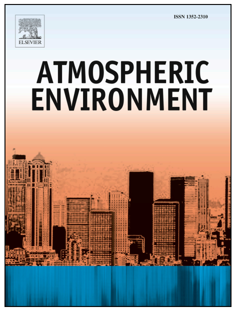
Journal Pre-proof



High PM_{10} concentrations in the city of Buenos Aires and their relationship with meteorological conditions

Andrea L. Pineda Rojas, Rafael Borge, Nicolás A. Mazzeo, Ramiro I. Saurral, Bruno N. Matarazzo, Jose M. Cordero, Emilio Kropff

PII: S1352-2310(20)30505-7

DOI: https://doi.org/10.1016/j.atmosenv.2020.117773

Reference: AEA 117773

To appear in: *Atmospheric Environment*

Received Date: 8 March 2020

Revised Date: 4 July 2020

Accepted Date: 9 July 2020

Please cite this article as: Pineda Rojas, A.L., Borge, R., Mazzeo, Nicolá.A., Saurral, R.I., Matarazzo, B.N., Cordero, J.M., Kropff, E., High PM_{10} concentrations in the city of Buenos Aires and their relationship with meteorological conditions, *Atmospheric Environment* (2020), doi: https://doi.org/10.1016/j.atmosenv.2020.117773.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2020 Published by Elsevier Ltd.

Andrea Pineda Rojas: Conceptualization, Methodology, Validation, Formal Analysis, Investigation, Resources, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization, Supervision, Project administration, Funding acquisition

Rafael Borge: Conceptualization, Methodology, Formal Analysis, Writing - Review & Editing, Visualization

Nicolás Mazzeo: Writing - Review & Editing, Visualization

Ramiro Saurral: Validation, Formal analysis, Writing - Review & Editing

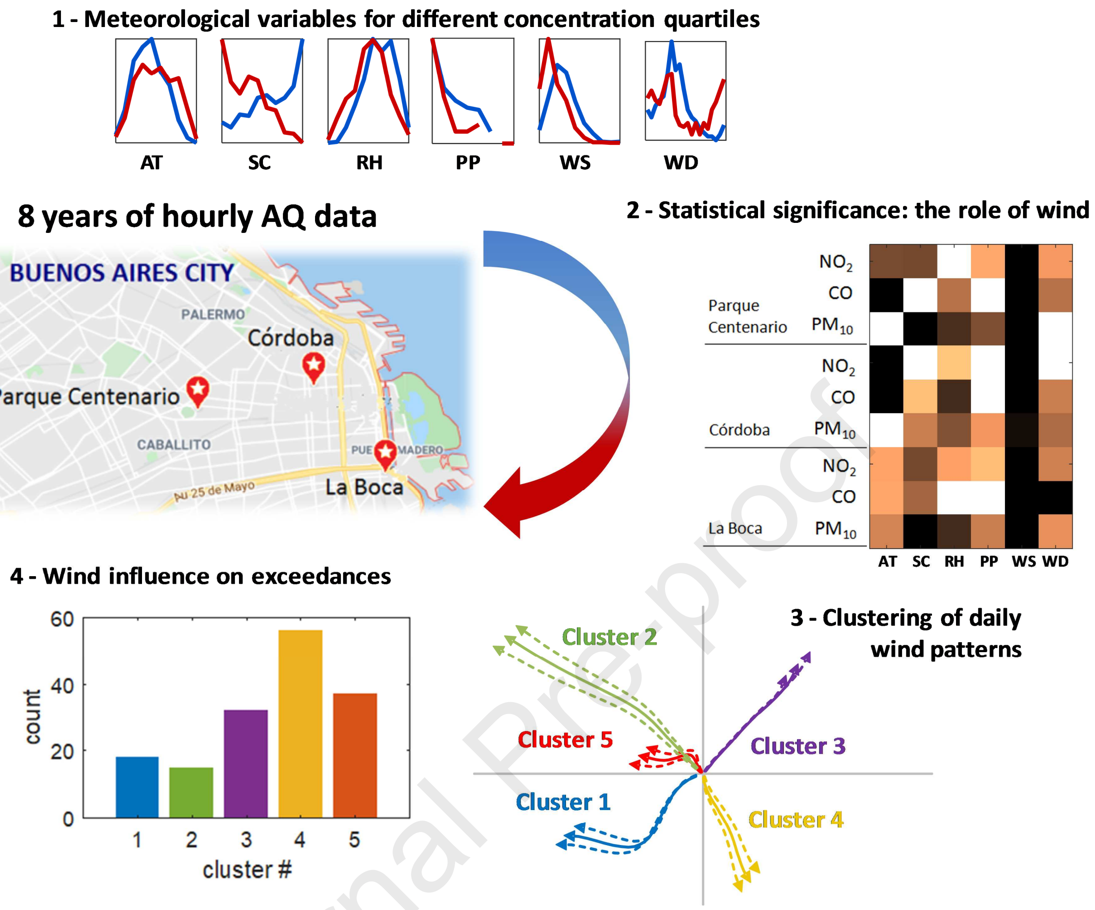
Bruno Matarazzo: Investigation, Visualization

Jose M. Cordero: Software, Visualization

Emilio Kropff: Methodology, Software, Validation, Formal Analysis, Investigation, Data Curation, Visualization, Writing - Review & Editing

Journal Pre-proof

Journal Pre-proof



Journal Pre-proof

Journal Pre-proof

HIGH PM_{10} CONCENTRATIONS IN THE CITY OF BUENOS AIRES AND THEIR RELATIONSHIP WITH METEOROLOGICAL CONDITIONS

Andrea L. Pineda Rojas [1]; Rafael Borge [2]; Nicolás A. Mazzeo [3]; Ramiro I. Saurral [1,4]; Bruno N. Matarazzo [4]; Jose M. Cordero [2], Emilio Kropff [5]

[1] Centro de Investigaciones del Mar y la Atmósfera, UMI-IFAECI/CNRS, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, CONICET, UBA, Buenos Aires, Argentina
[2] Laboratorio de Modelización Ambiental, Escuela Técnica Superior de Ingenieros Industriales, Universidad Politécnica de Madrid, España
[3] Departamento de Ingeniería Química, Facultad Regional Avellaneda, Universidad Tecnológica Nacional, CONICET, UTN, Buenos Aires, Argentina
[4] Departamento de Ciencias de la Atmósfera y los Océanos, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Buenos Aires, Argentina
[5] Fundación Instituto Leloir - IIBBA/CONICET, Buenos Aires, Argentina

Corresponding author: Andrea L. Pineda Rojas
E-mail: pineda@cima.fcen.uba.ar

Abstract

In this work, the first long-term (eight years) record of hourly concentrations of carbon monoxide (CO), nitrogen dioxide (NO_2) and particulate matter with diameter less than 10 µm (PM_{10}) from three sites in the city of Buenos Aires is analysed. Considering the short-term guidelines suggested by the WHO, the daily mean PM_{10} concentrations present a relatively large number of exceedances at the three sites. Different statistical techniques are combined to study the relationship between these relatively high PM_{10} concentrations and relevant surface meteorological variables. For all pollutants and sites, wind speed shows the largest differences between the lowest and highest concentration quartiles. To further explore its role on daily mean PM_{10} concentration, a k-means algorithm is applied, grouping days with similar surface 1h-wind sequences. Five wind sequence clusters are found, presenting distinctive air quality data features. Two clusters (1 and 2) show that PM_{10} exceedances occurring with winds entering the city from the river represent between 10-21% of total events at the three sites. The frequency of exceedance under these conditions decreases with the distance to the coast. For cluster 1, the hourly PM_{10} concentration profile and its associated daily wind sequence suggest an important contribution to exceedance events from the city's southernmost power plant. Two clusters (3 and 4), exhibiting continental winds, account for 49-59% of the exceedances and co-occur with relatively drier air conditions. The correlation between CO and PM_{10} for days belonging to cluster 3 supports the hypothesis of a potential remote or distributed source contribution with SW winds. For cluster 4, differences among sites in the number of events under NNW winds suggest an important contribution from the city's widest avenue to the PM_{10} levels at the most coastal site. A large contribution coming from urban sources is also indicated for these winds. Finally, cluster 5, exhibiting low wind speed sequences, accounts for 23-33% of the exceedances at the three sites. The average PM_{10} concentration increases with persistence of this cluster, which could be a driver for exceedances. These results contribute to show the importance of simple methods such as clustering analysis to obtain insights into air quality features such as exceedances and their potential drivers. They also suggest that further efforts in monitoring, modelling and emission estimates may help to better understand local, urban and regional source contributions to these events in the city of Buenos Aires.

Keywords: air quality data; Buenos Aires; exceedance conditions; meteorological data

1. INTRODUCTION

Air pollutants can cause adverse effects on human health and the environment if their concentrations are relatively large. According to the World Health Organisation (WHO), nine out of ten people worldwide breathe polluted air. The combined effects of poor outdoor and indoor air quality cause around seven million deaths annually (WHO, 2018). The largest negative impact of atmospheric pollution occurs in urban areas because they not only present the highest air pollutant levels but also the largest number of people exposed to them. Given the increasing evidence of negative effects on the health of the population (e.g., Lelieveld et al., 2019; Papadogeorgou et al., 2019; Sun and Zhu, 2019), more governments are taking action for which interaction with the scientific community may be crucial.

Although air quality improvement strategies rely on emission abatement measures, meteorological factors play a major role on high concentration values (Borge et al., 2018). Urban air quality is the result of a complex combination of different scale processes and local factors (Borge et al., 2016) that cannot be extrapolated across cities and must therefore be evaluated for each particular case. The study of the role of meteorological conditions on high concentration events requires long-term (i.e., several years) series of observations. Usually, air quality data availability is the main limitation for such analysis. Exceedances over a given threshold typically represent a small fraction of the data set (the higher the threshold the lower the fraction). A long-term record guarantees that all possible outcomes (resulting from the countless combinations of meteorological and emission conditions) are included in the analysis.

The city of Buenos Aires (CBA) is part of the Metropolitan Area of Buenos Aires (14,967,000 inhabitants, 3830 km^2), the third mega-city in Latin America (UN, 2019). Compared to other South American cities, the air quality and its relationship with meteorology in the CBA has been poorly studied. Most works (e.g., Bogo et al., 1999; Mazzeo et al., 2005; Arkouli et al., 2010) analyse air pollutant concentrations using observations from short-term (i.e., a few months) monitoring campaigns. A few studies using long-term series have focused on meteorological aspects of air pollution. For example, Venegas and Mazzeo (1999) analyse stagnation,

Journal Pre-proof

recirculation and ventilation conditions using two years of meteorological data from several sites of Argentina including Buenos Aires, and conclude that the frequency of occurrence of these conditions accounts for 10%, 6% and 40% of the time, respectively. Gassmann and Mazzeo (2000) study the air pollution potential using four years of observations and find that Buenos Aires has a frequency of days with low ventilation conditions of 8.5%. Finally, Mazzeo and Venegas (2004) analyse the persistence of different wind and stability conditions in the CBA using three years of meteorological data and highlight that while all wind directions present long persistence periods, some of them may lead to worse air pollution conditions.

At regulatory level, the environmental protection agency of the CBA (APRA, in Spanish) monitors ambient concentration levels of carbon monoxide (CO), nitrogen dioxide (NO_2) and particulate matter with an aerodynamic diameter less than or equal to 10 μm (PM_{10}) at three air quality stations, simultaneously since 2009. While current air quality is adequate to protect human health according to the present regulation (Res. 403/APRA/13), air quality standards in the CBA are expected to be updated in the coming years to follow the guidelines suggested by the World Health Organisation (APRA, personal communication). In this context, knowledge of the mechanisms responsible of high concentrations [i.e., those exceeding the WHO levels (WHO, 2005)] is fundamental. In addition, a comprehensive study of the relationship between air quality and meteorology may contribute to identify potential source regions of air pollutants.

