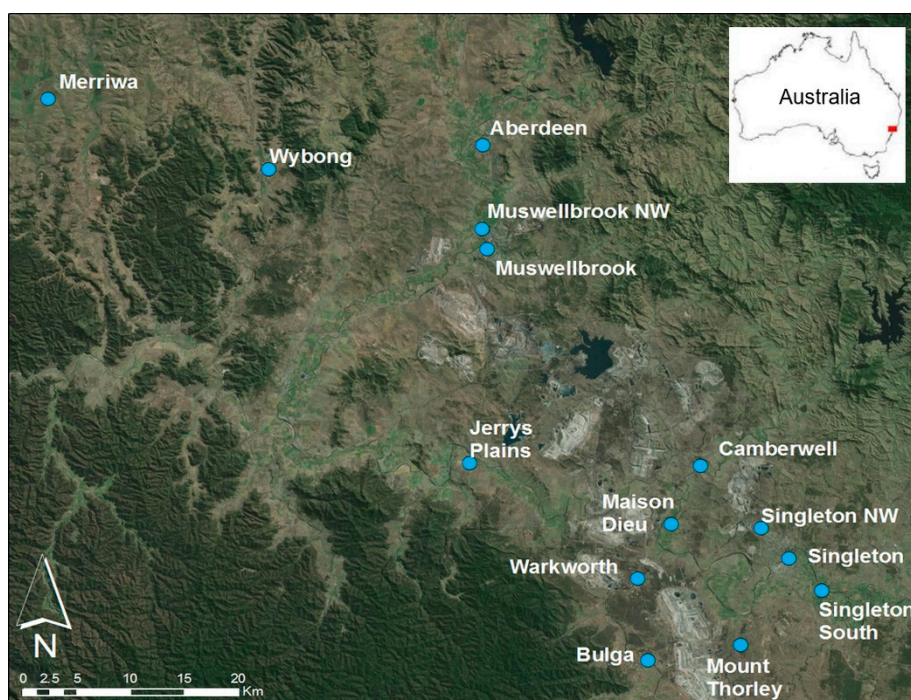
Figure 1. Upper Hunter Valley - locations of air quality monitoring stations. Source of base map source: Google.

The prevailing surface winds tend to follow the NW-SE orientation of the valley. The most frequent winds are north-westerlies in winter and south-easterlies in summer, with wind directions less defined in autumn and spring (DPE, 2022; OEH, 2017a; Holmes, 2008; Bridgman and Chambers, 1981). Nocturnal and early morning down-valley drainage flows, daytime up-valley winds or south-easterly sea breezes are also observed from time to time (OEH, 2017a; Hibberd et al., 2013; Holmes, 2008; Physick et al., 1991; Hyde et al., 1981). Wind strengths vary across the valley with most locations experiencing annual average wind speeds in the range of around 2~4 m/s. The precipitation in the region is generally low (compared to coastal regions to the east), and varies significantly across years, for example with the annual rainfall ranging from around 549 mm to 853 mm at Singleton and Muswellbrook during 2011–2015 (OEH, 2017). Higher rainfall tends to occur in summer and early autumn, and lower rainfall in winter and early spring.

Particle pollution is known to have adverse effects on human health and the environment (e.g., Hertzog et al., 2024; Huang et al., 2018; Segersson et al., 2017; Bartnet et al., 2006; Keywood et al., 2016; CSIRO, 2013; EPHC, 2010; Anderson et al., 2012). PM10 (air-borne particles with an aerodynamic diameter less than 10 micrometres) pollution has been a major air quality concern for local communities in the Upper Hunter Valley (NSWEPA, 2023; POEO Regulation, 2021; Keywood et al., 2020; NSWEPA, 2019; Morrison and Nelson, 2011; Conner et al., 2008). The main sources of PM10 emissions in the region include local open-cut coal mining activities (e.g., wheel generated dust, windblown dust from overburden) and surface soil erosions (windblown dust). For example, coal mining contributes around 88% of PM10 emissions in the combined Muswellbrook, Singleton and Upper Hunter local government areas (NSWEPA, 2019). Coal-fired electricity generation, agriculture, bushfires, prescribed hazard reduction burnings and state-wide dust storms also contribute to PM10 pollution in the region (DPE, 2022).

There have been a few studies examining PM10 pollution in the Upper Hunter Valley. Of these, most early investigations were based on data collected (with different types of instruments) at limited (small) number of locations through short-term campaign monitoring projects. These include those reported in SPCC (1982, 1983), Holmes and Associates (1996), Morrison and Nelson (2011), Hibberd et al. (2013) and relevant references therein. For example, SPCC (1982) reported a few observational and modelling studies, which concluded that localised dust pollution from open-cut coal mines and related developments continued to be issue of concern, and that there was unlikely to be serious

cumulative, region-wide problems resulting from dust emissions from mines. Holmes and Associates (1996) found that the increase in both deposition and concentration levels of dust over 1984-1994 were due to the increase in coal production and the severe drought affecting much of eastern Australia, and that the land affected by cumulative effects appeared to be primarily that owned by the coal mines. Holmes (2008) and Hyde et al. (1981) also suggested that local PM10 emissions in the valley can impact air quality in areas away from sources - that is, north-westerlies can transport dust generated in the upper end of the valley to areas near the bottom of the valley or further down over the metropolitan areas of Newcastle.

In partnership with the Upper Hunter coal and power industries, the NSW Government commissioned the Upper Hunter Air Quality Monitoring Network (UHAQMN) during 2010-2012 (Figure 1), to provide the community, industry and government with reliable and up-to-date information on air quality within the valley (OEH, 2017b). Pollution (including PM10) data from the network are reported as air quality categories (AQCs) in near real-time and in quarterly or annual data reports on the NSW government website (https://www.airquality.nsw.gov.au). Multi-year PM10 data summaries were available in two main reports, i.e., OEH (2017b) for data in 2011-2015 and DPE (2022) for data in 2011-2021. Three main findings are worth highlighting: 1) the annual PM10 concentrations in the UHAQMN were observed to be amongst the highest across the NSW Air Quality Monitoring Network (NSW AQMN, over 90 stations across NSW; Riley et al., 2020); 2) PM10 levels at some sites can exceed the national benchmark (i.e., Australian standard for PM10; NEPC, 2021) from time to time, and vary significantly from year to year due to impacts of draught conditions and occurrence of hazard reduction burnings or bushfires; 3) emissions from open-cut mines can lead to elevated PM10 pollution typically in the lower (southeast) end of the valley, in particular under the north-westerly winds.

OEH (2017a) performed a detailed analysis of the 2012-2015 (4-year) PM10 data from the UHAQMN in an air quality management campaign project. The results showed that: 1) poor air quality days generally occurred in spring and secondarily summer and autumn, with winter and February/March having relatively good air quality; 2) there were more poor air quality days in 2012–2013 but significantly reduced number of events in 2014–2015; and 3) the correlations between high PM10 pollution and individual meteorological variables were complex and non-linear, varying with time and location. The author also proposed the existence of two air quality clusters, i.e., the south-east (SE) and west-north-west (WNW) air quality subregions in the valley. Drawing up the project findings, the author indicated that the application of more sophisticated (holistic) analytic methods such as pattern recognition techniques may provide increased understanding of PM10 pollution in the study region.

Globally there are relatively few studies on air quality issues in rural valley environments, with much of the air quality literature primarily focusing on urban air pollution problems. Of those few studies, most work was undertaken with data for a small number of sites and based on correlation analysis, e.g., on spatial and/or temporal variations of PM10 by Mohd Shafie et al. (2022) and NPS (2023), and on correlations between PM10 levels and selected atmospheric parameters by Mannis (1988), Giri et al. (2008), Fortelli et al. (2016), Reisen et al. (2017), Czernecki et al. (2017) and Quimbayo-Duarte (2021). In summary, elevated PM10 pollution in those valleys are associated with (prolonged) dry conditions (low rainfall and humidity), low winds, thermal inversions (low mixing heights), and/or under the influence of high-pressure systems. Till now, to the best of our knowledge there is little or no research in the literature on the topic of examining spatial and temporal variability modes of PM10 pollution.

This project extends OEH (2017a), based on the 11-year (2012-2022) PM10 data from 14 stations in the UHAQMN and by applying advanced analytic methods, to holistically examine: 1) the spatial-temporal variation patterns of PM10 pollution in the Upper Hunter Valley; and 2) how elevated pollution days are related to local- and synoptic-scale meteorological configurations on the region. This paper is focused on presenting the investigation results of 1), i.e., on the spatial and temporal variation modes from the long-term multi-site PM10 data. The investigation is unique in at least three aspects: a) the air quality subregions initially proposed in OEH (2017a) are verified and analysed

based on a longer dataset using the rotated principal component analysis (RPCA); b) the temporal modes of PM10 pollution were identified for the air quality subregions through wavelet analysis and were illustrated via heat map visualisations; and c) the impact of exceptional events including dust storms and bushfires on PM10 pollution were also examined in some depth for the subregions. The results on the linkage between PM10 pollution and local and synoptic meteorological features will be reported separately in a companion paper of this text. The findings will used for air quality forecasting in NSW, and the methods and results can also be useful for air quality research in similar regions elsewhere.

2. Data

2.1. Air Quality Data

The 14 monitoring stations in the UHAQMN (Figure 1) can be grouped in station types that serve different purposes, as is described in Table 1 (OEH, 2012). Air quality data for these stations were obtained from the NSW Air Quality Data System (AQDS). These included: 1) daily (24-hr) average PM10 concentrations (μg/m^3) for each monitoring station; and 2) information on whether and when PM10 measurements at any of these stations were significantly impacted by exceptional events (further details in Section 2.4) such as hazard reduction burnings (HRBs), bushfires or widespread dust storms. The data were confirmed of high quality with missing values generally less than 2% at individual stations, except that there were up to 3.9% missing data at the Merriwa and Mt Thorley stations.

Table 1. Air quality station details and PM10 data used in this study. Source: adapted from OEH (2012).

