Characteristics of the east Mediterranean dust variability on small spatial and temporal scales

Yuval [a, *], Meytar Sorek–Hamer [a], Amnon Stupp [b], Pinhas Alpert [b], David M. Broday [a]

[a] Department of Civil and Environmental Engineering, Technion, Israel Institute of Technology, Haifa, Israel
[b] Department of Geophysics and Planetary Sciences, Tel-Aviv University, Tel Aviv, Israel

HIGHLIGHTS

- Using cluster analysis to detect dust presence in a comprehensive PM10 database.
- Dust events' duration was delineated on a high (half-hourly) temporal resolution.
- Small scale variability in dust presence was found in both time and space.
- The detected correlation of dust with elevation ASL hints about the dust dynamics.
- Important methodology for fine-tuning dust exposure for epidemiological studies.

ABSTRACT

The presence of naturally–occurring dust is a conspicuous meteorological phenomenon characterised by very high load of relatively coarse airborne particulate matter (PM), which may contain also various deleterious chemical and biological materials. Much work has been carried out to study the phenomenon by modelling the generation and transport of dust plumes, and assessment of their temporal characteristics on a large (>1000 km) spatial scale. This work studies in high spatial and temporal resolution the characteristics of dust presence on the mesoscale (>100 km). The small scale variability is important both for better understanding the physical characteristics of the dust phenomenon and for PM exposure specification in epidemiological studies. Unsupervised clustering–based method, using PM_{10} records and their derived attributes, was developed and applied to detect the impact of dust at the observation locations of a PM_{10} monitoring array. It was found that dust may cover the whole study area but very often the coverage is partial. On average, the larger the spatial extent of a dust event, the higher and more homogeneous are the associated PM_{10} concentrations. Dust event lengths however, are only weakly associated with the PM concentrations (Pearson correlation below 0.44). The large PM concentration variability during spatially small events and the fact that their occurrence is strongly correlated with the elevation above sea level of the reporting stations (Pearson correlation 0.87, p–value $< 10^{-5}$) points to small scale spatiotemporal dynamics of dust plumes.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Dust storms originating in the world's deserts transport particulate matter (PM) from the sources to distances of up to thousands of km and may collect along their trajectory additional pollutants and allergens (Goudie, 2014). The presence of dust in the earth's atmosphere is an interesting phenomenon that was studied extensively with the goal of improving our understanding of its formation, transport and characteristics (eg Prospero et al., 2002; Goudie and Middleton, 2006; Maghrabi et al., 2011; Israelevich et al., 2012). The health effects of dust raised concerns and many studies looked at its possible association with various medical outcomes (eg the recent reviews by Karanasiou et al., 2012; Goudie, 2014, and references therein). Most of the studies looked at the dust phenomenon variability on a large (>1000 km) spatial scale (Querol et al., 2009; Pey et al., 2013) or at one single location (Koçak et al., 2007; Ganor et al., 2009; Mallone et al., 2011; Krasnov et al., 2014).

To the best of our knowledge, the variability of dust presence at the mesoscale has not been looked at to date.

Many different methods have been utilised for studying airborne dust. Early work usually used chemical speciation in order to assess the contribution of natural sources to the total PM levels (eg, Ganor and Foner, 2001; Rodríguez et al., 2002). Particles with large fractions of Si, Al, Ca and other crustal components were attributed to mineral dust. Such methods involve costly and tedious analysis of filter data samples. Later studies sought methods that could be employed in a cheaper and more efficient manner. Following Escudero et al. (2005), in most cases an array of tools is used for the initial identification of dust episodes (Escudero et al., 2007a; Querol et al., 2009; Pey et al., 2013). These include back−trajectory analysis, satellite imagery, meteorological maps, and aerosol concentration maps from dust models. Mallone et al. (2011) used light detection and ranging (LIDAR) for detection of dust in Rome. The dust contribution is usually estimated using various measures of PM_{10} levels and their difference from some background values (Koçak et al., 2007; Escudero et al., 2007b; Querol et al., 2009; Dadvand et al., 2011; Mallone et al., 2011; Pey et al., 2013; Krasnov et al., 2014). Ratios of pollutant concentrations are used for verification (eg Mallone et al., 2011). In some cases only remote sensing (Zhang et al., 2006; Evan et al., 2006) or modelling outputs (Mitsakou et al., 2008; Jiménez-Guerrero et al., 2008) were used in dust studies.