In this work, we present a study of the first long-term air quality data set recorded in the CBA consisting of 8 years of 1-hour NO_2, CO and PM_{10} ambient concentration records measured by the APRA at the three monitoring sites. The methodology includes simple statistical analyses and clustering of daily wind sequences. The relationship between observed air quality features and meteorological conditions is discussed with a focus on the PM_{10} concentrations exceeding the daily mean WHO guideline. The objective is to gain knowledge on the underlying meteorological drivers of high air pollutant concentrations in the atmosphere of Buenos Aires that should be considered for air quality management in the city.

Air quality data are described in Section 2. The statistical analyses used four our research are presented in Section 3 and the results discussed in Section 4. A discussion on the advantages and limitations of the methodology is presented in Section 5 and the main conclusions of the work are summarised in Section 6.

2. AIR QUALITY DATA

Air quality data consist of eight years (01/05/2010 - 30/04/2018) of continuous hourly concentrations of CO, NO_2 and PM_{10} measured at three sites in the city of Buenos Aires (**Table 1**). These monitoring stations are regarded by the local environmental authority (APRA) as: urban background (CEN: Parque Centenario), urban traffic (COR: Córdoba) and urban industrial (LB: La Boca). Parque Centenario is located in the geographical centre of the city, in a residential-commercial area, 60 m away from a large park (see **Figure 1**). Córdoba station is located on a busy avenue through which 38,000 vehicles circulate daily in an east-west direction (Mazzeo and Venegas, 2012). This site is surrounded by buildings with varying heights (10-80 m) in a commercial area of the city. Finally, La Boca station is located near the coast (where three thermal power plants operate) and 160 m away from a highway.

Correlation between stations of hourly concentration values is greater than or equal to 0.5 for all pollutants and sites, except for CO between COR and LB sites (0.38) (see **Figure S.1**). Note that relatively high correlations within the city are expected because all stations are subject to similar influences of diurnal and seasonal variations in both the meteorological conditions and the emissions. However, correlations are not perfect, implying that local features still account for a substantial amount of the variation in concentration values. In general, at each site, the following correlation order is verified: $CO < NO_2 < PM_{10}$, as expected from the nature of these pollutants.

NO_2 and CO concentrations are larger in COR (**Figure 2**). As expected, in all sites, these species present diurnal variations which follow that of road traffic emissions with two maxima at peak hours and larger concentration during winter months mainly due to reduced dispersion conditions, as previously suggested in other studies (e.g., Venegas and Mazzeo, 1999; Mazzeo and Venegas, 2004). However, the amplitude of the mean monthly concentration values varies differently for the two pollutants. The maximum-to-minimum NO_2 concentration ratio is considerably larger at COR (104%) than at CEN (56%) and LB (70%); while the CO variation is rather similar across sites (74-85%). These differences must be related with local features.

PM_{10} presents hourly and monthly variations that are similar across sites and contrast with the traffic-related profile of the other species. On average, hourly PM_{10} concentration values show small variations during diurnal hours, except at LB where larger levels in the afternoon are observed, peaking at 18h. These characteristic small diurnal variations suggest a large contribution of sources other than traffic (secondary formation and/or non-local sources). The monthly variation of PM_{10} shows larger values in June, July and November, with a maximum-to-minimum PM_{10} concentration ratio varying between 26% (COR) and 45% (CEN). The relatively larger concentration values observed in warm months could be due to photochemical reactions leading to a larger fraction of secondary aerosols.

During weekdays, the concentration of the three species is larger as a result of greater human activities, mainly traffic. Despite similar temporal patterns, some differences among pollutants and sites are observed. At CEN, COR and LB, the weekend reductions relative to the weekday mean are respectively: 16%, 14% and 16% for NO_2 (ratio=1.14); 11%, 25% and 11% for CO (ratio=2.27); and 11%, 14% and 7% for PM_{10} (ratio=2).

3. METHODOLOGY

In order to explore the role of relevant meteorological variables on daily pollutant concentrations, simple statistical techniques and clustering analysis are combined. Due to the lack of meteorological observations at the air quality stations, in this work we assume that data from the meteorological station that is located in the local airport (see **Figure 1**) is representative of that at the three sites. These data consist of: air temperature (AT), sky cover (SC), relative humidity (RH), precipitation (PP), wind speed (WS) and wind direction (WD). The assumption of horizontally homogeneous ambient wind conditions in the area has been used satisfactory in air quality modelling studies (e.g., Pineda Rojas, 2014) and relies on the fact that the Metropolitan Area of Buenos Aires is located over flat terrain.

Days presenting two or more consecutive hours with missing data are discarded; otherwise linear interpolation is applied. Daily averages are computed for all variables, obtaining N=2347 days with complete (i.e., 24 hours) meteorological data and air quality data of each species in at least one site. Since wind is a circular variable, average values are obtained separately for its zonal and meridional components.

In order to identify the atmospheric variables that are mostly associated with low and high air pollutant levels, an analysis of their distributions by concentration quartiles for each species and site is first performed. For each meteorological variable, the statistical difference between the distributions of the 1st and 4th quartiles is assessed using the Mann-Whitney U-test. Significant effects indicate an influence of the meteorological variable on concentration levels for a given pollutant and site.

Since wind is the most important variable for dispersion of air pollutants, and its hourly variation (along with that of emissions) may contribute to the observed daily mean concentration values, a clustering analysis is then performed on the *daily 1h-wind sequence*. Clustering analysis is among the methods most widely used to study the relationship between air pollution and meteorological conditions using long-term series when a reduction of data dimensionality is required (e.g., Davies et al., 1998; Beaver and Palazoglu, 2006; Borge et al., 2007; Beaver et al., 2008; Rimetz-Planchon et al., 2008; Pakalapati et al., 2009; Khedairia and Khadir, 2012). This methodology allows grouping elements taken from a high dimensionality data set based on their similarity. In the case of daily 1h-wind sequence, each day is described by 48 variables (= 2 wind components x 24 hours). A k-means algorithm is applied to group them according to their wind sequences. The elbow method (Kaufman and Rousseeuw, 2009) is used to select the optimum number of clusters (k). It relies on the analysis of within-cluster sum of squares (WCSS) representing the sum of the squared distance between each member of the cluster and its centroid. WCSS is calculated for different numbers of cluster starting at k = 2 and plotted against k. According to this method, an optimum k value (i.e., that providing an optimum trade-off between a significant cluster separation and a manageable number of groups) is given by the inflexion point in the curve. In order to get a more complete description of the wind fields in the area that are associated with the clusters, the daily mean sea level pressure (SLP) field is computed for each cluster using data taken from the NCEP-NCAR II Reanalysis dataset (Kanamitsu et al., 2002).

Once days are classified according to the method described above, the statistical difference of air pollutant concentrations among clusters is assessed for each pollutant and site in order to

determine whether different wind patterns are associated with different air pollutant levels. For a given cluster, the centroid represents the wind dynamics during a stereotypical day belonging to it, and interactions with mean hourly variations of concentration values can be analysed. In addition, as highlighted by other authors (e.g., Carslaw and Beevers, 2013), further insights on the air quality data may be obtained through the evaluation of inter- and intra- cluster variability.
In order to elucidate the meteorological conditions associated with PM_{10} events, we analyse:
- differences among clusters in: i) the mean diurnal species concentration profiles, ii) the variation of mean pollutant concentration with the cluster persistence period, and iii) the PM_{10} exceedances over the WHO guideline at each site,
- within-cluster differences among exceedance and non-exceedance days of: iv) relevant meteorological variables and v) correlations between PM_{10} and CO daily mean concentrations.

While analyses i) to iii) are performed for each site, observations from all sites are pooled together in the analyses iv) and v) to gain a comprehensive view of the exceedances at city level and to increase the statistical power given the reduced overall number of events.

4. RESULTS

4.1. Local meteorological conditions associated with different air pollutant levels

In order to get a general picture of the relationship between air pollutant concentrations measured at each site and relevant surface meteorological variables recorded at AEP station, the distributions of the meteorological variables corresponding to each daily mean concentration quartile are analysed. As shown in **Figure 3**, NO_2 and CO upper quartile concentrations occur more frequently with lower daily mean air temperature (AT) values compared with their lower quartile levels. While the distributions of AT for the two extreme concentration quartiles show significant differences, those of sky cover (SC), relative humidity (RH) and precipitation (PP) present a more complex behaviour. For example, larger CO levels associated with larger RH values are only evident at CEN and COR stations. In the case of PM_{10}, differences in AT are not robust and those in SC and RH show a clear inverse relationship with the daily mean concentrations. Larger PM_{10} concentrations, occurring more frequently with lower SC (clearer sky conditions) and RH values, could be related to the role of precipitation. At the three sites, the impact of PP is stronger on PM_{10} than on NO_2 and almost no effect on CO concentrations is observed (**Figure 4**), which is expected due to the different efficiency of rain to remove these pollutants from the atmosphere (Yoo et al., 2014).
The three species show larger concentration values with lower values of wind speed (WS), and this variable is the one showing the most significant statistical difference between extreme quartile concentrations, as shown by its low p-value compared to other meteorological variables (**Figure 4**). The distributions of wind direction (WD) also show some similarities among pollutants and sites: in general, lower concentrations occur mostly with winds from the 1st and 2nd quadrants (i.e., winds entering the city from the river) and larger differences between the upper

and lower concentration quartiles (i.e., the difference between the red and blue curves in **Figure 3**) occur more frequently with winds from the 4th quadrant. An exception to this typical behaviour is observed in the WD distributions for CO and PM_{10} in COR which show greater upper - lower differences with S winds. This can be attributed to the street canyon effect that has been previously reported for this monitoring site (e.g., Mazzeo and Venegas, 2012; Venegas et al., 2014) and is expected due to the street having an east-west orientation and the monitoring station being located on its south side. NO_2 concentrations at traffic hotspots can be affected by a variety of combined micro-scale phenomena (e.g., Sanchez et al., 2017).The fact that these concentrations are less affected by the street canyon effect suggests that at COR they may be more influenced by urban background levels.

Considering the important role of wind on air pollutant concentrations and the fact that daily averages may prevent from understanding important features linked to hourly variations, the relationship between concentrations and surface wind sequences is investigated in the following sections.

4.2. Surface wind sequence patterns and associated synoptic pressure fields

As described in Section 3, clustering analysis is performed to find groups of similar daily sequences of hourly wind during the analysed period. A number of k (number of clusters) = 5 is identified as an adequate value for the k-means algorithm applied (see **Figure S.2**). **Figure 5** presents the accumulated hourly mean wind vector obtained for each of the five clusters along with the corresponding 95% confidence interval along the day. Clusters are ordered clockwise starting at the 1st quadrant (cluster 1) and ending at sequences including the lowest wind speeds (cluster 5). The first two clusters present larger hourly mean wind speeds. Cluster 1 shows uniform wind direction from NNE during night and early morning hours and with wind rotating clockwise around 15 h so that ESE direction dominates during the afternoon. Clusters 2, 3 and 4 show nearly constant wind directions from SE, SW and NNW, respectively. Finally, cluster 5 includes winds with the lowest mean speeds along the day. The direction is SSE at night, it rotates to ENE at around 12 h and to ESE at 18 h. Examples of how this wind sequence classification performs at individual days are shown in **Figure S.3**.