Station type	Station purpose	Station name*	Total number of days with valid daily PM10 data in 2012-2022	Total number of days with invalid or missing daily PM10 data in 2012-2022
Larger population centre	Monitoring air quality in the larger population centres	Aberdeen	3984	34
		Muswellbrook	3968	50
		Singleton	3978	40
Smaller population centre	Monitoring air quality in the smaller communities	Bulga	3967	51
		Camberwell	3970	48
		Jerrys Plains	3950	68
		Maison Dieu	3953	65
		Warkworth	3940	78
		Wybong	3962	56
Diagnostic	Providing diagnostic data that helps to diagnose the likely sources and movement of particles across the region as a whole; they do not provide information about air quality at population centres	Mt Thorley	3876	142
		Muswellbrook NW	3986	32
		Singleton NW	3978	40
Background	Providing background data at the top (Merriwa) and bottom (Singleton South) ends of the valley	Merriwa	3862	156
		Singleton S	3951	67

*NW: Northwest; S: South; there is a total of 130 exceptional event days (Section 2.2).

2.2. Definition of Exceptional Event Day, Normal Day, and Poor Air Quality Day

Of the up to 3986 days of valid daily measurements in 2012–2022, there were a total of 130 days when air quality in the Upper Hunter Valley was significantly impacted by air emissions from bushfire, planned hazard reduction burning (HRB) and/or continental-scale dust storm, with daily PM_{10} levels above the 24-hr average national benchmark (NEPC, 2021) level of 50 µg/m^3 at one or more stations in the UHAQMN. These 130 days are referred to as *exceptional event days* in this text. All other days, which were not significantly impacted by bushfire, HRB or continental-scale dust storm, are referred to as *normal days*, or interchangeably *non-exceptional event days*. Most of the exceptional event days occurred during the 2019-2020 spring-summer bushfires and secondarily the widespread dust storm events in 2018, both associated with impacts of prior prolonged droughts across Australia (DPE, 2019, 2020, 2022).

These definitions led to two derived daily PM_{10} datasets for 14 monitoring stations, i.e., *normal-day dataset* (excluding data for exceptional event days listwise across all stations) and *exceptional event day dataset*. The normal-day dataset reflects the air quality conditions mainly associated with PM10 emissions within the valley, primarily from open-cut mining activities and secondarily soil erosion, power generation and agriculture activities. The exceptional event day dataset is related to the significant air quality impacts due to emissions from (local or remote) exceptional events in the study region. For easy reference, the original PM10 dataset is referred to as *all-day* or *full* dataset (i.e., including data of both above datasets), depicting the full air quality conditions associated with effects from both local and broad-scale PM10 emissions.

To facilitate our analysis, a day was marked as poor air quality (pollution) day if PM10 concentration on the day was above the national benchmark level. This definition led to counts of poor air quality days for each monitoring station. The terms "PM10 pollution" and "air pollution" are used in an exchangeable manner in this text.

3. Methods

The investigation was conducted in four steps to examine the spatial and temporal variability modes of daily PM10 pollution. The first step was to describe the general properties of PM10 pollution in the Upper Hunter Valley, based on boxplots and general summary statistics for the long-term (2012-2022) PM10 measurements from 14 stations in the UHAQMN (Section 2.1). This sets out the context for the subsequent analyses and interpretation of results from other steps. The second step was to identify the spatial co-variation structure (i.e., spatial regionalisation) of PM10 pollution by applying the RPCA (rotated principal component analysis) to the multi-site PM10 data in 2012-20122 (Section 3.1). This analysis verifies the air quality clustering (subregions) previously identified in OEH (2017a) from a shorter (2012-2015) dataset. In the third step, wavelet analysis was applied to the principal component (PC) time series obtained from RPCA to determine the dominant temporal modes of PM10 pollution in each subregion, and how those modes changed over time (Section 3.2). Then in the fourth step, heat maps were used to illustrate the variation patterns of average PM10 levels and number of exceedance days in each subregion on the annual and interannual scales.

Throughout steps 1 to 4, the effects of exceptional events (vegetation fires and dust storms) on the findings were examined by repeating the relevant analyses for the all-day dataset and where appropriate the dataset for exceptional event days only. Due to the space limit, this text is focused on findings from the normal-day dataset (exceptional event days excluded), with some results from the all-day dataset given as Supplementary Materials to support our interpretations of findings. The implementation of RPCA and wavelet analysis are described next.

3.1. Rotated Principal Component Analysis (RPCA)

Principal component analysis (PCA) can be used for data reduction, variation mode identification, or feature classification (Jiang, 2010). In mode identification or feature classification, a rotation technique is often applied to PCs for easy interpretation of results (Richman, 1986). Here PCA was applied to the correlation matrix of the daily PM_{10} concentration time series (2012–2022) for

14 air quality stations to examine the structure of inter-site covariations, i.e., spatial regionalisation of PM10 pollution in the valley. Missing data points were not used, i.e., ignored listwise across stations, when calculating the correlation matrix to suppress the potential for introducing new noise due to the data imputation process (note: PCA was applied differently in Section 3.2 for the purpose of wavelet analysis). The retention of the number of principal components (PCs) was decided through a scree-test, following the combined use of the Cattell (1966) and North et al. (1982) methods (Jiang, 2011). A plot of initial eigenvalue and sampling error by PC-number was used for this purpose, keeping in mind that *degenerated multiplets* (i.e., unrotated PCs) should not be separated from one another and as much variance in the dataset as possible should be explained by the PCs retained. Varimax (orthogonal) rotation was then applied to the retained PCs to facilitate physically meaningful interpretations, while preserving the linear independency between PCs (Jiang et al., 2005).

In addition, an obliquely rotated PCA (without the orthogonality constraint for PC rotation; Jiang, 2011) was performed on the same data to confirm if the linearly independent PCs from the Varimax rotation could approximate the true (underlying) *simple structure* in the data (Harman, 1976). The RPCA was repeated for the daily PM10 data with or without exceptional event days included, i.e., all-day dataset vs normal-day dataset, to test the stability of the RPCA findings.

3.2. Wavelet Analysis

Wavelet analysis can generate a representation of a signal (stationary or non-stationary) simultaneously in the time and frequency domains, thereby allowing access to localised information about the signal (Adamowski and Chan, 2011). Wavelet analysis was performed to identify the dominant temporal variability modes of PM10 pollution if existing in the study region. Wavelet analysis requires that no missing data exist in the time series. Hence, an RPCA was performed on a processed daily PM10 dataset where the missing records (if any) at each station were filled with the overall median of the PM10 measurements in the 2012-2022 period for that station, resulting in PC time series (scores) without missing values but otherwise identical to those from the RPCAs described in Section 3.1. A wavelet analysis was then applied to these PC scores to determine the dominant modes of co-variability in PM10 pollution for each subregion, and how those modes changed over time.

We followed the procedure outlined in Torrence and Compo (1998), by using the Morlet wavelet with the following parameter configurations: sampling rate/resolution = 1 (day); frequency resolution = 0.25, i.e., four suboctaves (voices) per octave; lower and upper Fourier periods (scales) for wavelet decomposition set to 2 days) and 2048 days (~5.6 years), respectively. The lower and upper Fourier periods were chosen this way so that variability modes on the sub-weekly, annual (seasonality) and multi-year (5-6 years) scales, and anything in between will be considered. We assumed a red-noise process with the lag-1 serial correlation for each PC time series when testing the significance of wavelet spectrum power at the 0.05 level for a chi-square test.

The above analysis process was repeated on the PCs derived from the RPCA on the normal-day dataset, where both the missing data and the readings for exceptional event days were replaced with the median of PM10 measurements for normal days by individual stations. This was to assess the (potential) impact from the inclusion of high particle measurements in exceptional event days on the wavelet analysis results.

4. Results and Discussions

4.1. General Description of Daily PM10 Pollution and Impacts of Exceptional Events

Box plots and summary statistics for the normal-day dataset (i.e., exceptional event days excluded) are given in Figures 2 and 3 to illustrate the distributional properties of the long-term (2012-2022) PM10 concentration data from 14 stations in the UHAMQN. In general, relatively higher daily PM10 levels were recorded at four (direct) source-impacted locations, i.e., the Camberwell, Maison Dieu, Mt Thorley and Singleton NW stations, which are relatively close to open-cut mining sites in