All the above mentioned studies considered dust presence on a daily basis. The qualitative nature of the dust identification methods which they used, and the manual work which they require, set limits on the spatial and temporal extent of the phenomena which they were able to resolve. A daily temporal resolution seems compatible with epidemiological studies as health outcome rarely, if at all, are given at a higher resolution. It may also be sufficient for large scale studies concerned with dust contributions at the seasonal or annual time scales in a large spatial area. However, the turbulence associated with the transport of dust may be of much smaller temporal and spatial scales. Understanding the spatiotemporal variability of dust presence requires a much finer temporal resolution of the relevant data and an adequately spaced array of sensors. A scheme utilising those data have to function in an automatic manner. Ganor et al. (2009) introduced an algorithm to automatically detect dust presence using half−hourly PM_{10} data series. However, their algorithm was calibrated using daily dust observations at a specific location. Daily calibration data are too coarse to capture the high temporal variability that Ganor et al. (2009) expected to detect. (They assumed, correctly, that dust events may be as short as three hours.) Subsequent analysis also found some errors and omissions in the calibration data which were not detected during the original work. Moreover, as noted by Ganor et al. (2009), their calibration data may not be suitable for use in other locations, as indeed was found later by Viana et al. (2010).

This work explores the spatiotemporal variability of dust presence at the half−hourly temporal resolution in an area which is about 250 km by 40 km, using 19 monitoring stations reasonably spaced within it. The analyses are based on dust detection scheme that utilises only half−hourly PM_{10} concentration observations and their attributes. Unlike the above mentioned studies, our clustering classification scheme does not require calibration information. It is an unsupervised optimisation scheme (Liao, 2005; Hastie et al., 2009) which looks for hidden features in the data. It has been used extensively in the signal processing and pattern recognition fields (Liao, 2005) and we show it can be applied in environmental studies like ours which employ time series of substantial length.

2. Study area

The study area includes the Israeli shoreline and the western slopes of its coastal mountain range (see Fig 1). Local anthropogenic PM sources in the area are mainly transportation and large industrial plants. Additional major source of fine PM are secondary particulates transported from eastern and southern Europe (Asaf et al., 2008). Dust storms transport mineral dust to the region from the Sahara and the Arabian peninsula during the winter and the transition seasons. The dust contributes significantly to the total PM load (10−20% of the PM_{10} concentrations; Ganor et al., 2009) and there has been increasing interest in its impact on the population's health (eg Vodonos et al., 2014).

3. Data

PM monitoring and meteorological data were obtained from the Technion Centre of Excellence in Exposure Science and Environmental Health's air pollution monitoring data archive. The database includes all the half−hourly air quality monitoring data observed in Israel from 1997 to date. The data pass quality assurance and quality control processing before being released for use. The PM_{10} data used in this work are from standard monitoring stations (ie comply with the EU Council Directive 1999/30/EC for protection of human health), with at least 84% complete records for the years 2004−2013. The raw data were recorded by a tapered element oscillating microbalance (TEOM) devices (Thermo Scientific 1400) that provides a continuous direct mass measurement of particle mass. Table 1 provides descriptive statistics of the PM_{10} records and Fig. 1 shows the geographical locations of the stations. Note that the station numbers are ordered from south to north and that all the stations except 4 and 13 are within 35 km from the shoreline. Most of the stations are at elevations of less than 300 m above sea level. The only exception is station 4, located in Jerusalem at 786 m above sea level. Three of the PM_{10} stations (numbers 9, 10 and 15) also observe simultaneously $PM_{2.5}$. Most of the stations observe ambient temperature, and wind direction and speed. A few of them also observe relative humidity, barometric pressure and insolation.