In order to describe these clusters at synoptic scale, **Figure 6** displays the horizontal fields of the mean sea level pressure (SLP) averaged for each cluster using daily SLP data from the NCEP-NCAR II Reanalysis. The SLP pattern corresponding to cluster 1 (**Figure 6.a**) is characterised by a well-defined high pressure system over the South Atlantic waters, east of the study area, and winds prevailing mostly from the NE and ENE. A high pressure system is also present in cluster 2 (**Figure 6.b**), although its strength is weaker compared to cluster 1 and its location is displaced onto the continent, slightly to the south of the study region. In this scenario, prevailing winds come mainly from the E and SE. In turn, the SLP patterns corresponding to clusters 3 and 4 are associated with more dynamical conditions: cold fronts propagating over the study area (Escobar and Bischoff, 1999). In these two scenarios, surface winds have a continental origin and blow from the SW

(NW) in the case of cluster 3 (cluster 4). Both cases may lead to the uplift of fine soil particles, especially under dry soil conditions, providing a natural source of particulate matter (from the regions that are highlighted in the Figure) which may eventually reach the city of Buenos Aires and increase the concentration of PM_{10}. Lastly, the pattern associated with cluster 5 (**Figure 6.e**) is characterised by a high pressure system located over the city of Buenos Aires leading to light and variable winds over the study area. Overall, the obtained SLP fields are consistent with the surface wind sequences shown in **Figure 5** and provide an insight on the potential source regions and stagnation conditions that may contribute to increased pollution levels in the city. Still, it should be taken into account that hourly variations in the wind components (**Figure 5**) are expected on top of the synoptic, daily-mean wind direction, since the location of the wind measuring site close to the river shore makes it much more prone to be affected by the diurnal cycle of the river breeze.

4.3. Variation of air quality with surface wind sequence patterns

Once individual days have been distributed among the five wind sequence patterns described in the previous section, differences in the air pollutant concentrations among these clusters must be investigated in order to determine whether the classification can be useful to further explore the role of meteorology on air quality at the three sites. As shown in **Figure 7**, the concentrations of the three chemical species are larger (in terms of their means) in clusters 4 and 5 at CEN and LB stations, while at COR (the street canyon site) a less clear pattern is observed: NO_2 is larger in clusters 3 to 5, and CO and PM_{10} levels are higher in clusters 3 and 5. From this Figure, it can be concluded that average pollutant concentrations are relatively lower during surface winds conditions corresponding to clusters 1 and 2, and relatively higher for clusters 4 and 5, except at COR where cluster 3 is also associated with high concentration values.

It is interesting to note that differences among clusters are present not only in the concentration levels but also on the profiles of variation along the day (see **Figure 8**). For NO_2 and CO, for example, the concentration profile of cluster 1 presents a second daily peak considerably larger than the first one (except at LB), that of cluster 2 shows comparatively lower diurnal variation and those of clusters 3, 4 and 5 look like those of typical traffic emissions (see for example the vehicle emissions profile at COR presented by Venegas et al., 2014). In turn, the hourly variation of the PM_{10} concentration presents less smoothed profiles with no clear order in the cluster curves (they cross each other at different times of the day). At the three stations, in the night and during the morning, cluster 1 has the lowest concentrations and cluster 5 the largest one. Shortly after 6 h, PM_{10} concentrations of cluster 1 start to increase, reaching values comparable to those of cluster 5 towards 18 h. At LB, a clear contribution of cluster 1 to the highest 18h-peak (previously noticed in **Figure 2**) is observed. Other clusters present only small diurnal variations, while only cluster 5 at LB shows a profile similar to the emissions from the road transport.

Since persistence in wind conditions has been suggested to affect AQ in the CBA (Mazzeo and Venegas, 2004), the role of cluster persistence on average air pollutant concentrations was also

Journal Pre-proof

explored. The daily variation of the cluster number at each site (with a brief description) is presented in **Figure S.4**. Some common features are observed in the daily mean NO_2 and CO concentrations: they show significant decreases with persistence of cluster 2 at CEN, and increases with persistence of cluster 5 at COR (see **Figure 9** and **Table 2**). NO_2 also decreases with persistence of cluster 1 at all sites, while CO increases with persistence of cluster 5 at CEN. Daily mean concentrations of PM_{10} at the three sites decrease (increase) with persistence of clusters 2 (cluster 5), with the strongest statistical trend found for cluster 2. At LB, the most coastal site, an increase of the average PM_{10} concentrations with persistence of clusters 1 and 4 is also observed.

4.4. PM_{10} exceedances over the WHO guideline

While a few exceedances of the local air quality standard occur in the 8-year period under study (with frequencies ≤ 0.1, not shown), the WHO guideline (50 μg/m^3) is exceeded much more often, with frequencies varying between 5.8% in CEN and 8.2% in LB (computed over the total analysed days at each site), totalling 368 PM_{10} exceedances in 229 days (**Table 3**). 57% of these days presents exceedances at only one monitoring site, 25% at two sites and 18% at the three sites simultaneously (see **Figure 10**). The average daily concentration when the WHO standard is exceeded varies between 68-69 μg/m^3 across sites. When looking at the distribution of the PM_{10} exceedances by cluster (**Figure 11**), the largest number of exceedances at CEN and LB are related to cluster 4 (winds from the 4th quadrant), followed by cluster 5 (low wind speeds) and cluster 3 (winds from the 3rd quadrant), while a larger exceedance frequency for cluster 3 is found at COR. In total, 90%, 82% and 79% of the events occur in clusters 3 to 5 at CEN, COR and LB, respectively. Only a few exceedances occur in clusters 1 and 2, except at LB where these clusters amount for 21% of the events.

While the occurrence of clusters is quite homogeneous in the whole period (see upper panel of **Figure 11**), the relative frequency has considerable seasonal variations. **Figure 12** shows a clear dominance of clusters 1 and 2 during austral spring and summer (Sep-Feb) that could be explained by the fact that during the warm season, the climatological position of the Atlantic Subtropical High pressure system favours the occurrence of easterly winds over CBA. Clusters 3 and 4 dominate during the cold season (Jun-Aug) when the passage of cold fronts is much more frequent. Cluster 5 dominates in May but also shows relative peaks in February and August, which can be explained by the fact that higher pressures over CBA occur in both the cold and warm seasons. As shown in **Figure 12**, PM_{10} exceedances at CEN and COR occur more frequently during cold months, mainly related to clusters 3 to 5. In turn, at LB, the number of exceedance days is more homogeneously distributed both along the year and among the five clusters, except in June and July when a larger contribution from clusters 3 and 4 is observed. The monthly mean PM_{10} exceedance frequencies (computed over the monthly number of days with complete data) vary between 0.7-13.3% (CEN), 1.3-14.5% (COR) and 3.1-17.0% (LB).

Journal Pre-proof

Within each cluster, exceedances occur during a very small fraction of days, as evidenced by comparing the lower and upper panels of **Figure 11**. This suggests that, within a given cluster, variables other than the mean surface wind sequence may contribute to explain the occurrence of these events.

4.5. Within-cluster variation of other meteorological variables

This section analyses differences between meteorological variables occurring in exceedance vs. non-exceedance days for each cluster. Interestingly, exceedance days co-occur with relatively larger air temperature values under wind conditions of cluster 1 and the opposite is observed for clusters 4 and 5 (see **Figure 13**). This feature of cluster 1 suggests a potential contribution of secondary sources when the wind comes from the river and is not observed if the wind classification is not considered (not shown). Differences in the mean relative humidity are only evident in cluster 3 and 4; while those of wind speed are statistically significant in clusters 1, 3 and 4. Despite significant differences in SC and PP values between the lower and upper PM_{10} concentration quartiles (**Figure 4**), differences in their average values among exceedance and non-exceedance days are not significant under this wind classification. This could be due to the relatively low number of events.

The synoptic pressure fields (**Figure 6**) associated with SW winds in cluster 3 and NW winds in cluster 4 suggest that these continental winds are efficient in providing a natural source of particulate matter entering the city of Buenos Aires. In order to check this hypothesis, the correlation between daily mean concentrations of PM_{10} and CO during exceedance and non-exceedance days, are computed. As shown in **Table 4**, correlation for cluster 3 not only decreases but also becomes not-significant in exceedance days. This supports the hypothesis of an important contribution from remote emission sources under SW winds conditions. In turn, the correlation analysis does not support the hypothesis for cluster 4 because the correlation is higher during these PM_{10} events. Note that the relatively higher average concentrations of NO_2 and CO at CEN and LB for this cluster (**Figure 7**) suggest a larger contribution of urban pollution under NW winds, which could also be the case for PM_{10}.

5. DISCUSSION

Advantages of the methodology

In this work, different statistical techniques are combined to provide novel insights into the air quality (AQ) in the CBA. The rationale behind their choice is related to specific significant information provided at each step. First, meteorological variables with a high influence on daily concentration of pollutants are identified. To do this, the data distributions of each meteorological variable corresponding to the first and last pollutant concentration quartiles are compared. While significantly different distributions for several meteorological variables are found, the daily mean wind speed is identified as the most influential factor for all pollutants and sites. Next, to further understand the role of wind in determining daily pollutant concentration values and exceedances, the impact of different surface hourly wind sequences on daily concentration values is studied. A k-means algorithm is applied to group days having similar 1h-wind sequences with the aim of assessing potential differences in AQ data among these groups. The discriminating power of the clustering analysis comes from its capacity to classify multi-dimensional datasets, making no a priori assumptions as to how the data are distributed. The method proves itself useful if differences among the clusters are found in AQ variables, as exemplified in Carslaw and Beevers (2013). This is also the case for the data analysed here.

The main result of this work is that different typical hourly wind sequences, grouped by the k-means algorithm, are associated with distinct AQ features, as shown in figures 7 to 9 and 11 to 13. As an example of the advantage of the clustering approach applied to wind sequences, **Figure 14** shows that the daily mean PM_{10} concentration profile during exceedance days in cluster 1 presents strong hourly variations, which can be associated with the wind rotation characterising the cluster (**Figure 5**). In this case, a daily mean concentration above 50 $\mu g/m^3$ results from very large values (up to 140 $\mu g/m^3$) during the evening rather than moderately high but constant hourly concentration values along the day (**Figure 14**), as also suggested by examples of individual days belonging to the cluster. The interpretation of the mechanisms behind this particular class of exceedance would be obscured in a daily mean value analysis.

Cluster differences among sites highlight local features

Assuming that the wind measured at the local meteorological station may be representative for the whole city, cross-site differences in PM_{10} levels and in the number of daily exceedances for a given cluster are expected to result from differences in local source contributions. This simple reasoning allows us to draw some interesting conclusions regarding the impact of specific emission sources at different monitoring sites. First, the number of exceedances in clusters 1 and 2 decrease in the order: LB $\rightarrow$ COR $\rightarrow$ CEN (**Figure 11**). The location of the stations (**Figure 1**) and the dominant wind directions in these clusters (**Figure 5**) suggest an important contribution of emissions from the city's three power plants, located along the coastline. However, the clearly defined PM_{10} diurnal variation (**Figure 14**) occurring with the wind rotating conditions of cluster 1 (**Figure 5**) suggests that these exceedances are dominated by an emission source located at ESE of the city, highlighting the potential impact on AQ of the southernmost power plant (Costanera). Second, while COR and LB are approximately aligned along cluster 4's dominant wind direction (NNW), thus expected to receive the same air parcel from distant sources, they differ significantly in the number of PM_{10} exceedances on days belonging to this cluster (**Figure 11**). This difference can only be explained by the presence of an important PM_{10} emission source between the two sites. Arguably this source may be related to traffic emissions at Av. 9 de Julio (see **Figure 1**), regarded as the widest avenue in the world accommodating 14 car and bus lanes. Third, the number of exceedances under wind conditions of clusters 3 and 5 are quite homogeneous among sites, suggesting similar source contributions that are either distant or spatially distributed relative to the specific locations of the monitoring sites.