the southeast (lower end) of the valley (Figure 2). These stations had larger variabilities (indicated by larger inter-quartile ranges, i.e., box lengths) and more outlier/extreme values (indicated by dots and asterisks). PM10 levels at the Warkworth station were comparably high. In contrast, Merriwa (background station near the top of the valley) and secondarily Wybong, Jerrys Plains, Bulga and Aberdeen recorded relatively lower and less variable PM10 pollution, as indicated by the lower box positioning and relatively fewer outlier/extreme values. PM10 levels at the two largest population centres, Singleton and Muswellbrook, appeared in between the above two cases, with extreme values also observed at these locations. Minimum daily PM10 levels were generally similar across 14 stations, ranging 1.5 ~ 3.5 µg/m³ (Figure 3). Overall, these results are generally consistent with the distributional properties reported by OEH (2017a) for a shorter (normal days in 2012-2015) PM10 dataset. Holmes and Associates (1996) noted that the land affected by cumulative effects appeared to be primarily that owned by the coal mines. Essentially, as is for other regions (e.g., Jiang et al., 2014; Jiang et al., 2017), the significant inter-site variability in PM10 pollution reflects the combined effects of changes in local emissions (primarily from open-cut mining and soil erosion) and environmental factors such as meteorological conditions on different time scales.
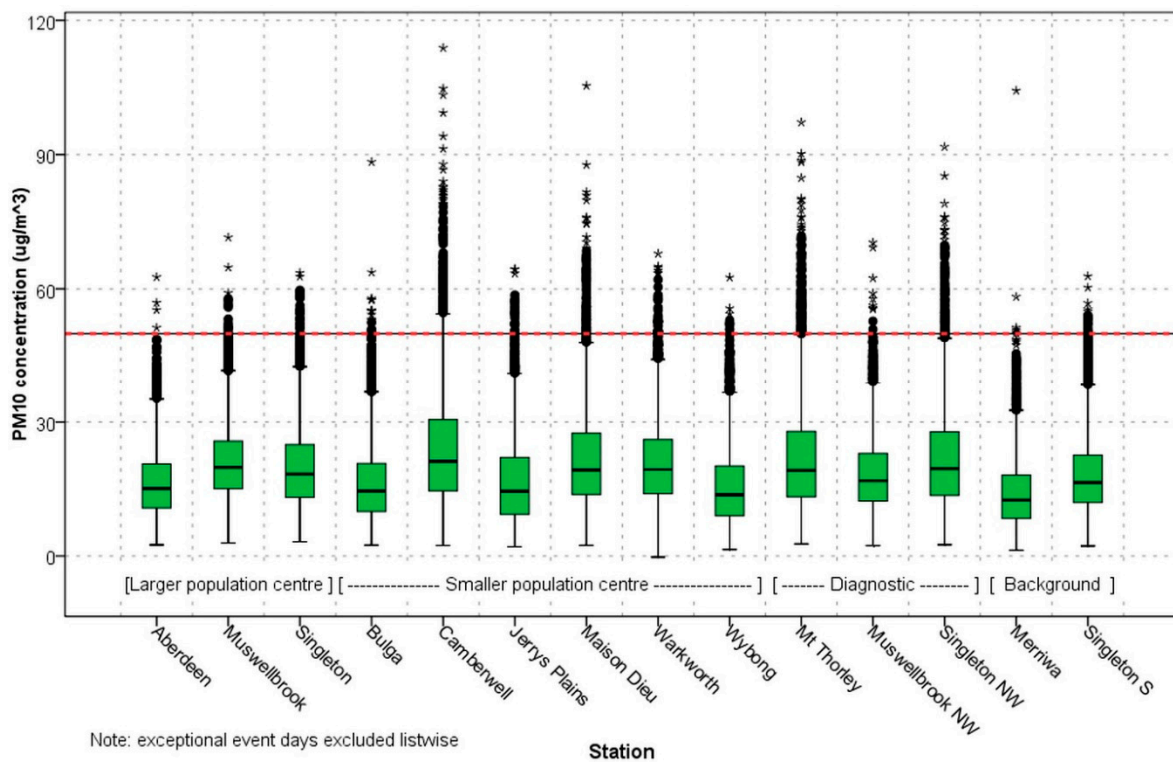
Figure 2. Box plots by site for daily PM10 measurements in 2012–2022 (excluding exceptional event days). The lower and upper boundaries of the box are respectively the 25th and 75th percentile; the horizontal line inside the box represents the median; asterisks represent extreme values, cases with values 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: the Australian national standard of 50 µg/m³ for daily PM10.

Station type	Station name	Normal-day dataset (excluding exceptional events)						All-day dataset (including exceptional events)					
		Number of days	Mean	Median	Minimum	Maximum	Std. Deviation	Number of days	Mean	Median	Minimum	Maximum	Std. Deviation
Larger population centre	Aberdeen	3854	16.4	15.1	2.5	62.6	7.6	3984	17.8	15.4	2.5	267.7	12.4
	Muswellbrook	3839	20.9	19.9	2.9	71.5	8.3	3968	22.3	20.1	2.9	231.3	12.4
	Singleton	3850	19.9	18.3	3.2	63.6	8.9	3978	21.1	18.7	3.2	206.1	12.3
Smaller population centre	Bulga	3843	16.4	14.6	2.4	88.4	8.7	3967	17.8	14.8	2.4	196.6	12.9
	Camberwell	3841	24.3	21.2	2.4	113.8	13.3	3970	25.9	21.8	2.4	294.4	17.5
	Jerrys Plains	3825	16.8	14.5	2.1	64.4	9.6	3950	18.3	14.9	2.1	226.7	14.4
	Maison Dieu	3823	21.8	19.3	2.4	105.4	11.2	3953	23.5	19.6	2.4	446.1	16.8
	Warkworth	3813	20.9	19.4	0.0	67.9	9.5	3940	22.3	19.7	0.0	181.5	13.4
	Wybong	3834	15.5	13.7	1.5	62.6	8.6	3962	17.0	14.0	1.5	373.6	14.2
Diagnostic site	Mt Thorley	3750	22.1	19.2	2.7	97.2	12.2	3876	23.6	19.6	2.7	231.4	15.9
	Muswellbrook NW	3859	18.4	16.8	2.3	70.4	8.6	3986	19.9	17.1	2.3	244.6	13.3
	Singleton NW	3850	22.0	19.6	2.5	91.7	11.5	3978	23.4	20.1	2.5	237.4	14.7
Background site	Merriwa	3735	14.1	12.5	1.3	104.3	7.8	3862	15.8	12.8	1.3	620.7	16.5
	Singleton S	3822	18.3	16.4	2.2	62.8	8.5	3951	19.6	16.8	2.2	212.6	12.4

Figure 3. Summary statistics by station for the normal-day dataset (exceptional days excluded listwise) compared with the all-day dataset (exceptional days included). Values are colour coded for each dataset separately: green – relatively low value; yellow – near medium value; red – relatively high value. Data: daily PM10 concentrations ($\mu g/m^3$) for stations in the UHAQMN.

An analysis was also performed for the all-day (full) dataset (including exceptional event days), as illustrated in Figures 3 and Figure S1 in Supplementary Material. The distributional properties of PM10 data were found very similar to those described above for the normal-day dataset. The exception is that the full dataset showed slightly higher means, medians and standard deviations, significantly (around two or more times) higher maximum values (Figure 3), and many more days with outlier or extreme values (Figure S1). These differences between datasets were expected, essentially reflecting the significant impact of exceptional events on regional air quality, such as the widespread dust storms in 2018 and the spring-summer bushfires across large areas of NSW in 2019-2020 (DPE, 2019, 2020; Watt et al., 2019), as is further illustrated on the interannual scale in Sections 4.3 and 4.4.

We also compared the summary statistics between the normal-day and exceptional event-day datasets (Figure 4). The mean and median of PM10 concentrations for exceptional event days were above the national benchmark (50 $\mu g/m^3$) at all stations, equivalent to around 2~3 times increase over those for normal days. The minimum PM10 pollution for exceptional event days reached around half of the national benchmark level, or around 5~10 times increase over the that for normal days. The maximum and standard deviation values for exceptional event days were also around 1~4 and 2~7 times higher, respectively, compared to normal days. Overall, the four source-impacted stations were most significantly affected, but with Merriwa recording the highest maximum PM10 level due to the impact of a wide-spread dust storm on 11 January 2020 (DPE, 2021). This finding indicates the accumulated (combined) impacts on air quality from local (primarily associated with open-cut mining and soil erosion), remote (mainly continental dust storms) and incidental (HRBs or bushfires) particle emissions.

Station type	Station name	Number of days	Exceptional event day dataset					Proportional increase over normal-day dataset				
			Mean	Median	Minimum	Maximum	Std. Deviation	Mean	Median	Minimum	Maximum	Std. Deviation
Larger population centre	Aberdeen	130	57.7	50.4	17.2	267.7	36.3	2.5	2.3	5.9	3.3	3.7
	Muswellbrook	129	62.6	55.6	18.1	231.3	31.8	2.0	1.8	5.2	2.2	2.8
	Singleton	128	58.3	50.6	19.7	206.1	29.7	1.9	1.8	5.2	2.2	2.3
Smaller population centre	Bulga	124	59.3	49.4	17.2	196.6	34.9	2.6	2.4	6.1	1.2	3.0
	Camberwell	129	73.8	65.3	18.5	294.4	42.1	2.0	2.1	6.8	1.6	2.2
	Jerrys Plains	125	66.0	55.0	17.6	226.7	37.6	2.9	2.8	7.4	2.5	2.9
	Maison Dieu	130	73.9	58.9	23.3	446.1	48.0	2.4	2.1	8.7	3.2	3.3
	Warkworth	127	64.9	55.4	23.4	181.5	31.2	2.1	1.9	0.0	1.7	2.3
	Wybong	128	61.4	51.0	13.4	373.6	45.3	3.0	2.7	8.2	5.0	4.3
Diagnostic site	Mt Thorley	126	67.7	60.3	16.4	231.4	36.7	2.1	2.1	5.1	1.4	2.0
	Muswellbrook NW	127	63.8	54.4	18.7	244.6	36.6	2.5	2.2	7.1	2.5	3.3
	Singleton NW	128	64.6	58.8	19.3	237.4	32.3	1.9	2.0	6.6	1.6	1.8
Background site	Merriwa	127	64.9	50.9	11.7	620.7	63.9	3.6	3.1	8.1	5.0	7.2
	Singleton S	129	59.2	49.8	20.2	212.6	30.9	2.2	2.0	8.0	2.4	2.6

Figure 4. Summary statistics by station for exceptional event days (totalling N=130 days) and proportional increases over the normal-day dataset given in Figure 3. Values are colour coded separately for the exceptional event dataset and the proportional increase ratios: green – relatively low value; yellow – near medium value; red – relatively high value. Data: daily PM10 measurements in 2012-2022.