Daily synoptic system classification at 12:00 UTC for the eastern Mediterranean during the study period 2004−2013 was calculated following Alpert et al. (2004). The method is based on a semi−objective classification of geopotential height, temperature and the horizontal wind components at the 1000 hPa level. Alpert et al. (2004) defined 19 synoptic systems characteristic to the eastern Mediterranean, which can be lumped into six groups: Red Sea Troughs (RST), Persian Trough (PT), High to the West (HW), Siberian Highs (SH), Winter Lows (WL) and Sharav Low (SL). Detailed description of the synoptic systems and their grouping as well as the classification can be found in Alpert et al. (2004).

4. Methods

4.1. Dust presence detection using cluster analysis

The major characteristic of naturally−occurring dust is an unusually high PM load. Dust-related PM tends to be of the coarser fractions (Koçak et al., 2007) and thus the main variable which we used for the detection of dust was the PM_{10} concentration. For each PM_{10} monitoring location, the half-hourly time points in 2004−2013 were grouped to dust−related or normal points based on the PM_{10} concentration, and the corresponding running maximum and minimum series, as explained below. Additional meteorological and air quality variables were also considered but were not used in the final classification for three reasons: (a) with the exception of the $PM_{2.5}/PM_{10}$ concentration ratio, none of them seems to contribute a considerable power for the clustering process beyond what the PM_{10} concentration and its attributes provide. (b) using additional variables, with values in a dynamic range much
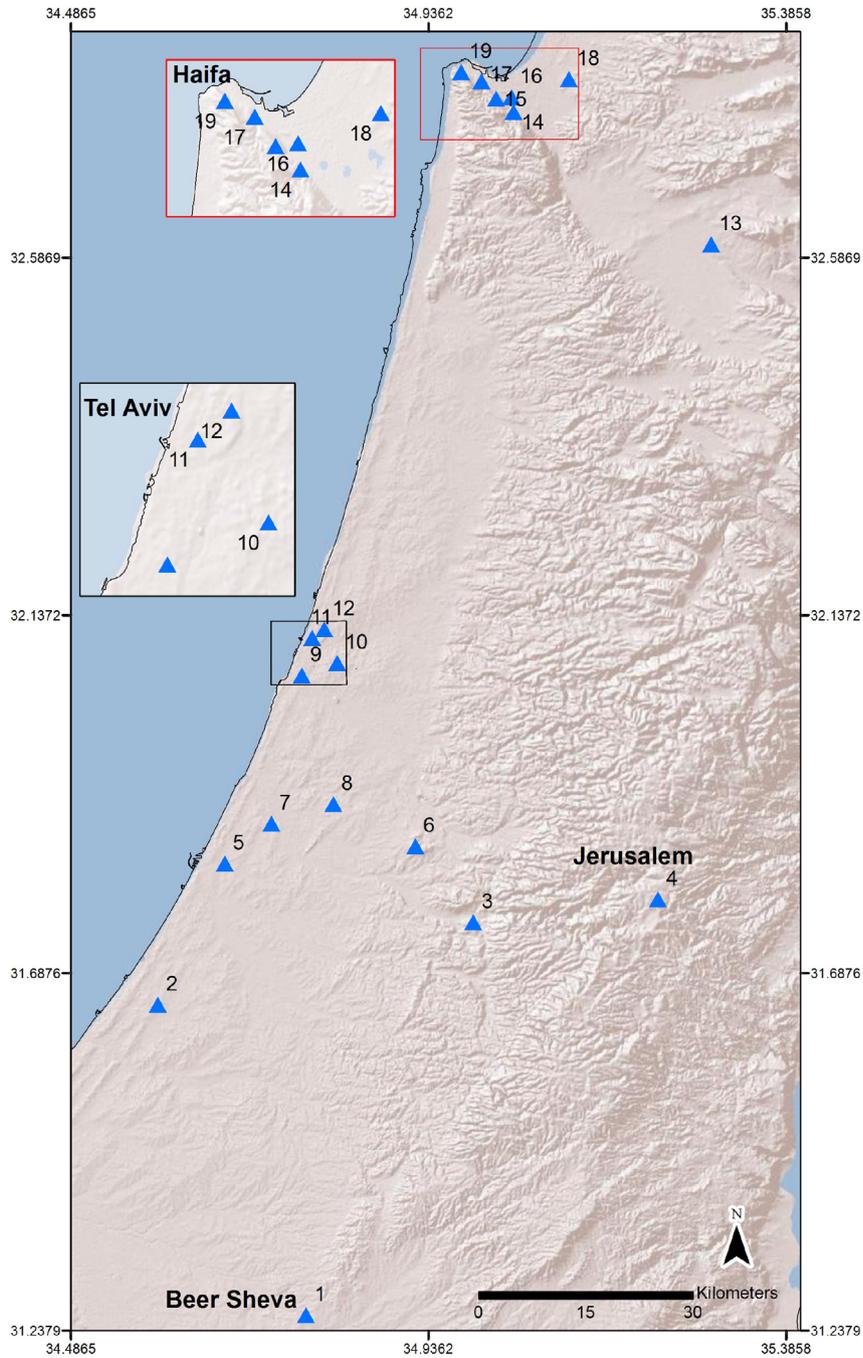
Fig. 1. The study area map showing the shoreline, the locations of the PM$_{10}$ monitoring locations (triangles) and their numbers, and the topography.