More and better data are essential

Several statistical tests in this work suggest significant effects when considering the complete 8-year dataset but non-significant effects when analysing exceedance days alone. This brings up the issue, often overlooked in the literature, of which of these two datasets best describes urban AQ. This is the case, for example, for the difference in PP between the two extreme PM_{10} quartiles (**Figure 4**), or for correlations between daily mean PM_{10} and CO concentrations (**Table 4**). A strong possibility behind the loss of significance is the simple fact that there is less data when only exceedences are considered. Note that lack of significance does not provide positive evidence in favour of the null hypothesis, but rather states that the null hypothesis cannot be rejected with the present data. This also highlights the need for larger records of AQ data, to be able to analyse exceedance events and the mechanisms behind them with enough statistical power, especially if data are further divided into smaller groups representative of specific conditions (as in the clustering analysis).

The 2010-2018 data suggest that the total number of exceedances over the WHO suggested guideline for daily PM_{10} concentration (50 $\mu g/m^3$) is decreasing over time at CEN and COR. In contrast, exceedances at LB seem to increase since 2014 (**Figure S.4**). This calls for further efforts in monitoring and analysing particulate matter levels and their main drivers around this site. Understanding the role of specific emission sources on PM_{10} concentrations requires detailed knowledge about emission sources in the whole Metropolitan area of Buenos Aires, currently not available. In order to correctly account for the local contribution (using air quality models) and confirm the hypothesis of potentially important contributions from specific sources, in-situ wind measurements at the air quality monitoring sites are also needed.

6. CONCLUSIONS

In this work, a relatively large number of daily mean PM_{10} concentrations exceeding the World Health Organisation (WHO) guideline (50 $\mu g/m^3$) is found at the three air quality (AQ) sites in the city of Buenos Aires (CEN: 110, COR: 100 and LB: 158) in a 8-year period (2010-2018). Simple quartile and clustering analyses of the three pollutants (NO_2, CO and PM_{10}) measured at the sites and relevant meteorological variables measured at the AEP station, are combined to study the main drivers of these events.

The analysis of meteorological values corresponding to each pollutant concentration quartile shows different relationships for two groups of pollutants. On one hand, NO_2 and CO (local pollutants) present larger daily mean concentrations most frequently with lower temperatures (which is probably related to the lower dispersion capacity of the atmosphere during winter). On the other hand, PM_{10} most frequently exhibits higher concentration levels with lower sky cover and relative humidity, highlighting the role played by wet removal. The statistical comparison of the extreme quartile distributions confirms that the impact of precipitation is stronger on PM_{10} than on NO_2, and that wind speed is the variable showing largest differences between low and high concentration distributions for all pollutants and sites.

In order to better understand the role of wind on the daily mean pollutant concentrations, a k-means algorithm is applied to find groups of days having similar surface 1h-wind sequences. Five wind sequence patterns (clusters) are obtained and significant differences in the AQ among them are found at the three sites. From the combined analysis of: a) PM_{10} event frequency distribution by cluster, b) synoptic pressure fields associated with the clusters, c) cluster profiles of pollutant hourly concentrations, d) pollutant levels vs cluster persistence period, and e) intra-cluster differences of relevant variables between exceedance and non-exceedance days, we arrive to the following conclusions:

- Two clusters (1 and 2) represent winds entering the city from the river (i.e., bringing "clean air") that are associated with a breeze-type circulation. Exceedances under these conditions account for 10-21% of total events at each site. Those presenting with winds from the 1st quadrant (cluster 1) co-occur with relatively lower wind speeds and higher temperatures (compared with similar wind conditions during non-exceedance days) which suggests a potential contribution of photochemical formation. The decrease in the number of exceedances with the distance to the coast and the strong diurnal profiles of both PM_{10} concentrations and wind during exceedance days, suggest an important contribution from the southernmost power plant.
- Two clusters (3 and 4) represent continental winds from the SW and NW sectors which are associated with cold fronts. Exceedances in these cases represent 49-59% of total events and co-occur with lower relative humidity values. SW winds (cluster 3) are suggested to provide a natural source of PM_{10} for the city of Buenos Aires; while a larger contribution from urban sources to the concentrations of the three pollutants is suggested under NW wind conditions (cluster 4). In this last case, a large difference in the number of exceedances between COR and LB highlights an important contribution from Av. 9 de Julio to the PM_{10} events at LB.
- The fifth cluster (5) represents days with low wind speeds and calm conditions. PM_{10} exceedances in these situations account for 23-33% of total events at the three sites, and they co-occur with relatively lower air temperatures and wind speeds compared with non-exceedance days.
- At the three sites, the average PM_{10} concentrations are positively correlated with persistence of cluster 5 and negatively correlated with that of cluster 2. A positive correlation is also observed in cluster 1 (breeze-prone conditions) at the most coastal site. These relationships are not statistically significant when only exceedance days are considered. Larger records of AQ measurements will help to determine the reason, as discussed in the previous section.

Further research, for example utilizing air quality deterministic modelling tools, may contribute to confirm the role of different emission sources within the Metropolitan area of Buenos Aires. To do this, both a high resolution emissions inventory and further monitoring efforts (including a larger number of air quality monitoring sites and in-situ wind measurements) are necessary.

Acknowledgement

The authors are grateful with the 'Agencia de Protección Ambiental (APRA)' of the city of Buenos Aires and the 'Servicio Meteorológico Nacional (SMN)' for providing the air quality and meteorological data, respectively. This work has been funded by the Project PICT2015-1676 from the 'Agencial Nacional de Promoción Científica y Tecnológica (ANPCyT)'.

References

Arkouli M., Ulke A. G., Endlicher W., Baumbach G., Schultz E., Vogt U., Müller M., Dawidowski L., Faggi A., Wolf-Benning U. &Scheffknecht G., 2010. Distribution and temporal behaviour of particulate matter over the urban area of Buenos Aires. Atmospheric Pollution Research, 1(1), 1-8.

Beaver S., Palazoglu A. 2006. Cluster analysis of hourly wind measurements to reveal synoptic regimes affecting air quality. J. Appl. Meteorol. Clim. 45, 1710-1726.

Beaver S., Palazoglu A., Tanrikulu S. 2008. Cluster sequencing to analyze synoptic transitions affecting regional ozone. J. Appl. Meteorol. Clim. 47, 901-916.

Bogo H., Negri R. M., San Román E. 1999. Continuous measurement of gaseous pollutants in Buenos Aires City. Atmos. Environ. 33, 2587-2598.

Borge R., Begoña Artíñano B., Yagüe C., Gomez-Moreno F.J., Saiz-Lopez A., Sastre M., Narros A., García-Nieto D., Benavent N., Maqueda G., Barreiro G., de Andrés J.M., Cristóbal A. 2018. Application of a short term air quality action plan in Madrid (Spain) under a high-pollution episode - Part I: Diagnostic and analysis from observations. Sci. Tot. Environ. 635, 1561-1573.

Borge R., Narros A., Artínano B., Yagüe C., et al. 2016. Assessment of micro- scale spatio-temporal variation of air pollution at an urban hotspot in Madrid (Spain) through an extensive field campaign. Atmos. Environ. 140, 432-445.

Borge R., Vardoulakis S., Lumbreras J., Rodríguez M.E., 2007. Analysis of long-range transport influences on urban PM10 using two-stage atmospheric trajectory clusters. Atmos. Environ. 41, 4434-4450.

Carslaw D.C., Beevers S.D. 2013 Characterising and understanding emission sources using bivariate polar plots and k-means clustering. Environmental Modelling & Software 40, 325-329.

Davies J.M., Eder B.K., Nychka D., Yang Q. 1998. Modeling the effects of meteorology on ozone in Houston using cluster analysis and generalized additive models. Atmos. Environ. 32(14/15), 2505–2520.

Escobar G., Bischoff S. 1999. Meteorological situations associated with significant temperature falls in Buenos Aires: an application to the daily consumption of residential natural gas. Meteorol. Appl. 6, 253-260.

Gassmann, M. I., Mazzeo, N. A., 2000. Air pollution potential: Regional study in Argentina. Environmental Management, 25(4), 375-382.

Kanamitsu M., Ebisuzaki W., Woollen J., Yang S.K., Fiorino M., Potter G.L. 2002. NCEP-DOE AMIP-II Reanalysis (R-2). Bull. Amer. Meteor. Soc., 83, 1631-1644.

Kaufman L., Rousseeuw P.J., 2009. Finding groups in data: an introduction to cluster analysis, John Wiley &Sons. ISBN: 9780471878766. DOI:10.1002/9780470316801.

Khedairia S., Khadir M.T. 2012. Impact of clustered meteorological parameters on air pollutants concentrations in the region of Annaba, Algeria. Atmos. Res. 113, 89–101.

Lelieveld J., Klingmüller K., Pozzer A., Pöschl U., Fnais M., Daiber A., Münzel T. 2019. Cardiovascular disease burden from ambient air pollution in Europe reassessed using novel hazard ratio functions. European Heart Journal 40, 1590–1596.

Mazzeo N.A., Venegas L.E. 2004. Some aspects of air pollution in Buenos Aires city. Int. J. Environ. Pollut. 22(4), 365-378.

Mazzeo N.A., Venegas L.E. 2012.Hourly NO_x concentrations and wind direction in the vicinity of a street intersection. International Journal of Environment and Pollution 48 (1-4), 96-104.

Mazzeo N.A., Venegas L.E., Choren H. 2005. Analysis of NO, NO_2, O_3 and NO_x concentrations measured at a green area of Buenos Aires City during wintertime. Atmos. Environ. 39, 3055-3068.

Pakalapati S., Beaver S., Romagnoli J.A., Palazoglu A. 2009. Sequencing diurnal air flow patterns for ozone exposure assessment around Houston, Texas. Atmos. Environ. 43, 715–723.

Papadogeorgou G., Kioumourtzoglou M-A., Braun D., Zanobetti A. 2019. Low Levels of Air Pollution and Health: Effect Estimates, Methodological Challenges, and Future Directions. Current Environmental Health Reports https://doi.org/10.1007/s40572-019-00235-7

Pineda Rojas A.L. 2014. Simple atmospheric dispersion model to estimate hourly ground-level nitrogen dioxide and ozone concentrations at urban scale. Environ. Model. Softw. 59, 127-134.

Rimetz-Planchon J., Perdrix E., Sobanska S., Bremard C. 2008. PM_{10} air quality variations in an urbanized and industrialized harbor. Atmos. Environ. 42, 7274–7283.

Sanchez B., Santiago J.L., Martilli A., Martin F., Borge R., Quaassdorff C., de la Paz D. 2017. Modelling NO_x concentrations through CFD-RANS in an urban hotspot using high resolution traffic emissions and meteorology from a mesoscale model. Atmos. Environ. 163, 155-165.

Sun Z., Zhu D. 2019. Exposure to outdoor air pollution and its human health outcomes: A scoping review. PLoS ONE 14(5): e0216550.

UN. 2019. World Urbanization Prospects: The 2018 Revision. Department of Economic and Social Affairs, Population Division (ST/ESA/SER.A/420), United Nations, 126 pp.

Venegas L.E., Mazzeo N.A. 1999. Atmospheric stagnation, recirculation and ventilation potential of several sites in Argentina. Atmospheric Research 52. 43–57.

Venegas L.E., Mazzeo N.A., Dezzutti M.C., 2014. A simple model for calculating air pollution within street canyons. Atmos. Environ. 87, 77–86.

Venegas L.E., Mazzeo N.A., Pineda Rojas A.L. 2011. Chapter 14: Evaluation of an emission inventory and air pollution in the Metropolitan Area of Buenos Aires. En: D. Popovic (ed.) Air Quality-Models and applications, Editorial In-Tech, 261-288.

WHO. 2005. Air quality guidelines. Global update 2005. Particulate matter, ozone, nitrogen dioxide and sulfur dioxide. World Health Organisation, 496 pp.

Journal Pre-proof

WHO. 2018. Ambient (outdoor) air pollution: https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health.
Yoo J-M., Lee Y-R., Kim D., Jeong M-J., Stockwell W.R., Kundu P.K., Oha S-M., Shin D-B., Lee S-J. 2014. New indices for wet scavenging of air pollutants (O_3, CO, NO_2, SO_2, and PM_{10}) by summertime rain. Atmos. Environ. 82, 226-237.