4.2. Spatial Pattern - Identification of Two Air Quality Subregions

The RPCA for the normal-day dataset (Section 3.1) led to the retention of two leading principal components (PC1 and PC2). The two Varimax rotated PCs explain around 88% of the total variance in the dataset, with 45% by PC1 and 43% by PC2. Figure 5 shows the loadings of two PCs, equivalent to Pearson correlations between each PC time series and daily PM10 concentrations at individual stations. The PC loadings (correlations) identify two distinct air quality clusters: 1) the WNW subregion, with high loadings on PC1 for stations in the northern and western parts of the valley; 2) the SE subregion, with high loadings on PC2 for stations in the south-eastern part of the valley. Hence, the variability of PM10 pollution in the valley is summarised into two linearly independent (dimensionless) time series, i.e., PC1 and PC2 scores, despite of the significant inter-site difference in PM10 concentrations (as shown in the previous section).
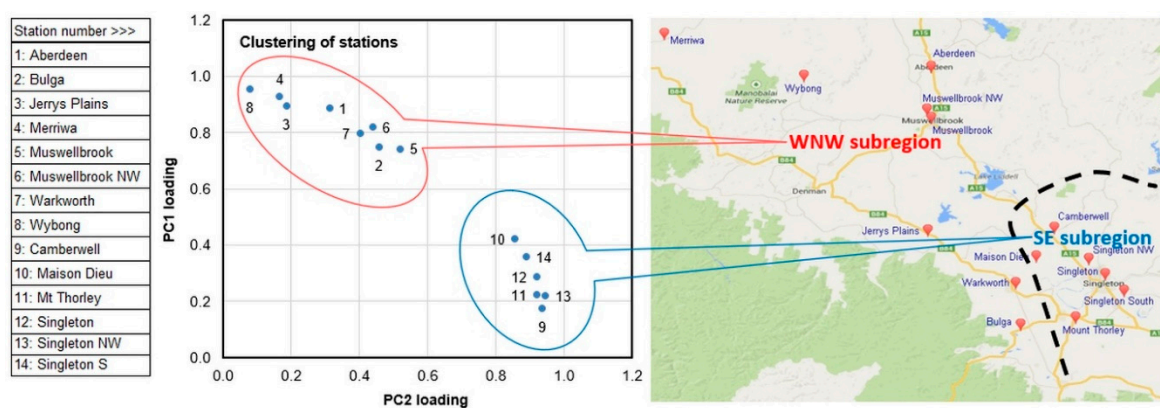
Figure 5. Identification of two air quality subregions in the Upper Hunter Valley based on Varimax rotated principal component analysis (RPCA) of daily PM10 data in 2012-2022 (exceptional event days excluded). Left panel: key of station number; middle panel: scatter plot of loadings for first two rotated principal components (PC1, PC2); right panel: map showing two air quality subregions with red balloons indicating station locations in the UHAQMN. Source of base map source: Google.

11

RPCA on the all-day dataset (including exceptional event days) could produce similar results, except that the two leading PCs are reversed in order and two clusters show slightly reduced separations (Figure S2). Obliquely rotated PCA (Jiang, 2011) on the relevant datasets generated very similar (almost identical) results, confirming the existence of a strong *underlying simple structure (clusters)* (Harman, 1976) in the daily PM10 dataset, i.e., the existence of two distinct air quality subregions in the valley.

These results are highly consistent with OEH (2017a), who initially proposed the two-subregion property of PM10 pollution from a shorter (4-year) normal-day dataset for the study region. Hence, based on the longer (11-year) dataset, the present analysis has verified the stability and robustness of two air quality subregions in the Upper Hunter Valley, despite of the impact of (a small number of) exceptional events.

The division of the valley into two air quality subregions can somehow be expected, primarily due to the valley's NW-SE oriented slope terrain and the prevalence of valley-following air follows in the lower boundary layer. Previous work revealed that the most frequent winds are north-westerlies and south-easterlies in the valley (e.g., DPE, 2022; OEH, 2017b; Holmes, 2008; Bridgman and Chambers, 1981). Northwesterly winds can blow dust generated in the upper (NW) part of the valley southeastward (down slope), contributing to elevated PM10 concentrations in areas near the bottom end (i.e., the SE subregion) of the valley (Hyde et al., 1981). In contrast, southeasterly winds may transport pollutants from the lower valley northwestward (up-slope), resulting in elevated PM10 pollution at the upper end (i.e., the WNW subregion) of the valley. Some recent observational case studies (OEH, 2017b; DPE, 2022) also indicated that emissions from open-cut mines can lead to elevated PM10 pollution typically in the lower (southeast) part of the valley in particular under north-westerly winds.

4.3. Temporal Pattern – Identification of Key Variation Modes

The previous section has shown that the variability of PM10 pollution in the study region can be expressed with two linearly independent PC time series, that is, PC1 and PC2 scores represent the distinct covariational features of PM10 pollution at stations in the WNW and SE subregions. Wavelet analysis was used to identify the dominant variability modes in the PC time series and how those modes changed over time. Figure 6 and Figure 7 illustrate the wavelet analysis results for PC1 (Figure 6a; WNW subregion) and PC2 (Figure 7a; SE subregion) scores from the RPCA of the normal-day dataset (Section 3.2). It is most distinguishable that the wavelet power spectrum peaks near the 360-day time scale (annual mode, significant at the 0.05 level for a Chi-square test) persistently across all years for both subregions (Figure 6b and Figure 7b). The higher variance occurred in spring to summer for the WNW subregion (Figure 6c), but in winter to spring for the SE subregion (Figure 7c).

The annual variability mode in PM10 pollution can be linked to the seasonal changes of weather and climatic conditions that influence pollutant (PM10) emission and dispersion conditions in the valley. Higher rainfall (hence lower dust generation and higher wet deposition) tends to occur in summer and early autumn, and lower rainfall (hence higher dust generation and lower wet deposition) in winter and early spring (OEH, 2017b). The most frequent winds in the valley are north-westerlies in winter and south-easterlies in summer, with wind directions less defined in autumn and spring (Holmes, 2008; OEH, 2017a). Consequently, the prevailing northwesterly winds and higher PM10 emissions in winter and spring provide a high potential for transporting or accumulating PM10 pollution over the SE subregion. In contrast, the prevailing southeasterly flows or sea breezes in summer tends to transport pollutants northwestward, potentially resulting in elevated PM10 pollution in the upper part of the valley (i.e., in the WNW subregion). This seasonality in PM10 pollution will be further illustrated in Section 4.4 for two air quality subregions.

There are also intermittent wavelet power peaks at around 120 days, 30~90 days, and shorter (less than 30 days) time scales (Figure 6b, Figure 7b), indicating the signals of triannual, intra-seasonal and shorter term variation modes in the PC time series. These variation modes appeared stronger in the SE subregion (PC2 time series) than the WNW subregion (PC1 time series). The signal strengths changed dramatically across years, manifesting phase difference between two subregions. For

example, the signal for the variability mode of around 120 days in the WNW subregion was more intense in summertime across years 2013-2015 and 2017-2018, but a lot weaker during some periods in 2021-2022 (Figure 7b). In comparison, this variability mode in the SE subregion showed relatively higher power during wintertime in 2012, 2013 and 2018, but lower power during some periods in 2016 and 2019-2022. Of note is the high wavelet power at around 5-6 years for both subregions (Figure 6b; Figure 7b), implying a signal at the interannual scale in the PC time series. However, most of the spectra at this time scale fall within the cone of influence (COI) region (under the curved line), which are within the uncertainty range associated with the edge effect in Fourier transforms (Torrence and Compo, 1998).

Figure 6. WNW subregion temporal variability patterns. (a) The first principal component (PC1) scores used for the wavelet analysis, derived from RPCA on the normal-day dataset where missing data and data for exceptional event days 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 serial correlation 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.
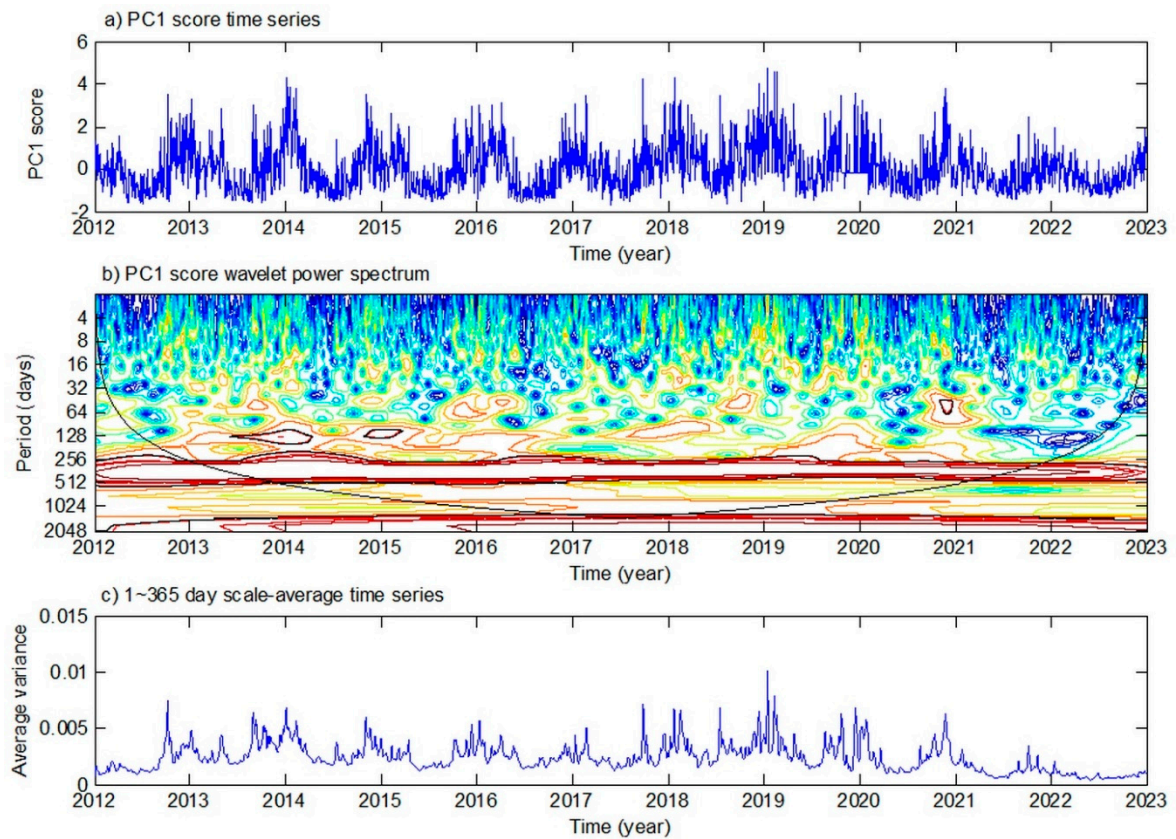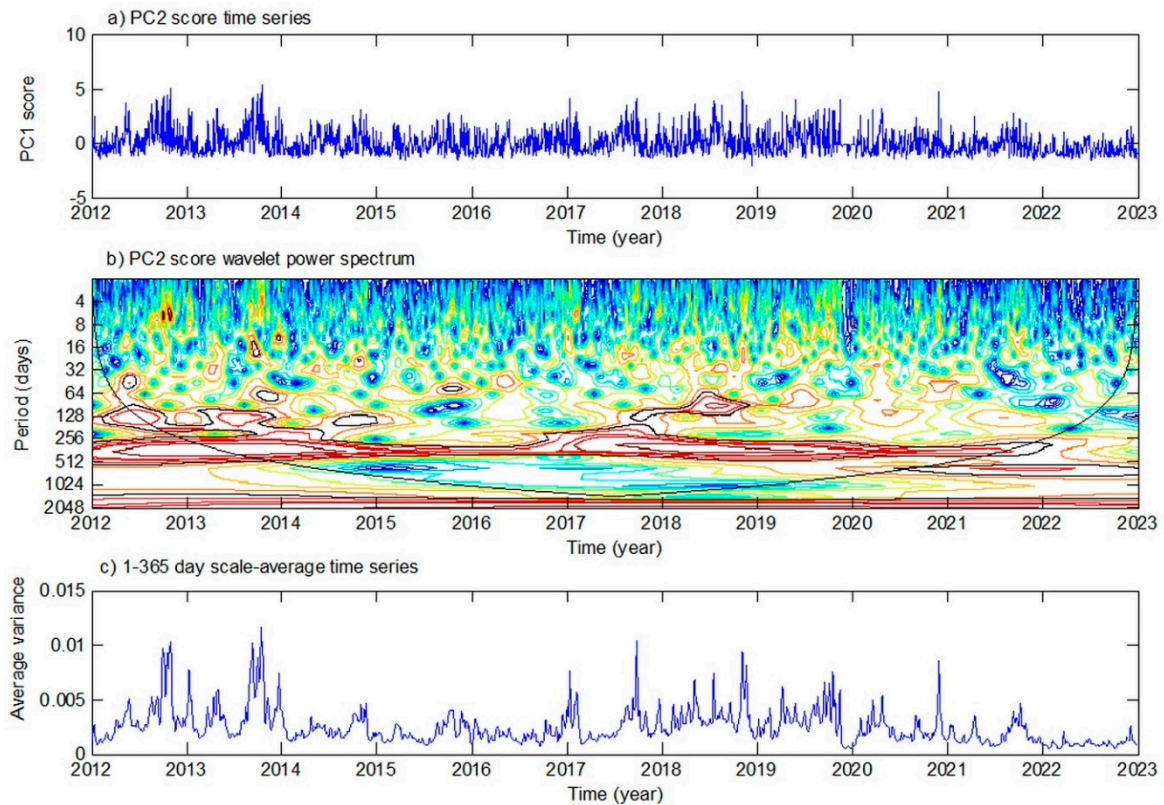
Figure 7. SE subregion temporal variability patterns. (a) The second principal component (PC2) scores used for the wavelet analysis, derived from RPCA on the normal-day dataset where missing data and data for exceptional event days were replaced with overall median for each station. (b) The local normalised wavelet power spectrum of (a) using the Morlet wavelet. The contour lines are at normalised variances of 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.