different than that of the PM$_{10}$ concentrations, requires a normalisation of the values, for which some undesirable arbitrary decision with a very large impact on the clustering must be made. (c) using additional variables in the classification process would exclude their use for independent assessment of the classification.

The running PM$_{10}$ maxima and minima series were calculated such that each element in the series is given by $y_i = ext \{x_{i-q}, x_{i-q+1}, \cdots, x_{i+q-1}, x_{i+q}\}$, where **x** is the PM$_{10}$ series, $2q + 1$ is the operation's window width and $ext\{\}$ takes the extremum, either maximum or minimum, of a sequence of values. The window width should be related to the mean duration of dust events. We used $q = 7$ time points (3.5 h) which results in a 7.5 h window width.

Sensitivity analysis did not reveal significant differences in the results using $3 \leq q \leq 9$.

Cluster analysis divides a set of state vectors into subsets, or clusters, such that members of each cluster are more similar in some mathematical sense to each other than to members of other clusters. We used for that purpose the $K-$means clustering method (Liao, 2005). The time−dependent state vectors form in our case the three−dimensional data array $\mathbf{Z}_i, i = 1, \cdots, N$, where the PM$_{10}$ concentration and the corresponding running maximum and minimum values are the components at each time point, and N is the number of half−hourly time points in the study period 2004−2013. Our first goal was a partition of the N state vectors into

Table 1
Descriptive statistics and station information. The table provides the mean, standard deviation, median, 95th percentile, 99th percentile and maximal values ($\mu g m^{-3}$) of each monitoring station's PM_{10} record, its percentage of missing data and the elevation above sea level (ASL) of the monitoring station.