Journal Pre-proof

Journal Pre-proof

TABLES

Table 1: Basic statistics obtained from hourly time series of NO_2 (ppb), CO (ppm) and PM_{10} ($\mu g/m^3$) concentrations at each monitoring site (CEN: Parque Centenario, COR: Córdoba, LB: La Boca) during the analysed period (2010-2018). [The number of observations (N) and data availability (%) at each site is indicated]

	NO_2 (ppb)			CO (ppm)			PM_{10} ($\mu g/m^3$)		
	CEN	COR	LB	CEN	COR	LB	CEN	COR	LB
Hourly values									
Min	0	0	1	0.05	0.05	0.05	4	4	4
Median	18	23	16	0.50	0.50	0.26	24	24	24
Mean	19	26	19	0.56	0.61	0.34	30	30	31
Max	112	190	129	6.80	5.21	8.76	492	744	880
N	58499	46674	57131	51669	54604	50831	56245	52284	60915
%	83	67	81	74	78	72	80	75	87
Daily values									
Min	3	2	2	0.16	0.08	0.07	9	9	7
Median	18	22	16	0.52	0.6	0.34	25	26	27
Mean	18	25	18	0.56	0.64	0.39	29	29	30
Max	54	102	75	1.85	2.59	2.9	280	174	243
N	1809	1497	1800	1726	1587	1332	1882	1657	1936
%	62	51	62	59	54	46	64	57	66

Journal Pre-proof

Table 2: Pearson correlation coefficient between the average concentration and the cluster persistence period, for each pollutant and site. Positive (negative) correlation values indicate positive (negative) trends in the curves of **Figure 9**. Bold numbers indicate that the correlation is significant (i.e., the p-value obtained from the Kendall test is lower than 0.05)

	Cluster				
	1	2	3	4	5
NO_2					
CEN	**-0.13**	**-0.13**	0.09	**0.09**	0.06
COR	**-0.12**	-0.04	0.02	0.08	**0.10**
LB	**-0.10**	-0.07	-0.01	0.07	0.02
CO					
CEN	-0.06	**-0.18**	0.02	**0.09**	**0.20**
COR	0.02	0.02	-0.05	0.04	**0.12**
LB	**0.11**	-0.07	-0.08	0.05	0.01
PM_{10}					
CEN	-0.03	**-0.22**	0.00	0.05	**0.08**
COR	0.05	**-0.21**	-0.03	0.06	**0.10**
LB	**0.10**	**-0.20**	0.02	0.07	**0.09**

Table 3: Number and relative frequency (%) of daily mean PM_{10} concentration exceedances over the guideline suggested by the WHO (50 µg/m^3) at each monitoring site, in the studied period. [N: number of complete days (with 24 hours of data)]

Monitoring site	N	Number of events	Frequency (%)
CEN	1882	110	5.8
COR	1657	100	6.0
LB	1936	158	8.2

Table 4: Pearson correlation coefficient between PM_{10} and CO daily mean concentrations, for each cluster, stratifying by WHO guideline non-exceedance (Non-EXC) and exceedance (EXC) days. Bold numbers indicate that the correlation is statistically significant ($p < 0.05$)

Days	Cluster					
	1	2	3	4	5	All
Non-EXC	**0.15**	**0.19**	**0.33**	**0.25**	**0.28**	**0.29**
EXC	0.41	**-0.43**	0.16	**0.45**	0.12	**0.19**

Journal Pre-proof

Figures

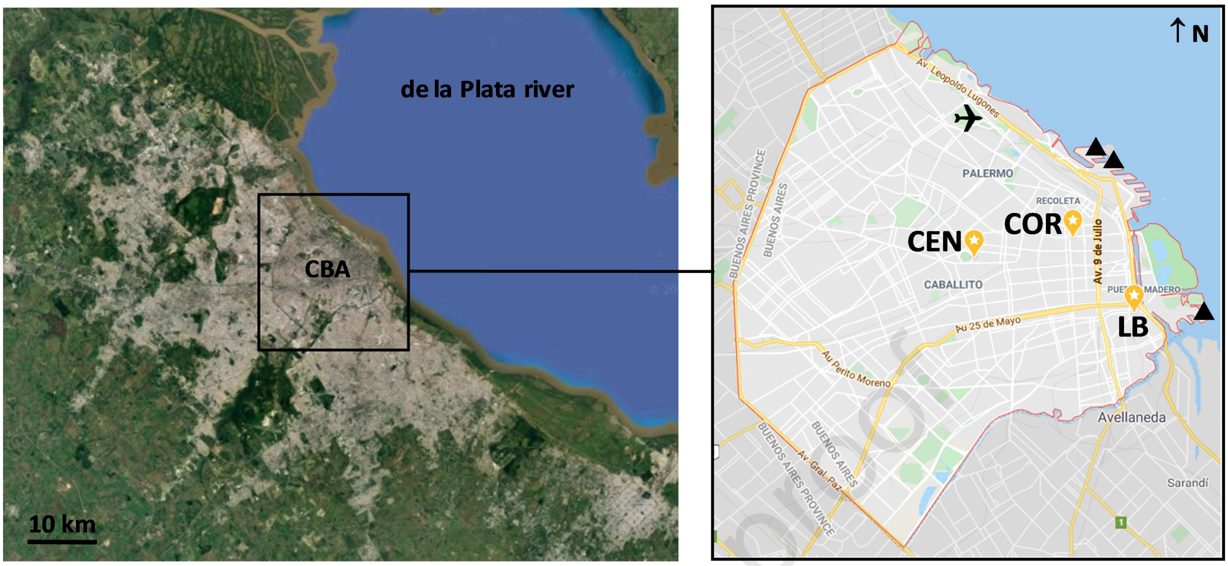


Figure 1: Map of the Metropolitan Area of Buenos Aires, including the city of Buenos Aires (CBA) and location of the three air quality monitoring stations (CEN: Parque Centenario, COR: Córdoba, LB: La Boca) from the `Agencia de Protección Ambiental (APRA)'. The locations of the AEP meteorological station (Domestic Airport) and three thermal power plants (triangles) are also indicated. [*Source: Google Maps*]

Journal Pre-proof

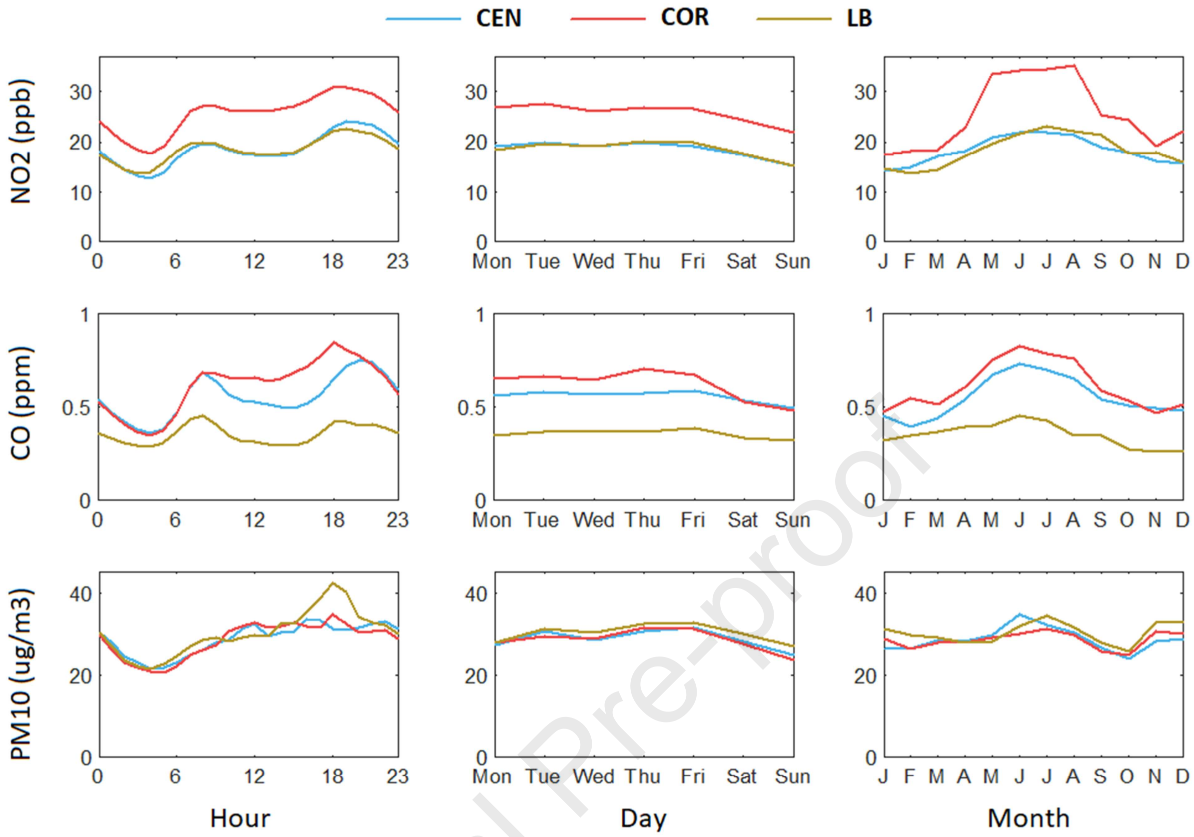


Figure 2: Temporal variation of air pollutant concentrations at hourly, monthly and weekly scales, at each monitoring site (CEN: Parque Centenario, COR: Córdoba, LB: La Boca).