The variability modes at times scales of 30-90 days and 120 days in PM10 pollution are not yet readily understood. Many studies have examined the signal of intraseasonal oscillations at time scales of 30-90 or 30-60 days in tropical atmospheric or oceanic variables in tropical atmospheric or oceanic variables, which can often be linked to the Madden-Julian Oscillation (MJO) in the tropics and the extra-tropical teleconnection patterns such as El Niño-Southern Oscillation (ENSO) (e.g., Zhang, 2005; Weickmann and Berry, 2009; Gushchina and Dewitte, 2012). Such oscillations are also investigated extensively in atmospheric variables associated with monsoon activities (e.g., Zhou and Chan, 2008; Kikuchi et al., 2012) and to a lesser extend in those at higher latitudes (e.g., Rimbu et al., 2012; 2013). For example, Rimbu et al., (2012) revealed two intraseasonal variability patterns in synoptic observations (temperature, wind speed, sea level pressure) at a high-latitude Antarctic research station, Neumayer (70°S, 8°W). The dominant pattern manifests with out-of-phase variations of temperature and wind speed with sea level pressure at time scales of around 40 and 80 days, which can be related to tropical intraseasonal oscillations via large-scale eastward propagating atmospheric circulation wave-trains. In contrast, the second pattern was characterised by the in-phase variations between temperature, wind and sea level pressure at time scales of around 35, 60 and 120 days, which can be attributed to the Southern Annular Mode/Antarctic Oscillation (SAM/AAO). Drawing upon these studies, one may speculate whether the PM10 variability modes at time scales of 30-90 and 120 days (and potentially at the multi-year time scale) could be related to

the influence of large-scale climate drivers such as MJO, ENSO and SAM, which are known to modulate the weather and climate in Australia. This aspect deserves further attention in future work.

A wavelet analysis was also performed on the two leading PCs from the RPCA on the all-day PM10 dataset (measurements for exceptional event days included), with results shown in Figures S3 and S4 (note that the order of two PCs was reversed). Overall, the wavelet power of the PC time series could identify generally similar dominant modes to those described above. However, clearly there are significant distortions (increases) in the wavelet power peaks during mid-2017 to mid-2020 for the SE subregion (PC1 times series; Figure S3) and during mid-2019 to mid-2020 for the WNW subregion (PC2 time series; Figure S4). The distortions appear more intense in the SE subregion than the WNW subregion, essentially reflecting the significant impact of high PM10 measurements during the widespread dust storms in 2018 and the unprecedented Black Summer bushfires in 2019-2020 across large areas of NSW.

In addition, the shorter-term (under 30 days) variability in PM10 pollution is attributable to day-to-day changes in emission conditions and the effects of local and synoptic weather variability, as is demonstrated in some previous studies for other regions (e.g., Jiang et al., 2014; Jiang et al., 2017). This aspect will be discussed in a companion paper of this text.

4.4. Illustrating the Annual and Interannual Variability in PM10 Pollution

The knowledge of the spatial and temporal variability patterns (Sections 4.2 and 4.3) can be used to facilitate the summarisation of PM10 data with increased clarity. To demonstrate this utility, we use PC scores to visualise the general annual (seasonality) and interannual variation patterns of PM10 pollution for two subregions in the same (one) framework. Figure 8 shows the heat maps of mean PC1 and PC2 scores for normal days (exceptional event days excluded), by month and year separately for each subregion. The analysis on PC scores for the all-day (full) dataset makes no significant difference to the findings reported here.

It is visually clear that mean PC scores identify very distinct annual variation patterns (seasonality) in PM10 pollution between two subregions. In the WNW subregion (Figure 8a), positive PC1 mean scores (hence high mean PM10 levels) occurred mostly in warmer months (late spring to summer, highest in November to February), but negative scores (low mean PM10 levels) in cooler months (May to September, lowest in June). In the SE subregion, however, the seasonal variability pattern appears complex (Figure 8b). Positive mean PC2 scores (thus higher mean PM10 levels) tended to occur in some cooler months (in particular, May and July-October) and negative scores (thus lower PM10 levels) in some warmer months (November-April) in most years. As expected, the mean scores for November and December in 2019 are also high, reflecting the broad-scale higher mean PM10 pollution associated with impacts of the unprecedented 2019-2020 spring-summer bushfires in NSW (DPE, 2020).

The interannual changes in PC scores appear similar between two subregions (Figure 8). Higher values occurred in 2018-2019 and secondarily 2012-2013 (this variability signal is slightly weaker in the WNW subregion), but lower values in 2021-2022 and secondarily 2015-2016, with 2022 being the cleanest year on record (DPE, 2023). Of note is that June has generally negative mean scores across all years in both subregions, indicating generally better air quality for this time of the year.

(a) WNW subregion - mean PC1 scores (exceptional events excluded)

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2012	.	0.0	-0.1	0.0	-0.8	-0.9	-0.8	-0.9	-0.3	0.4	0.9	1.1
2013	0.8	0.1	0.3	0.2	-0.2	-0.9	-0.8	-0.8	-0.1	0.0	0.0	0.8
2014	1.8	1.4	-0.1	-0.3	-0.7	-0.9	-0.8	-0.4	-0.4	0.4	1.3	0.2
2015	-0.1	0.5	0.8	-0.4	-0.8	-0.9	-0.9	-0.9	-0.5	0.5	0.0	0.5
2016	0.3	0.9	0.7	1.0	-0.7	-1.0	-1.0	-0.9	-0.9	-0.6	0.2	0.8
2017	0.5	0.9	-0.1	-0.1	-0.4	-0.4	-0.8	-0.5	0.0	0.4	0.4	0.5
2018	1.5	1.3	0.5	0.7	0.5	-0.6	-0.3	0.0	-0.1	.	0.9	1.1
2019	1.6	1.3	0.9	0.6	-0.3	-0.4	-0.4	0.2	0.5	1.3	0.6	4.3
2020	1.8	0.8	0.2	-0.1	-0.6	-0.7	-0.7	-0.7	0.1	0.0	1.0	0.5
2021	-0.1	0.1	-0.3	-0.1	-0.5	-0.8	-0.8	-0.6	-0.4	-0.2	-0.1	0.1
2022	0.3	-0.1	-0.2	-0.3	-0.5	-0.8	-0.6	-0.7	-0.6	-0.3	0.0	0.4

(b) SE subregon - mean PC2 scores (exceptional events excluded)

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2012	.	-0.3	-0.4	-0.1	1.1	-0.3	0.0	1.2	1.4	1.2	0.3	0.0
2013	0.0	-0.6	-0.3	0.1	0.3	-0.1	0.3	1.1	1.1	1.3	-0.4	-0.1
2014	-0.4	-0.6	-0.5	-0.2	0.4	0.0	0.4	-0.3	0.0	0.5	0.1	-0.3
2015	-0.4	-0.8	0.2	-0.5	-0.2	-0.1	-0.2	0.4	-0.1	0.2	-0.6	-0.3
2016	-0.4	-0.2	-0.6	-0.2	0.7	-0.4	-0.1	0.1	0.0	0.2	0.4	0.5
2017	0.1	-0.2	-0.5	-0.5	0.0	-0.1	0.7	0.7	1.3	-0.1	-0.8	-0.1
2018	-0.1	-0.2	-0.4	0.2	1.0	-0.2	1.3	1.0	-0.7	.	0.3	-0.4
2019	0.1	-0.4	-0.4	-0.2	0.6	0.3	0.6	0.9	0.3	0.3	1.4	1.0
2020	-0.4	-0.1	-0.4	0.5	-0.1	-0.2	-0.3	0.0	0.0	-0.3	0.0	-0.5
2021	-0.3	-0.7	-0.6	-0.2	-0.2	-0.4	0.1	0.3	0.1	-0.1	-0.6	-0.6
2022	-0.7	-0.8	-0.8	-0.9	-0.5	-0.1	-0.7	-0.5	-0.6	-0.7	-0.5	-0.6

Figure 8. Mean PC1 and PC2 scores by month and year for two subregions. PC1 and PC2 are from RPCA on the normal-day PM10 dataset (exceptional event days excluded). Blank cell indicates there are insufficient data point for valid calculation. Colour scale: green - relatively low value; yellow – near medium value; red - relatively high value.