Station no.	Mean	Std	Median	95%	99th%	Max	% Missing	ASL (m)
1	55.0	122.3	36.2	124.0	379.8	3997.3	16.2	251
2	52.4	98.6	36.1	123.6	361.1	2947.2	9.4	56
3	50.5	72.8	35.2	124.0	365.2	1553.2	5.0	297
4	57.4	123.9	37.3	135.6	417.4	4398.5	4.0	786
5	54.5	91.1	39.1	126.5	331.9	3194.1	4.3	7
6	52.9	88.3	36.5	130.4	362.4	2254.9	5.6	234
7	52.5	87.9	37.7	119.8	322.2	2912.1	14.0	19
8	50.7	86.3	35.1	118.2	329.0	2276.6	9.0	57
9	58.3	89.5	41.9	134.1	340.6	3000.4	6.1	22
10	52.5	87.4	36.7	124.7	332.8	2899.2	4.8	56
11	50.8	82.6	36.5	115.9	299.6	3127.8	13.2	6
12	57.7	90.2	39.9	141.0	390.3	2060.5	13.2	23
13	46.8	76.7	33.1	112.6	290.2	2306.9	12.7	53
14	51.1	78.1	36.3	124.5	360.7	1665.3	5.8	101
15	44.1	85.4	30.4	111.9	303.5	2837.9	1.6	201
16	44.7	74.4	31.9	107.2	267.4	2980.3	3.7	14
17	50.5	90.0	35.2	124.7	323.9	3599.6	3.4	65
18	48.9	66.7	36.8	111.9	310.0	1323.9	1.4	38
19	48.4	88.4	32.9	122.1	321.8	3656.2	10.3	200

K clusters $S = S_1, S_2, \cdots, S_K$ with centres $\mu_i, i = 1, \cdots, K$, such that the sum of the within-clusters sum of squared distances from the centres is

$$\arg \min_{S} \sum_{j=1}^{K} \sum_{\mathbf{Z} \in S_j} ||\mathbf{Z} - \mu_j||^2. \tag{1}$$

We carried this out using the Matlab© kmeans function (Matlab, 2013). The K—means minimisation is not guaranteed to arrive at the global minimum and thus the final values of the clusters' centres depend on the initial values from which the search for the minimum starts. In our case the sensitivity to the initial values was very weak and the corresponding elements of the clusters' centres were all within 1% of each other for many trial runs that started with different random initial values.

The decision about the value of K can be made based on various criteria. Using $K = 2$ did not result in the desired partition to dust-related and non—dust time points. In all the stations one of the clusters was clearly composed of only dust—related time points (PM_{10} concentrations above 500 $\mu g m^{-3}$). However, this cluster was very small. The other, much larger, cluster included mainly normal time points but also many time points which were obviously dust-related (PM_{10} concentrations in the 100 s of $\mu g m^{-3}$). A common method to decide about an appropriate value for the number of clusters K is using the very intuitive yet rigorous F statistic, which is the ratio of the between—clusters variability to the within—clusters variability (Liao, 2005). The K maximising the F statistic is the most desirable one in the pure statistical sense. We found clear maxima in the curves of F as a function of K for the data of all the stations (different maximum for each station). Using the maximising K, time points classified to different clusters had specific meteorological attributes that correspond to known meteorological scenarios with respect to dust presence. A few of the clusters were always distinctively dust-related while others were distinctly associated with normal conditions (based on their meteorological characteristics, the associated synoptic system and their PM_{10} attributes). However, there were always a cluster or two which could not be classified clearly. These clusters were in many cases quite large so their classification as dust—related or not could have a very large impact on the subsequent analyses. Hence our final scheme used a partition to the very large number of $K = 150$ clusters in all the stations. Each of these clusters happens to be a sub—cluster of the partition based on the F statistic. The classification of the clusters to