Journal Pre-proof

Journal Pre-proof

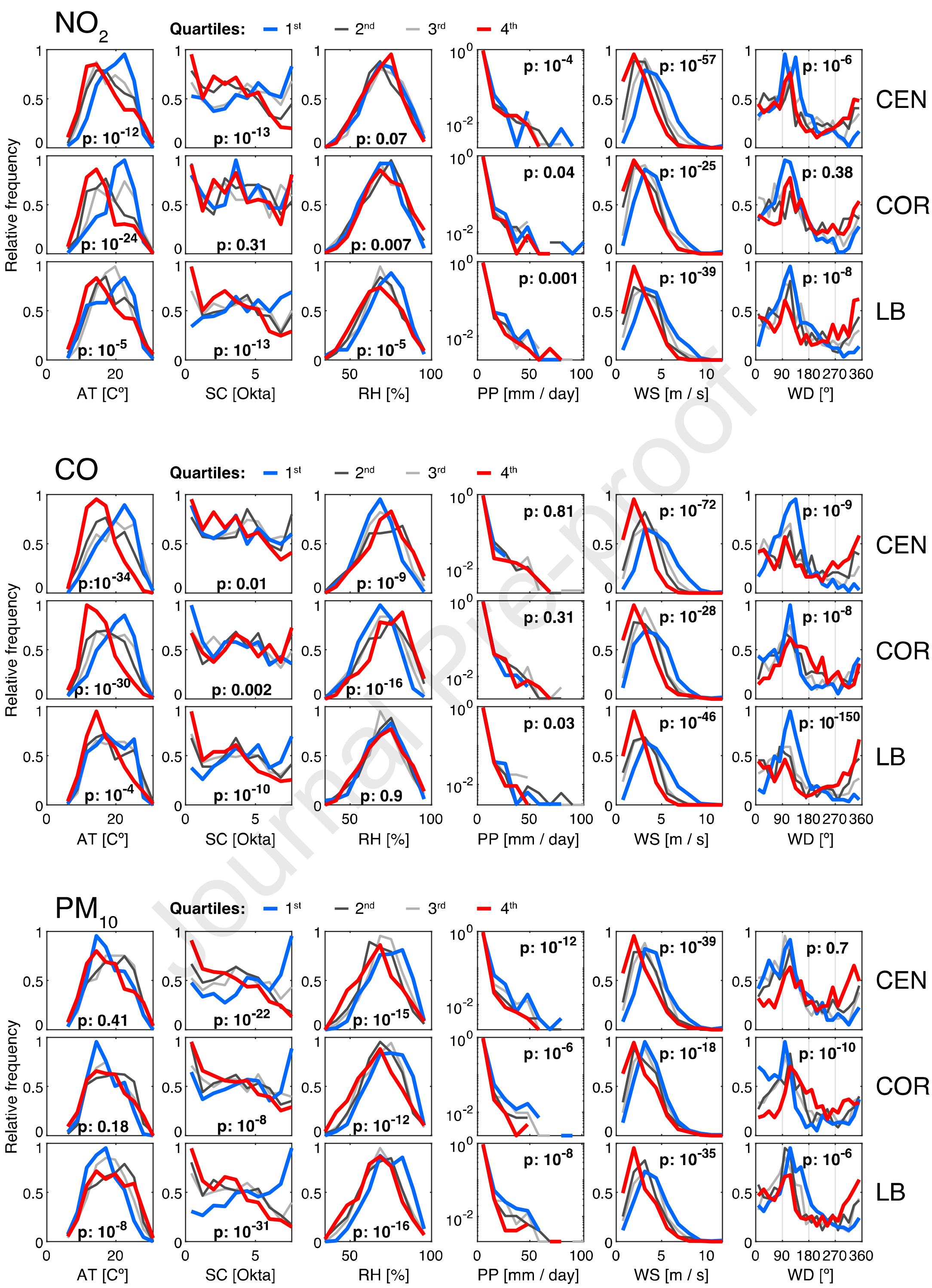


Figure 3: Distributions of surface daily mean meteorological data associated to each quartile interval of daily mean (a) NO_2, (b) CO and (c) PM_{10} concentrations, at each monitoring site. The p-value in each figure indicates the statistical significance of the difference between the average values of the 1st (blue) and 4th (red) quartiles.

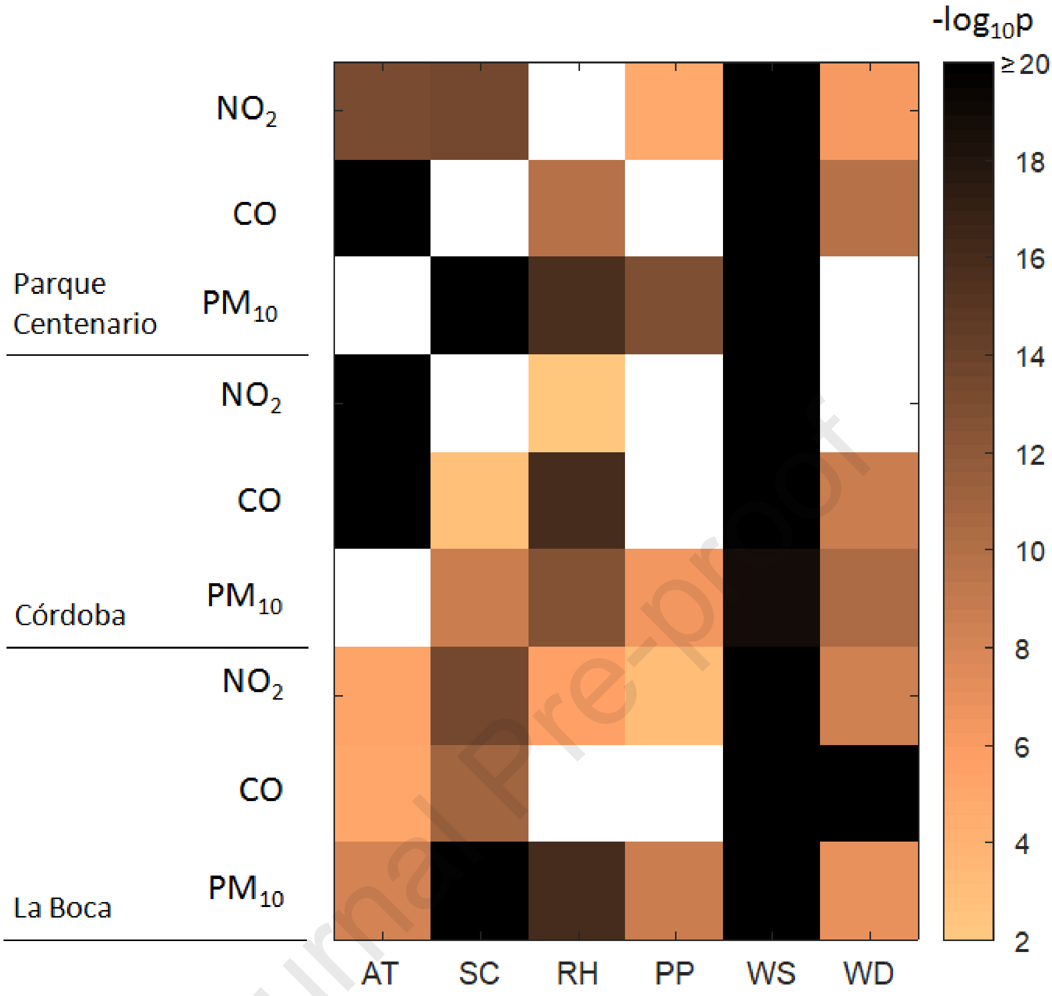


Figure 4: Statistical comparison between the mean meteorological variables corresponding to the lower and upper quartile distribution of the daily mean concentration of each pollutant (blue and red series in Figure 3, respectively), at each site. [significance level of the difference (-log10 of p-value); white indicates no statistically significant difference ($p > 0.01$) and darker colours indicate more significant difference].

Journal Pre-proof

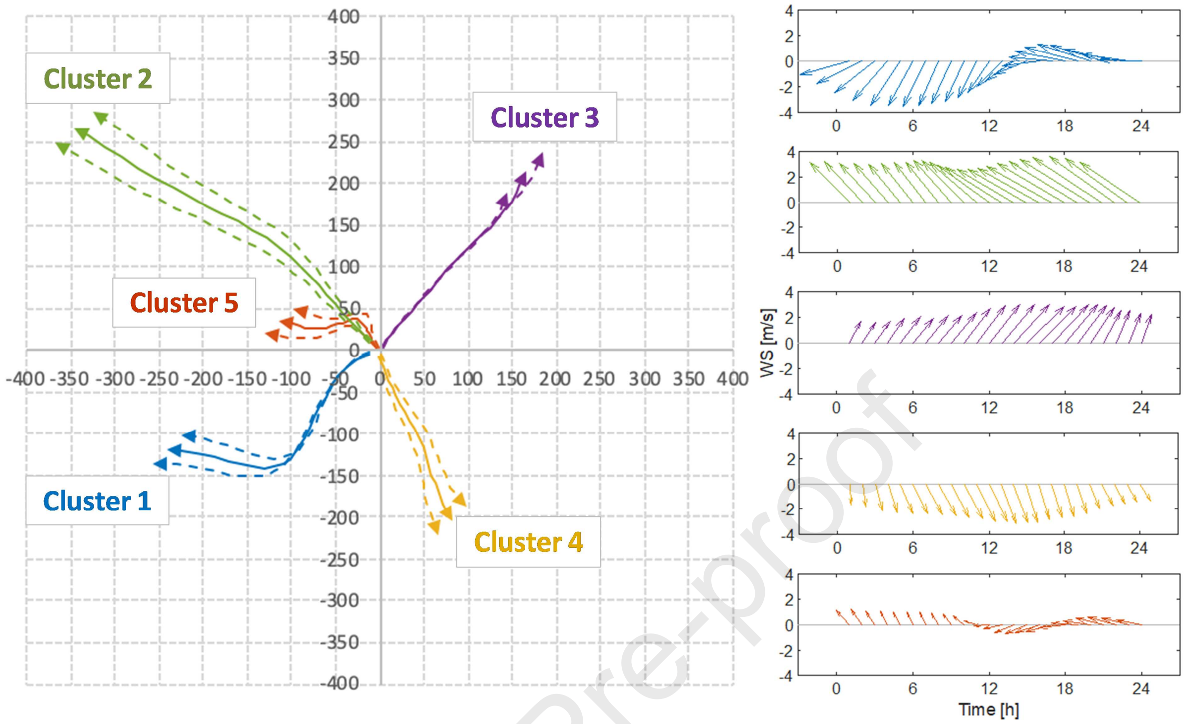


Figure 5: Accumulated hourly mean wind vector (solid lines) and corresponding 95% CI (dashed lines) along the day for each cluster (left) and their corresponding mean daily wind sequences (right).

Journal Pre-proof

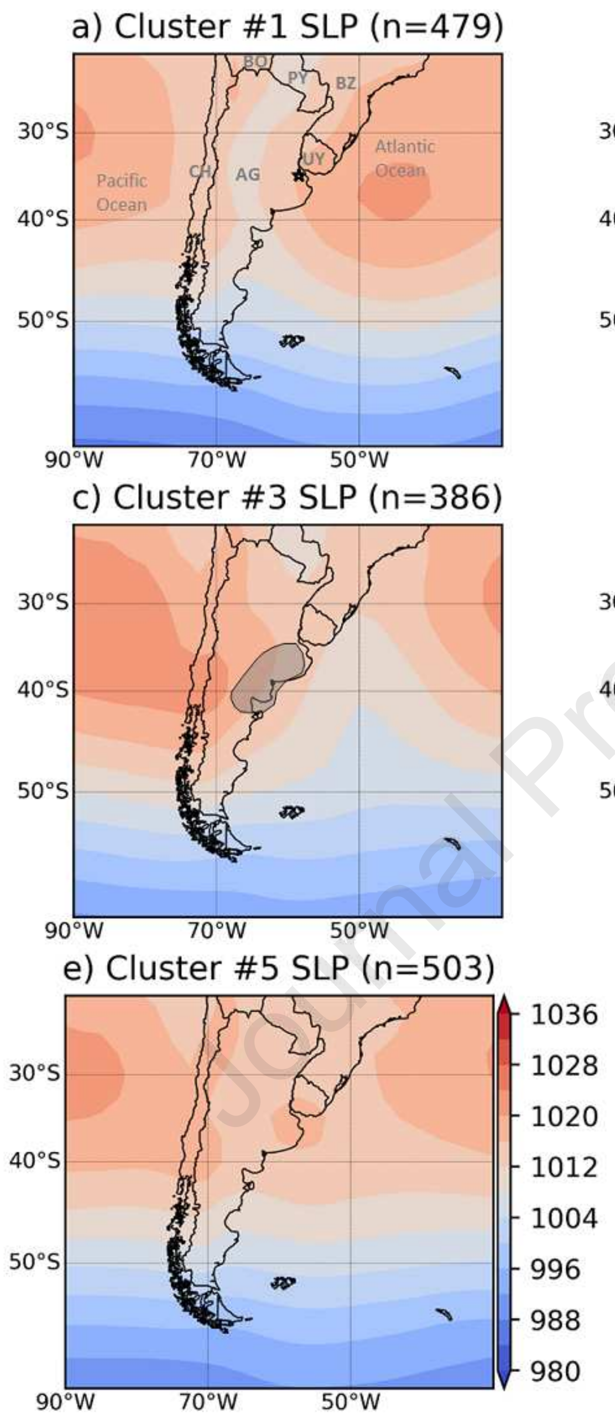


Figure 6: Sea level pressure (SLP) averaged for each cluster a) 1 to e) 5. The number of days within each cluster is indicated between brackets. SLP units are hPa. Shaded regions indicate potential natural sources of PM for the area of the city of Buenos Aires. In Figure 6.a), Argentina (AG), the CBA (star) and neighbour countries (UY: Uruguay, PY: Paraguay, BZ: Brazil, BO: Bolivia, CH: Chile) are indicated.