4.4.1. Mean PM10 Levels and Total Number of Poor Air Quality Days

In this section, the annual and interannual variability patterns in mean PM10 concentrations and total number of poor air quality days are illustrated at the station level for two air quality subregions (Figures 9 and 10). The results confirm the above identified variability patterns in PC scores, but with greater details on PM10 pollution at individual locations/stations.

In the WNW subregion (Figure 9a), the mean PM10 levels and number of poor air quality days were generally higher in warmer months (i.e., late spring to summer, highest in November to January) but lower in cooler months (lowest in June). In the SE subregion (Figure 9b), however, two statistics were generally higher in August-November (late winter to spring) and May (highest in September and October) but lower in February, March and June (lowest in June), with near average values in December to January.

a. WNW subregion

		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	All months
Mean PM10 levels ($\mu g/m^3$)	Merriwa	19.9	18.9	15.6	14.6	11.1	8.0	8.7	10.2	13.0	15.3	17.6	18.5	14.1
	Wybong	22.4	20.4	17.2	16.2	12.1	8.7	9.3	11.1	13.9	17.1	19.6	20.8	15.5
	Bulga	21.5	18.5	16.7	15.9	13.9	10.1	11.9	14.2	16.5	19.2	20.2	19.9	16.4
	Aberdeen	20.5	19.8	17.5	17.5	15.0	11.3	12.1	13.6	15.8	17.4	18.9	18.7	16.4
	Jerrys Plains	23.0	20.9	18.1	16.8	13.2	9.0	10.1	12.0	16.5	20.0	21.6	22.2	16.8
	Muswellbrook NW	22.6	21.2	18.7	19.1	16.7	12.8	14.1	15.9	18.0	20.3	21.4	21.0	18.4
	Warkworth	26.4	23.4	21.1	19.7	17.7	13.9	15.1	17.6	21.7	24.6	25.1	25.8	20.9
	Muswellbrook	23.6	21.9	19.9	21.2	20.8	17.7	19.0	20.0	20.6	21.9	23.0	21.9	20.9
Poor air quality days (count)	Aberdeen	0	2	0	0	0	0	0	1	0	1	0	0	4
	Merriwa	1	0	0	0	0	0	1	0	0	1	0	2	5
	Muswellbrook NW	0	2	0	1	0	0	1	0	0	2	3	1	10
	Wybong	3	4	0	0	0	0	1	0	1	1	0	0	10
	Muswellbrook	1	2	0	1	2	0	2	1	1	2	1	0	13
	Bulga	4	2	0	1	0	0	1	0	0	2	2	3	15
	Jerrys Plains	5	3	0	1	0	0	0	0	2	3	6	1	21
	Warkworth	6	6	2	1	0	0	0	0	2	6	5	6	34

b. SE subregion

		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	All months
Mean PM10 levels ($\mu g/m3$)	Singleton S	19.4	17.0	16.2	17.3	18.6	14.5	16.9	19.8	19.8	20.9	19.8	19.1	18.3
	Singleton	20.3	17.9	17.1	19.1	21.9	17.3	20.1	22.4	21.6	22.1	20.0	19.0	19.9
	Maison Dieu	24.2	20.3	19.5	20.8	21.8	15.8	19.3	23.4	24.7	25.7	23.4	22.8	21.8
	Singleton NW	21.6	18.7	18.6	21.2	25.0	18.3	21.8	25.4	25.2	24.8	22.5	21.2	22.0
	Mt Thorley	20.6	18.4	18.4	21.2	25.1	17.8	21.7	26.3	26.3	25.7	22.8	21.0	22.1
	Camberwell	23.6	20.0	19.9	22.3	26.3	20.0	25.3	28.8	28.3	28.3	25.5	22.9	24.3
Poor air quality days (count)	Singleton S	0	2	0	2	1	0	1	0	1	5	2	0	14
	Singleton	3	0	0	0	0	0	1	0	5	7	5	1	22
	Maison Dieu	7	5	1	4	2	0	5	6	18	25	10	5	88
	Singleton_NW	5	4	2	7	11	0	4	11	19	24	9	5	101
	Mt Thorley	3	6	1	9	15	1	12	20	27	25	11	6	136
	Camberwell	13	7	5	9	11	2	19	27	36	35	17	9	190

Figure 9. Monthly mean PM10 levels and total number of poor air quality days (with PM10 levels > 50 $\mu g/m^3$) for stations in the (a) WNW and (b) SE subregions. Rows are sorted by the "All months" column (multi-year station means or total number of poor air quality days). Data: daily PM10 measurements in 2012–2022 (excluding exceptional event days). Colour scale: green – relatively low value; yellow – near medium value; red – relatively high value.

Broadly the interannual variation patterns are similar between two subregions, with higher values in 2018-2019 and secondarily 2012-2013 (this variability signal is slightly weaker in the WNW subregion) but lower values in 2021-2022 and secondarily 2015-2016 (Figure 10). Consistently, OEH (2017a) also noted the higher PM10 pollution in 2012 and 2013 than other years during 2012-2015. OEH (2014) and DPE (2019, 2020) suggested that the (prior) higher-than-average temperature and prolonged drought conditions across NSW (broadly most of the Australian continent) had contributed to the poorer air quality observed in years of 2012-2013 and 2018-2019. DPE (2022, 2023) noted that the cooler and wetter climate conditions contributed to improved air quality in 2021-2022.

Overall, (as expected) the SE subregion has generally higher mean PM10 pollution and more exceedance days (highest at the four source-impacted sites) than the WNW subregion, as is consistent with local air quality experience (DPE, 2022; Holmes and Associates, 1996). Of the three larger population centre sites, higher PM10 pollution occurred at Muswellbrook and Singleton. In contrast, Aberdeen experienced relatively better air quality, with the mean PM10 levels and total number of exceedance days comparable to Merriwa (background site). Also of note is that June appears the cleanest across all months, recording the lowest mean pollution and least number of poor air quality days across almost all stations.

The distribution patterns identified for the normal-day dataset were further confirmed through a similar analysis on the all-day dataset (Figures S5 and S6). In comparison, however, the mean PM10 levels and poor air quality days for the all-day dataset show significant increase for November-January and 2018-2020, primarily associated with the broad-scale impact of dust storms in 2018 and the Black Summer Bushfires in 2019-2020 spring-summer months (DPE, 2020).

(a) WNW subregion

		2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	All years
Mean PM10 levels (µg/m3)	Merriwa	14.2	14.5	15.0	12.8	13.4	13.8	17.9	17.7	13.8	11.6	11.2	14.1
	Wybong	15.4	14.9	16.8	14.4	15.2	16.2	20.3	19.4	14.8	12.5	11.7	15.5
	Bulga	18.7	18.5	17.5	14.7	15.8	16.9	20.1	19.2	16.3	12.9	10.6	16.4
	Aberdeen	17.0	16.9	17.7	14.9	15.5	17.2	20.9	21.3	15.1	12.8	12.3	16.4
	Jerrys Plains	10.8	17.8	18.0	15.2	16.7	17.7	22.8	21.8	18.1	13.5	13.3	16.8
	Muswellbrook NW	19.1	18.4	19.0	16.3	16.4	18.1	23.5	25.0	18.2	15.5	14.3	18.4
	Warkworth	21.1	20.7	20.5	17.9	18.4	21.4	25.1	24.1	21.3	20.6	19.3	20.9
	Muswellbrook	21.8	22.0	21.3	18.8	19.1	21.3	25.8	26.2	20.0	18.2	16.6	20.9
Poor air quality days (count)	Aberdeen	0	0	0	0	0	0	0	4	0	0	0	4
	Merriwa	1	0	1	0	0	0	3	0	0	0	0	5
	Muswellbrook NW	1	0	0	0	0	0	2	5	1	0	1	10
	Wybong	1	0	2	0	1	1	2	3	0	0	0	10
	Muswellbrook	1	1	0	0	0	0	5	5	1	0	0	13
	Bulga	2	4	2	1	0	0	3	3	0	0	0	15
	Jerrys Plains	0	3	5	0	0	1	5	4	3	0	0	21
	Warkworth	0	4	3	2	0	0	9	9	3	3	1	34

(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	19.6	18.1	16.5	17.8	18.9	21.5	22.1	18.0	16.4	14.0	18.3
	Singleton	22.3	22.7	20.9	18.9	19.1	20.2	22.5	22.4	18.7	17.3	14.5	19.9
	Maison Dieu	25.8	24.8	22.5	20.0	20.1	22.4	26.4	26.7	20.0	17.5	14.0	21.8
	Singleton NW	25.9	25.1	22.6	20.5	21.6	22.1	25.4	25.8	20.6	18.7	15.2	22.0
	Mt Thorley	24.8	23.9	21.4	19.4	22.4	24.6	27.1	27.4	20.4	18.8	14.2	22.1
	Camberwell	26.4	26.9	24.5	21.6	24.2	26.5	29.3	29.6	22.5	20.4	16.0	24.3
Poor air quality days (count)	Singleton S	2	3	0	0	0	0	2	3	3	1	0	14
	Singleton	6	9	0	0	0	0	3	3	1	0	0	22
	Maison Dieu	20	22	5	3	0	6	16	10	6	0	0	88
	Singleton_NW	29	22	5	2	2	7	15	14	5	0	0	101
	Mt Thorley	28	20	3	5	4	17	26	24	5	4	0	136
	Camberwell	23	31	11	9	8	27	36	34	8	3	0	190

Figure 10. 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 (excluding exceptional event days). Colour scale: green - relatively low value; yellow – near medium value; red - relatively high value.