dust or not dust was based on cut—off values for the cluster centre elements. The large number of small clusters enabled tagging them as associated with dust or not with a fine resolution; the statistical analyses presented later were only very weakly sensitive to the number of clusters for $K = 100$ and above. The cut—off conditions we used were as follows: a cluster was define as impacted by dust if (a) the PM_{10} and maximal PM_{10} concentrations of its centre were above or equal 110 $\mu g m^{-3}$ and 250 $\mu g m^{-3}$, respectively, or (b) both the PM_{10} and minimal PM_{10} concentrations of a cluster's centre were above or equal 110 $\mu g m^{-3}$. The cut—off values were based on our experience and the general idea underlying the method of Ganor et al. (2009). The somewhat arbitrary nature of the cut-off conditions have some impact on the number of detected dust time points. However, the spatiotemporal characteristics of the dust phenomenon manifested themselves clearly and with minimal sensitivity for a wide range of cut—off values.
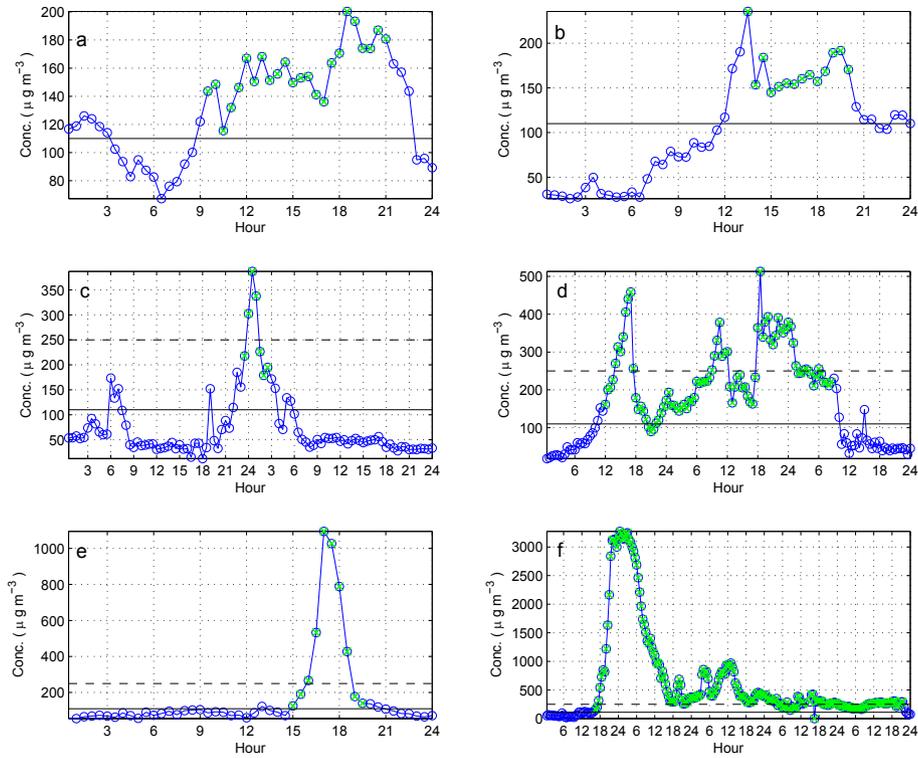
4.2. Dust events

The cluster analysis considers the input sets $\mathbf{Z}_i$ with no reference to their temporal order. However, in most cases the process resulted in continuous sequences of dust-related time points, as can be seen in Fig. 2. The intensity of dust presence during certain synoptic systems fluctuates significantly, resulting sometimes in short dipping of the PM_{10} concentration to normal values, followed by renewed high levels (eg Fig. 2d). In cases where the end of a dust-flagged sequence of time points was less than six time points (three hours) from the starting point of the next sequence, the time points in between were flagged as dust—related. We refer to continuous sequences of time points flagged as dust—related as dust events. Only for comparisons with previous studies we refer also to dust days, which we consider to be 24-h periods, starting at 00:00, during which dust was detected.