Journal Pre-proof

Journal Pre-proof

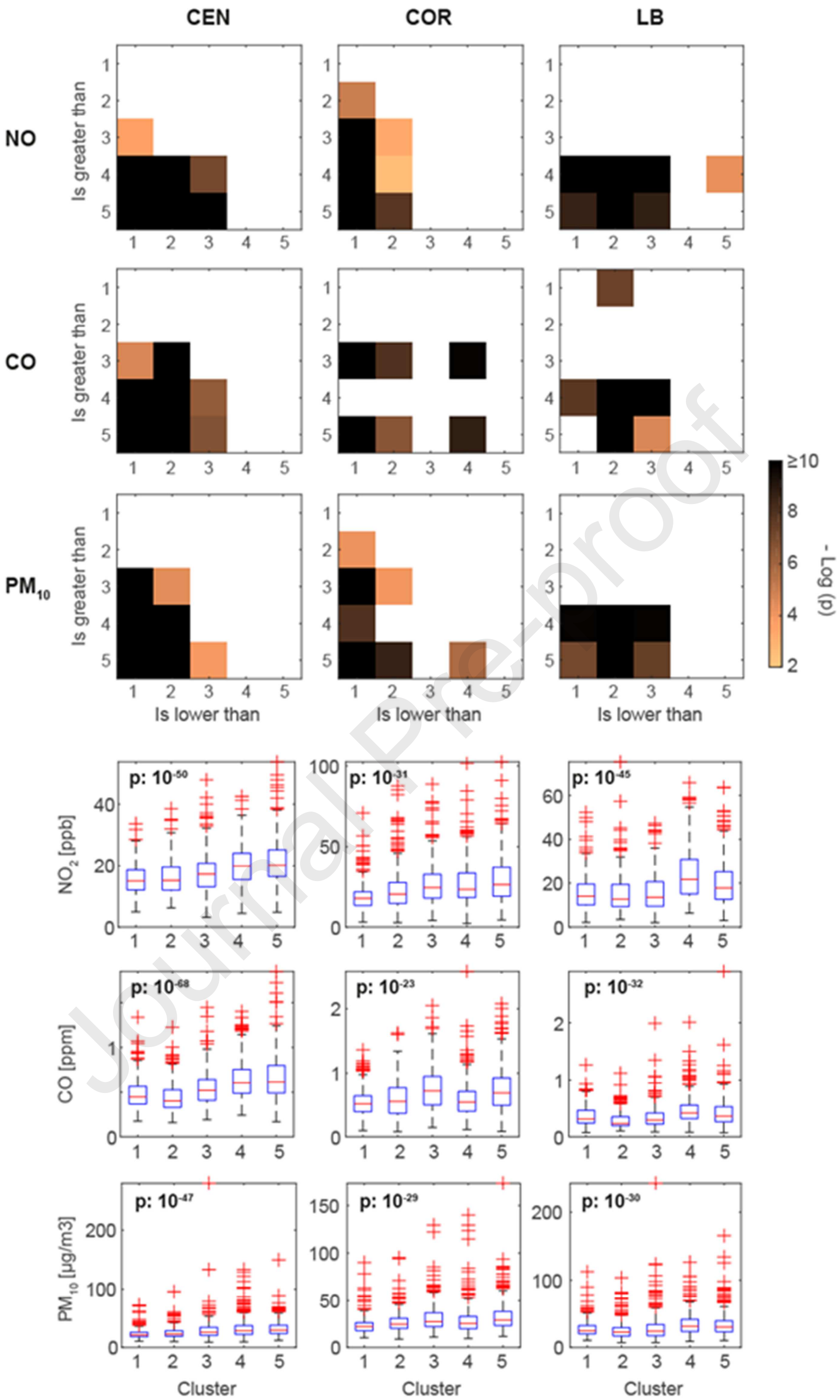


Figure 7: Matrix of p-values obtained from the Dunn-Sidàk multi-comparison test applied to the cluster mean concentrations (upper panel; white indicates no statistical difference between the mean pollutant concentrations of two clusters) and box-plot of the cluster concentrations (lower panel; the p-values obtained from the Kruskal-Wallis test is indicated), for each pollutant and monitoring site.

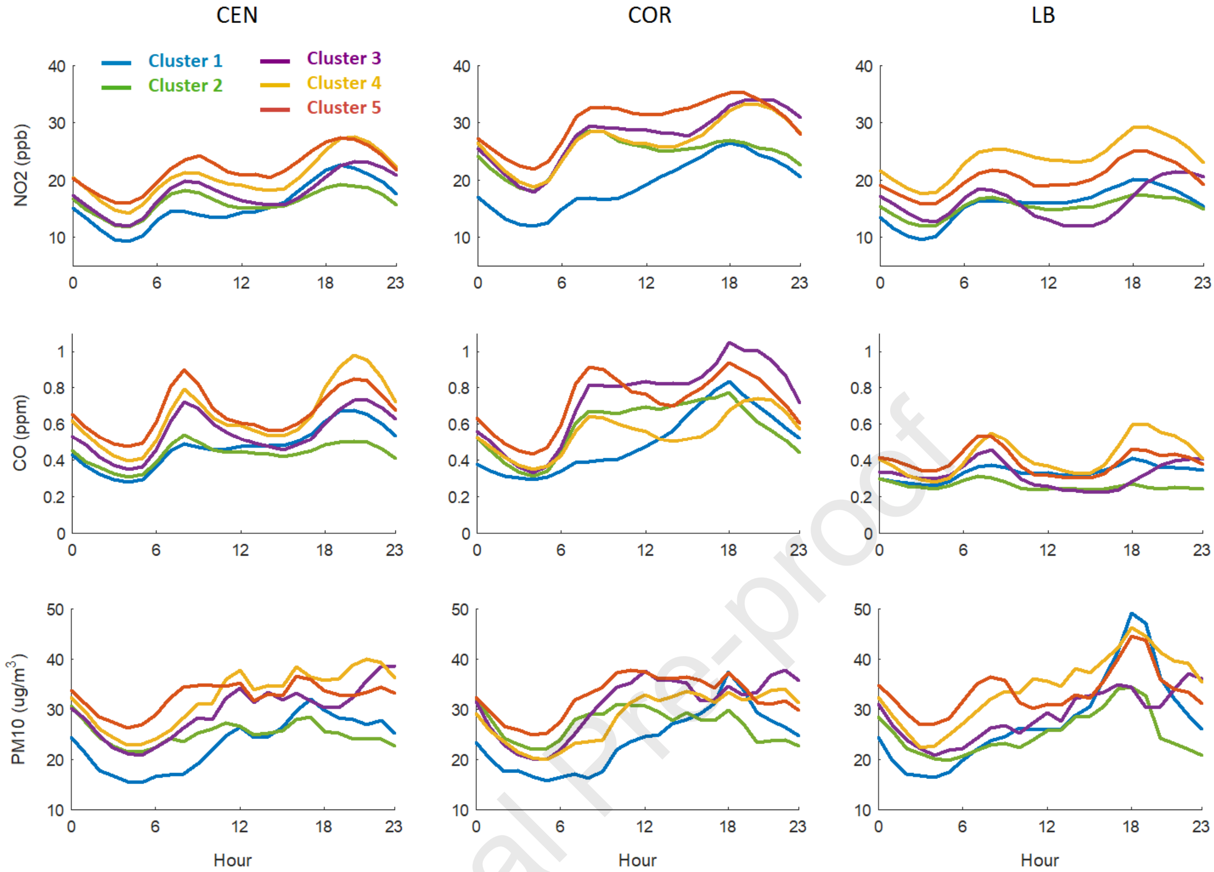


Figure 8: Hourly variation of air pollutant concentrations at each monitoring site (CEN: Parque Centenario, COR: Córdoba, LB: La Boca), by cluster.

Journal Pre-proof

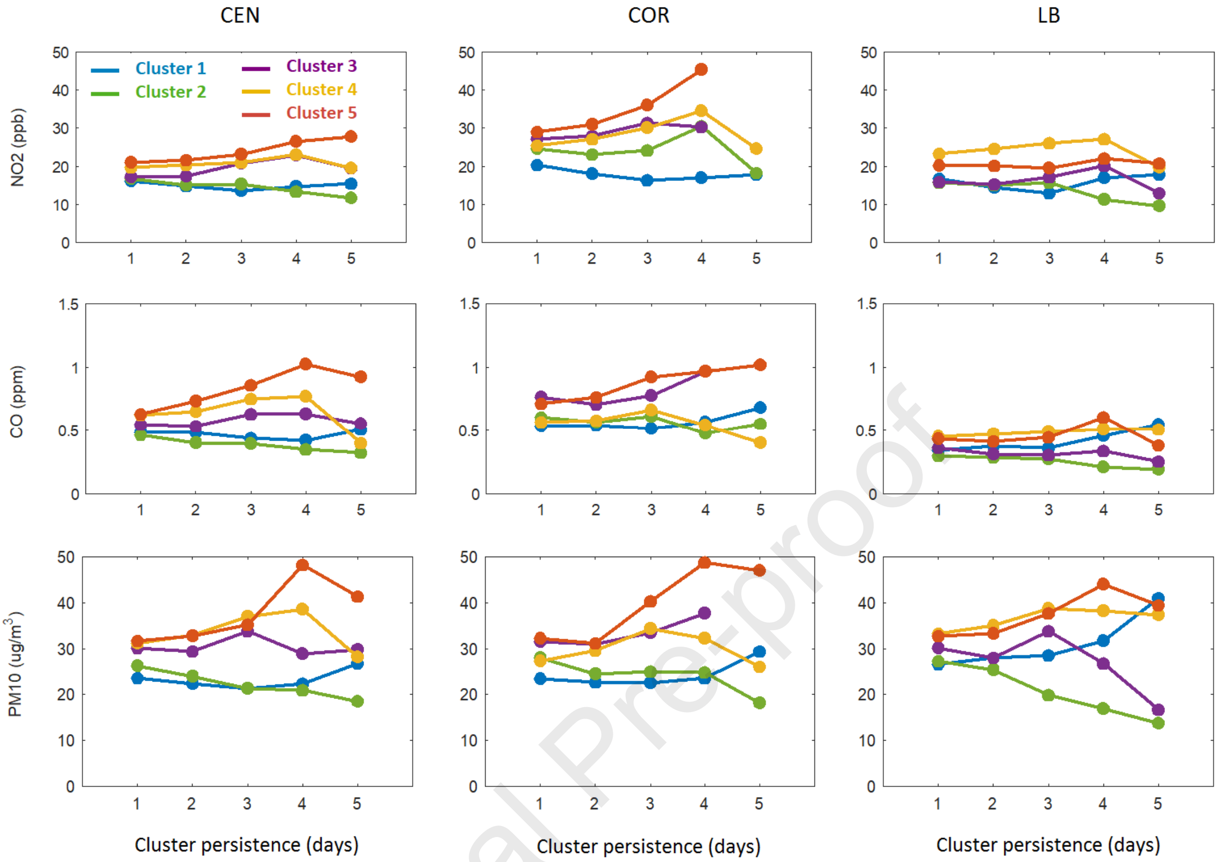


Figure 9: Mean pollutant concentration occurring with different wind sequence persistence periods, by cluster.