5. Summary and Conclusion

The present study has examined the spatial-temporal variability of PM10 pollution in the Upper Hunter Valley, based on the long-term (2012-2022), multi-site air quality data from the UHAQMN and application of the RPCA (rotated principal component analysis) and wavelet analysis techniques. The impact of exceptional events has also been examined at some depths, with some results given as Supplementary Materials due to space limit. The main findings are summarised below:

(1) The RPCA has identified two air quality subregions in the Upper Hunter Valley, one in the west/northwest (WNW) part, the other in the southeast (SE) part of the valley. This finding verifies the previous work by OEH (2017a), confirming the spatial regionalization property of

PM10 pollution in the valley despite of the significant impacts from exceptional events such as widespread dust storms or vegetation fires. Hence, it is possible to characterise the air quality *variability* in the valley with two linearly independent and dimensionless PC score time series, or alternatively PM_{10} time series from a subset of (representative) monitoring stations from the subregions.

(2) Wavelet analysis has identified the annual cycle (and potentially interannual variability) to be most persistent mode(s) of temporal variability of PM10 pollution in two subregions. There were also intermittent signals at time scales of around 120 days (triannual), 30~90 days (intraseasonal) and under 30 days, with the mode intensity changing dramatically across time. These intermittent signals in PM10 pollution are not yet readily understood – which deserve further attention in future research.

(3) The knowledge on the temporal and special variability modes can be used to facilitate the summarisation of PM10 data with increased clarity, as has been demonstrated for illustrating the seasonal and interannual variability patterns for two air quality subregions. The seasonal variation patterns differed between subregions - higher pollution occurred in warmer months (summer in particular) in the WNW subregion, but in late winter to spring in the SE subregion. The interannual variation patterns were broadly similar for two subregions, with higher PM10 pollution in 2018-2019 and 2012-2013 (but lower PM10 pollution in other years, whereas 2022 observed the lowest PM10 pollution on record).

(4) Relatively higher daily PM10 levels were recorded in the SE southeast of the valley, highest at four direct source-impacted locations, which are relatively close to the open-cut mining sites scattered near the bottom end of the valley. The stations in the WNW subregion had generally lower PM10 pollution, in particular with one background site (Merriwa) and one larger population site (Aberdeen) recording relatively good air quality. The other two larger population centre sties, Singleton and Muswellbrook, experienced generally moderate PM10 pollution levels when compared to other stations.

(5) The exceptional events, including widespread dust storms in 2018 and bushfires in 2019-2020 had significant impacts on regional air quality, resulting in significant increases in minimum, mean and maximum PM10 levels, as well as the number of poor air quality days, across all monitoring stations in the valley. The impact was most significant in the SE subregion in particular at the four source-impacted stations, indicating the combined impacts from local, remote and/or incidental emissions. This suggests a need for enhanced management of open-cut mining activities even during exceptional event days.

(6) The inclusion or exclusion of exceptional event days in the RPCA and wavelet analysis performed in the study did not make significant difference to the findings reported here. This suggests that the identified spatial and temporal variability modes in PM10 pollution are primarily associated with changes in local PM10 emission and the influence of metrological conditions in the Upper Hunter Valley.

In conclusion, the present study has illustrated a quantitative holistic approach for examining the spatial-temporal variability of PM10 pollution in the Upper Hunter Valley. The findings from this study will be used to improve air quality reporting and forecasting in NSW. Future work may point to many directions. For example, one area is to investigate the underlying mechanisms associated with the PM10 variation modes at time scales of 30-90 days, 120 days and multi-years, in particular when even a longer dataset become available. Previous studies showed that both local-, synoptic- and large-scale climatic conditions could affect local air quality in a region (e.g., Jiang et al., 2014; Jiang et al., 2017). Research in this aspect is ongoing, with details to be presented in a companion (Part II) paper of this text. The methodology and results can also be useful for air quality research in similar regions elsewhere.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org: Figure S1 Box plots of daily PM_{10} data by station for data in 2012–2022 (including exceptional event days). The lower and upper boundaries of the 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 exceptional event days). Left panel: key of station number; middle panel: scatter plot of loadings for first two rotated principal components (PC1, PC2); right panel: map showing two air quality subregions with red balloons indicating station locations in the UHAQMN. Base map source: Google. Figure S3 SE subregion temporal variability patterns. (a) The first principal component (PC1) scores used for the wavelet analysis, derived from RPCA of all-day dataset (exceptional event data included) 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. (a) The second principal component (PC2) scores used for the wavelet analysis, derived from of all-day dataset (exceptional event data included) 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 Monthly mean PM10 levels and total number of poor air quality days (with PM10 levels > 50 $\mu g/m^3$) 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. Figure S6 Annual mean PM10 levels and total number of poor air quality days (with PM_{10} levels > 50 $\mu g/m^3$) 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.

Author Contributions: Conceptualisation: N Jiang; methodology, data curation and analysis: N Jiang; writing - original draft preparation: N Jiang; writing - review and editing: N Jiang, M Riley, M Azzi, P Puppala, H Duc, G Di Virgilio. All authors have read and agreed to the submitted/published version of the manuscript.

Data Availability Statement: The data presented in this study are publicly available at https://www.airquality.nsw.gov.au/air-quality-data-services/air-quality-api.

Conflicts of Interest: The authors declare that they have no conflict of interest.

References

ABARES. 2023. About my region – Hunter Valley (excluding Newcastle) New South Wales. Australian Bureau of Agricultural and Resource Economics and Sciences (ABARES): Canberra, Australia. Retrieved 27 September 2023, from: http://www.agriculture.gov.au/abares/research-topics/aboutmyregion.

ABS. 2023. Search Census data by geography - Census 2021: https://abs.gov.au/census/find-census-data/search-by-area. Australian Bureau of Statistics (ABS): Canberra, Australia. Retrieved 11 October 2023.

Adamowski JF, Chan HF. 2011. A wavelet neural network conjunction model for groundwater level forecasting. Journal of Hydrology 407: 28-40. DOI: 10.1016/j.jhydrol.2011.06.013

Anderson JO, Thundiyil JG, Stolbach A. 2012. Clearing the Air: a Review of the Effects of Particulate Matter Air Pollution on Human Health. Journal of Medical Toxicology 8: 166-175.

Barnett AG, Williams GM, Schwartz J, Best TL, Neller AH, Petroeschevski AL, Simpson RW. 2006. The effects of air pollution on hospitalizations for cardiovascular disease in elderly people in Australia and New Zealand cities. Environmental Health Perspectives 114(7): 1018-1023.

Bridgman HA, Chambers AJ, 1981. Air Quality in the middle Hunter: the extensive study periods. A report to the New South Wales State Pollution Control Commission. University of Newcastle: Newcastle, Australia.

Cattell, R.B., 1966. The scree test for the number of factors. *Multivar. Behav. Res*. **1:** 245-276.

Connor L, Higginbotham N, Freeman S, Albrecht G. 2008. Watercourses and Discourses: Coalmining in the Upper Hunter Valley, New South Wales. Oceania 78 (1): 1-128. https://onlinelibrary.wiley.com/doi/10.1002/j.1834-4461.2008.tb00029.x.

CSIRO. 2013. Inquiry into the impacts on health of air quality in Australia. Senate Standing Committees of Community Affairs. Commonwealth Science and Industrial Research Organisation (CSIRO) Submission 12/472: Acton, Australia.

Czernecki B, Półrolniczak M, Kolendowicz L, Marosz M, Kendzierski S, Pilguj N. 2017. Influence of the atmospheric conditions on PM10 concentrations in Poznań, Poland. Journal of Atmospheric Chemistry 74: 115–139.

DPE. 2019. Annual Air Quality Statement 2018. State of NSW Department of Planning and Environment. Retrieved 10 May 2023, from: https://www.environment.nsw.gov.au/research-and-publications/publications-search/nsw-annual-air-quality-statement-2018.

DPE. 2020. *Air quality special statement spring-summer 2019–20*. Department of Planning and Environment. Retrieved 1 October 2023 from: https://www.environment.nsw.gov.au/topics/air/nsw-air-quality-statements/air-quality-special-statement-spring-summer-2019-20.

DPE. 2021. Annual Air Quality Statement 2020. State of NSW Department of Planning and Environment. Retrieved 12 May 2022, from: https://www.environment.nsw.gov.au/topics/air/nsw-air-quality-statements/annual-air-quality-statement-2020.

DPE. 2022. Upper Hunter Air Quality Monitoring Network. 5-year review 2022. State of NSW Department of Planning and Environment. Retrieved 25 March 2023, from https://www.environment.nsw.gov.au/research-and-publications/publications-search/upper-hunter-air-quality-monitoring-network-5-year-review-2022.

DPE. 2023. Annual Air Quality Statement 2022. State of NSW Department of Planning and Environment. Retrieved 25 Sept 2023, from https://www.environment.nsw.gov.au/topics/air/nsw-air-quality-statements/annual-air-quality-statement-2022.

DPI, 2018. New South Wales Department of Primary Industries Performance, Data & Insights. NSW Department of Primary Industries: Orange, NSW. Retrieved 27 September 2023, from https://www.dpi.nsw.gov.au/__data/assets/pdf_file/0005/841163/15116_PDI2018_DecEdition.pdf.

EPHC. 2010. Expansion of the multi-city mortality and morbidity study. Executive summary and summary report. Environment Protection and Heritage Council (EPHC): Canberra, Australia.

Fortelli A, Scafetta N, Mazzarella A. 2016. Influence of synoptic and local atmospheric patterns on PM10 air pollution levels: a model application to Naples (Italy), Atmospheric Environment 143: 218-228, https://doi.org/10.1016/j.atmosenv.2016.08.050.

Giri D, Krishna Murthy V, Adhikary PR. 2008. The Influence of Meteorological Conditions on PM10 Concentrations in Kathmandu Valley. International Journal of Environmental Research 2(1): 49-60.

Gushchina D, Dewitte B. 2012. Intraseasonal Tropical Atmospheric Variability Associated with the Two Flavors of El Niño. Monthly Weather Review 140: 3669-3681. DOI: 10.1175/MWR-D-11-00267.1

Harman HH. 1976. *Modern factor analysis*. Third edition. University of Chicago Press: Chicago.