5. Results

### 5.1. Characteristics of normal and dust—related periods

The presence of high loads of aerosols and their relatively coarse size distribution are the major characteristics of natural-ly—occurring dust. Figs. 3 and 4 demonstrate that the cluster analysis classified the time points during the study period into two

Fig. 2. Examples of PM$_{10}$ concentration time series on days when dust events were detected in station 9. The time points flagged as dust—related are highlighted in green. The solid and dashed lines mark the values used in the criteria defining the cluster centres as dust event (see Section 4.1 for more details). (a) Event from 09:00 to 21:00 on 04-01-2004 (12.0 h), synoptic system group: Winter Low. (b) Event from 13:00 to 20:00 on 02-02-2006 (7.0 h), synoptic system group: Sharav Low. (c) Event from 23:00 on 26-05-2005 to 02:00 on 27-05-2006 (3.5 h), synoptic system group: unknown, the very short event's duration did not include the 12:00 UTC hour, the time for which the synoptic systems were classified. (d) Event from 12:00 on 21-10-2011 to 08:00 on 23-10-2011 (45 h), synoptic system group: Red Sea Trough. (e) Event from 15:00 to 19:00 on 31-03-2005 (5 h), synoptic system group: Sharav Low. (f) Event from 16:00 on 25-02-2006 to 21:00 on 01-03-2006 (101.5 h), synoptic system group: changing from Winter Low (26-02 and 27-02) to Siberian High (28-02) and to Winter Low again (01—03). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

groups with clearly different distributions with respect to these two characteristics. Figs. 3 and 4 as well as the other figures in this section and the next one use the data from stations 9. It is centrally located within the study area and have records with high completeness for most of the variables. The results for other

stations are very similar. Fig. 3 shows the empirical density distributions of PM$_{10}$ concentrations at normal (non—dust) and dust—related time points. The range of PM$_{10}$ concentrations during non—dust times is between zero and 140 μgm^{-3}, with only 241 (0.14%) of the values above the median of the concentrations during
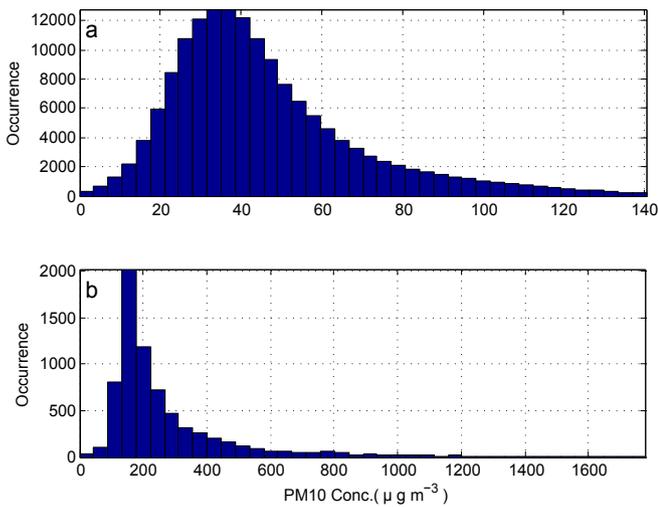
Fig. 3. The empirical probability density distributions of PM$_{10}$ concentrations in station 9 of (a) the sample of normal time points and (b) the sample of time points flagged as associated with dust. Values above the 99% percentile of the samples were not used for the plots for visual purposes.
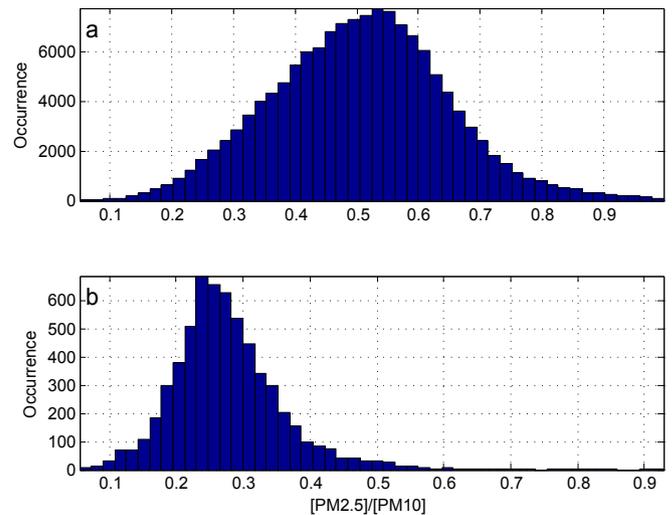
Fig. 4. The empirical probability density distributions of the PM$_{2.5}$/PM$_{10}$ concentration ratio in station 9 of (a) the sample of normal time points and (b) the sample of time points flagged as associated with dust.