Journal Pre-proof

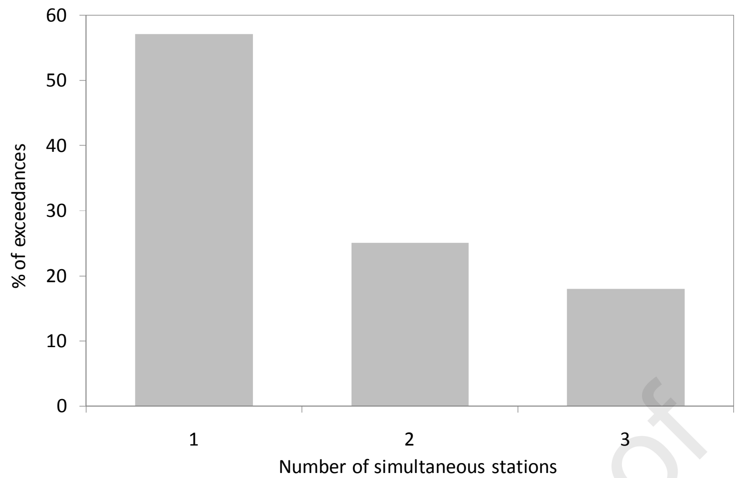


Figure 10: Percentage of days with PM_{10} concentration exceeding the WHO guideline (50 μg/m^3, 24 h) at one, two and the three sites simultaneously, relative to the total number of exceedance days (229).

Journal Pre-proof

Journal Pre-proof

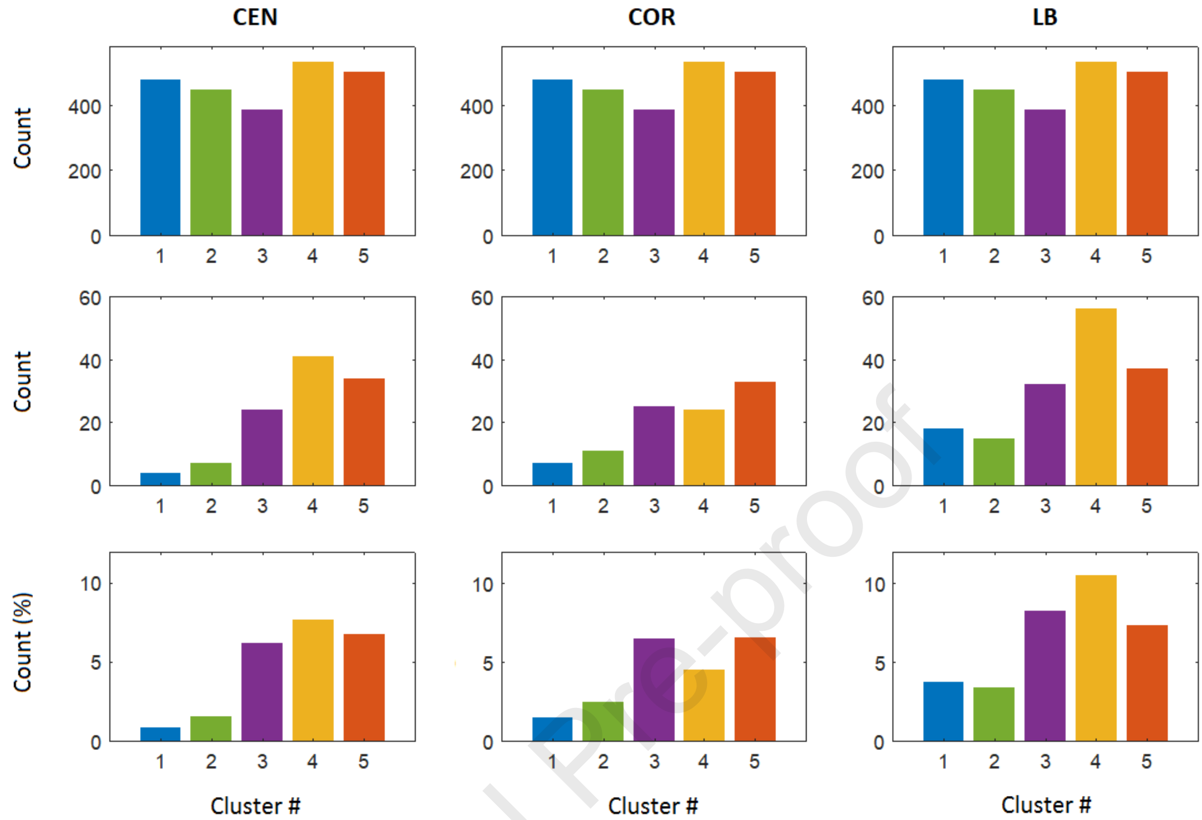


Figure 11: For each cluster (colour code): day count (upper panel), exceedance day count (middle panel) and percentage of days with exceedance (lower panel), at each monitoring site (columns, as indicated).

Journal Pre-proof

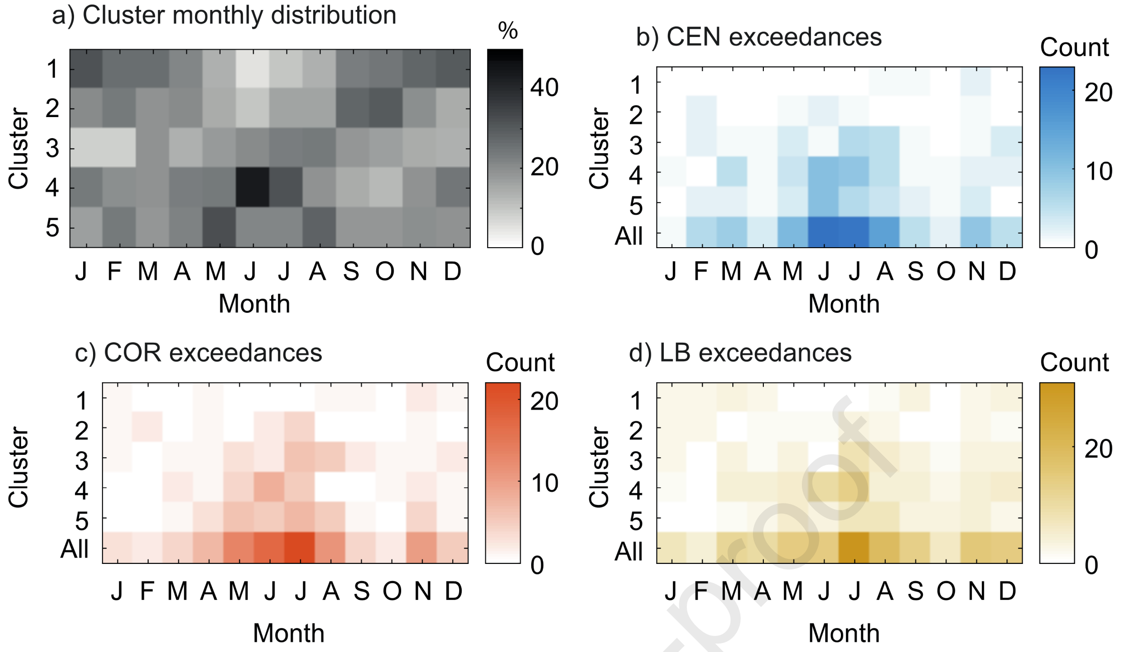


Figure 12: Monthly distributions of: (a) cluster occurrence (%); and number of PM_{10} exceedance days by cluster at the three monitoring sites: (b) CEN, (c) COR and (d) LB.

Journal Pre-proof

Journal Pre-proof

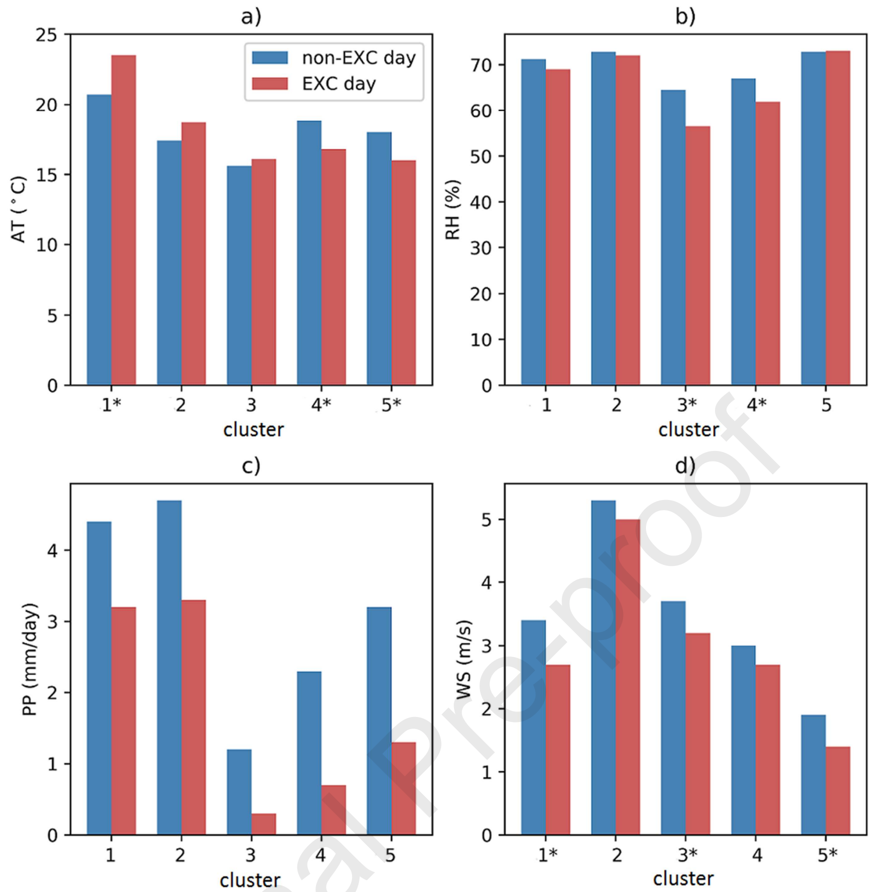


Figure 13: Mean values of a) temperature (AT, °C); b) relative humidity (RH, %); c) precipitation (PP, mm/day); and d) wind speed (WS, m/s) considering days with PM_{10} concentration below (blue) and above (red) the 50 μg/m^3 threshold in at least one of the monitoring stations, by cluster. Asterisks next to the cluster number indicate that the differences are statistically significant at the 99% confidence level according to a two-sided t-Student test.

Journal Pre-proof

Journal Pre-proof

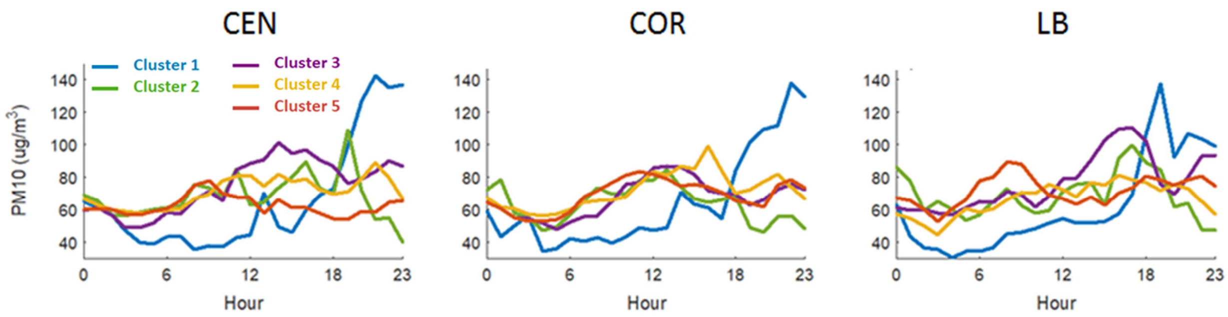


Figure 14: Hourly variation of PM_{10} concentrations during exceedance days at each monitoring site (CEN: Parque Centenario, COR: Córdoba, LB: La Boca), by cluster.

Journal Pre-proof

Highlights

- Hourly AQ data and surface wind sequence are analysed over a 8-year period

- Large number of PM_{10} exceedances over the WHO guideline at the three sites

- Five 1-h wind sequence clusters with distinctive AQ features

- Wind and AQ variations suggest specific local source contributions

- Two clusters highlight potential urban and regional contributions to the events

Journal Pre-proof

Declaration of interests: none

☐The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Journal Pre-proof