Hertzog L, Morgan GG, Yuen C, Gopi K, Pereira GF, Johnston FH, Cope M, Chaston TB, Vyas A, Vardoulakis S, Hanigan IC. 2024. Mortality burden attributable to exceptional PM2.5 air pollution events in Australian cities: A health impact assessment. Heliyon 10 (2): 2024. https://doi.org/10.1016/j.heliyon.2024.e24532.

Hibberd MF, Selleck PW, Keywood MD, Cohen DD, Stelcer E, Atanaclo AJ. 2013. Upper Hunter Particle Characterisation Study. Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia.

Holmes 2008, Upper Hunter Valley Monitoring Network Design, 15th February 2008, prepared for NSW Department of Environment and Climate Change, Sydney.

Holmes and Associates. 1996. Air quality study: cumulative effects due to atmospheric emissions in the Upper Hunter Valley, NSW. New South Wales Department of Urban Affairs and Planning: Sydney, Australia.

Huang R, Ning H, He T, Bian G, Hu J, Xu G. 2018. Impact of PM_{10} and meteorological factors on the incidence of hand, foot, and mouth disease in female children in Ningbo, China: a spatiotemporal and time-series study. Environmental Science and Pollution Research International: 2018 July 1. doi: 10.1007/s11356-018-2619-5.

Hyde R, Malfroy H, Watt GN, Maynard J. 1981. The Hunter Valley meteorological study: Interim Report to the New South Wales State Pollution Control Commission on Mesoscale Meteorology in the Hunter Valley. Macquarie University: Sydney, Australia.

Jiang N. 2010. *Application of Two Different Weather Typing Procedures, An Australian Case Study*. VDM Verlag Dr. Mueller: Germany.

Jiang N 2011, A new objective procedure for classifying New Zealand synoptic weather types during 1958–2008. *International Journal of Climatology* 31: 863–879. Doi: 10.1002/joc.2126.

Jiang N, Dirks KN, Luo K. 2014. Effects of local, synoptic and large-scale climate conditions on daily nitrogen dioxide concentrations   in Auckland, New Zealand. *Int. J. Climatol.* **34**: 1883–1897. doi: 10.1002/joc.3808.

Jiang N, Hay JE, Fisher GW. 2005. Effects of meteorological conditions on concentrations of nitrogen oxides in Auckland. Weather and Climate, 24: 15-34.

Jiang N, Scorgie Y, Hart M, Riley ML, Crawford J, Beggs PJ, Edwards GC, Chang L, Salter D, Di Virgilio G. 2017. Visualising the relationships between synoptic circulation type and air quality in Sydney, a subtropical coastal-basin environment. International Journal of Climatology 37: 1211-1228. doi: https://doi.org/10.1002/joc.4770.

Keywood MD, Emmerson KM, Hibberd MF. 2016. Ambient air quality: Health impacts of air pollution. In: Australia state of the environment 2016. Australian Government Department of the Environment and Energy: Canberra, Australia. Doi: 10.42216/94/58b65c70bc372.

Keywood MD, Hibberd MF, Selleck PW, Desservettaz M, Cohen DD, Stelcer E, Atanacio AJ, Scorgie Y, Chang L. 2020 Sources of Particulate Matter in the Hunter Valley, New South Wales, Australia. *Atmosphere 11 (*4). https://doi.org/10.3390/atmos11010004.

Kikuchi K, Wang B, Kajikawa Y. 2012. Bimodal representation of the tropical intraseasonal oscillation. Clim Dyn 38: 1989–2000. https://doi.org/10.1007/s00382-011-1159-1.

Mannis P. 1988. CSIRO eprint matterilals,, retrieved 25 September 2023, from https://www.cmar.csiro.au/e-print/internal/manins_x1988c.pdf

Mohd Shafie, S.H., Mahmud, M., Mohamad, S. et al. 2022. Influence of urban air pollution on the population in the Klang Valley, Malaysia: a spatial approach. Ecol Process **11:** 3 (2022). Doi: https://doi.org/10.1186/s13717-021-00342-0

Morrison AL, Nelson PF. 2011. Quantification, speciation and morphology of respirable silica in the vicinity of open-cut coal mines in the Hunter Valley, NSW. ACARP project C18024 final report. Australian Coal Research Limited: NSW, Australia.

NPS. 2023. National Park Service, department of the Interior, USA. Retrieved 25 September 2023, from https://www.nps.gov/deva/learn/nature/airquality.htm.

NSWEPA. 2019. *2013 Calendar Year Air Emissions Inventory for the Greater Metropolitan Region in NSW*. NSW Environment Protection Authority Sydney, Australia. Retrieved 11 October 2023 from: https://www.epa.nsw.gov.au/your-environment/air/air-emissions-inventory/air-emissions-inventory-2013.

NSWEPA. 2023. Tackling coal mine dust. New South Wales Environment Protection Authority, Sydney, Australia. Retrieved 11 October 2023 from: https://www.epa.nsw.gov.au/your-environment/air/regional-air-quality/tackling-coal-mine-dust.

NEPC. 2021. National Environment Protection (Ambient Air Quality) Measure, Compilation No. 3. Australian National Environment Protection Council Retrieved June 22nd, 2022, from https:// www. legis lation. gov. au/ Detai ls/ F2021C00475.

North GR, Bell TL, Cahalan RF. 1982. Sampling errors in the estimation of empirical orthogonal functions. *Mon. Wea. Rev.* **110:** 699-706.

OEH. 2012. Upper Hunter Air Quality Monitoring Network – 2012 Annual Report, NSW Office of Environment and Heritage, Sydney, https://www.environment.nsw.gov.au/topics/air/air-publications.

OEH. 2014. Annual Air Quality Statement 2013. State of NSW Department of Planning and Environment. Retrieved 10 May 2023, from: https://www.environment.nsw.gov.au/research-and-publications/publications-search/air-quality-statement-nsw-2013.

OEH. 2017a. Upper Hunter Dust Risk Forecasting Scheme Development. Final report to the NSW Environment Protection Authority. ISBN 978-1-76039-995-5. NSW Office and Environment and Heritage: Sydney, Australia. Accessed 30 November 2023 at: https://www.environment.nsw.gov.au/-/media/OEH/Corporate-Site/Documents/Air/upper-hunter-dust-risk-forecasting-scheme-development-final-report-170707.pdf.

OEH. 2017b. *Better evidence, stronger networks, health communities*. Five-year review of the Upper Hunter Air Quality Monitoring Network. ISBN 978-1-76039-979-5. Office of Environment and Heritage: Sydney, Australia

Olguin CR, Everingham J. 2015. *ACARP C22029 Managing Cumulative Impacts in Mixed-Industry Regions - Upper Hunter Valley Case Study, New South Wales*. Centre for Social Responsibility in Mining Sustainable Minerals Institute, The University of Queensland, Australia.

Physick WL, Noonan JA, Manins PC. 1991. Air quality modelling study of the Hunter Valley. Phase 1: Emitters in the Upper Hunter. Electricity Commission of New South Wales: Sydney, Australia.

POEO Regulation. 2021. Protection of the Environment Operations (General) Regulation 2021. New South Wales Government: Sydney. Retrieved 11 October 2023 from: https://legislation.nsw.gov.au/view/whole/html/inforce/current/sl-2021-0486.

Quimbayo-Duarte J, Chemel C, Staquet C, Troude F, Arduini G. 2021. Drivers of severe air pollution events in a deep valley during wintertime: A case study from the Arve river valley, France, Atmospheric Environment 247: 118030, https://doi.org/10.1016/j.atmosenv.2020.118030.

Reisen F, Gillett R, Choi J, Fisher G, Torre P. 2017. Characteristics of an open-cut coal mine fire pollution event. Atmospheric Environment 151: 140-51. DOI: https://doi.org/10.1016/j.atmosenv.2016.12.015.

Richman MB. 1986. Rotation of principal components. J. Climatol. 6: 293–335.

Riley M, Kirkwood J, Jiang N, Ross G and Scorgie Y. 2020. Air quality monitoring in NSW: From long term trend monitoring to integrated urban services'. *Air Quality & Climate Change* 54(1): 44-51.

Rimbu N, Lohmann G, König-Langlo G, Necula C, Ionita M. 2013. Daily to intraseasonal oscillations at Antarctic research station Neumayer. Antarctic Science **26(02):** 193-204. https://doi.org/10.1017/S0954102013000540.

Rimbu N, Lohmann G, König-Langlo G, Necula C, Ionita M Monica. 2012. 30 years of synoptic observations from Neumayer Station with links to datasets. PANGAEA, https://doi.org/10.1594/PANGAEA.804156

Segersson D, Eneroth K, Gidhagen L, Johansson C, Omstedt G, Nylén AE, Forsberg B. 2017. Health Impact of PM_{10}, $PM_{2.5}$ and Black Carbon Exposure Due to Different Source Sectors in Stockholm, Gothenburg and Umea, Sweden. *Int J Environ Res Public Health* 14(7): 742. doi: 10.3390/ijerph14070742.

SPCC. 1982. Air pollution dispersion in the Hunter Valley. ISBN 0 7240 5856 7. New South Wales State Pollution Control Commission.

SPCC. 1983. Air pollution from coal mining and related developments. ISBN 0 7240 5936 9. New South Wales State Pollution Control Commission.

Torrence C, Compo GP. 1998. A Practical Guide to Wavelet Analysis. Bulletin of American Meteorological Society 79: 61-78.

Watt S, Rahman M, Jiang N. 2019. A meteorological analysis of elevated particle pollution during the November 2018 and February 2019 dust storms in New South Wales. Air Quality and Climate Change 53 (3): 9-12.

Weickmann K, Berry E. 2009. The tropical Madden-Julian oscillation and the global wind oscillation. Monthly Weather Review 137: 1601-1614. DOI: 10.1175/2008MWR2686.1

Zhang C. 2005. Madden-Julian Oscillation. Review of Geophysics 43: RG2003. doi:10.1029/2004RG000158.

Zhou W, Chan JCL. 2005. Intraseasonal oscillations and the South China Sea summer monsoon onset. International Journal of Climatology 25: 1585-1609. DOI: 10.1002/joc.1209.