dust time points. During time points classified as impacted by dust, the PM_{10} concentrations are up to values above 1600 μgm^{-3}, with only 23 (0.013%) of the values below the median of the concentrations during normal times. Fig. 4 shows the distributions of the $PM_{2.5}/PM_{10}$ concentration ratios during the normal and dust–related time points. The medians of the PM fractions' ratios are 0.508 and 0.266 for normal and dust–related distributions, respectively. Only in 3.8% of the normal time points the ratio was lower than the median value during dust–related time points and only in 1.2% of the dust–related time points the ratio exceeds the median ratio during normal times. Many meteorological variables also exhibited different distributions during dust presence and normal times, however, the distributions are not very distinct as in Figs. 3 and 4. The distributions of ambient temperature, relative humidity and insolation for dust–related and normal times are shown in the electronic Supplementary Material in Figs. SM1, SM2 and SM3, respectively.

Dust presence is usually associated only with a few types of synoptic systems while very rarely, or never existing, in others. Fig. 5 shows the prevalence of the synoptic system groups during dust-related and normal times. Fig. SM4 in the Supplementary Material shows the corresponding figure for each synoptic system separately. As the synoptic system classification was available only for 12:00 UTC, we used only the daily time point at that hour to carry out the counting to produce Fig. 5 and SM4. There is a clear difference between the prevalence of the synoptic system groups during dust–related and normal times. The relative prevalence of the HW and PT groups, mainly spring and summer phenomena and the most common during normal times, is considerably smaller during dust time points. The prevalence of the RST, SH and especially the WL groups, mainly fall and winter phenomena known as possible dust–related, increases. Occurrences of the SL synoptic system group are rare and it is thus not highly represented in any of the Fig. 5 plates. However, as mentioned by Ganor et al. (2010), it should be noted that in relative terms its presence during dust–related time points is the highest: 29% of the SL occurrences are associated with dust compared to corresponding percentages of 6.0, 1.4, 1.9, 3.5 and 11.7% for the RST, PT, HW, SH and WL groups, respectively.
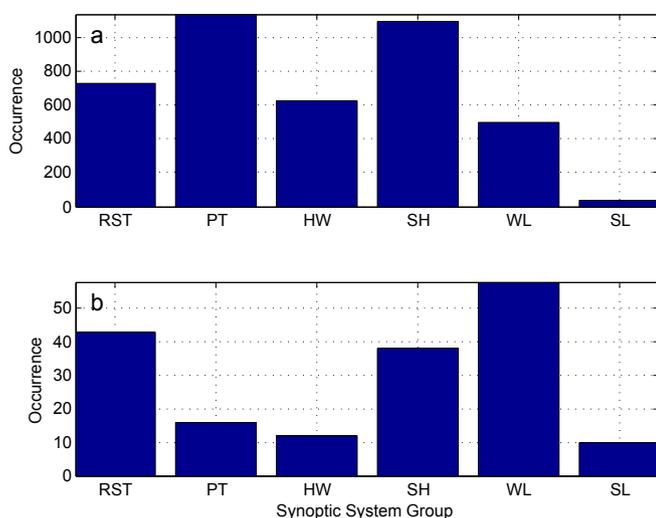
5.2. Temporal analysis

Fig. 2 shows six examples of PM_{10} concentration time series observed during dust events. The events are very different one from the others in their length, peak PM_{10} level and PM_{10} concentration variability. Figs. SM5 and SM6 in the electronic Supplementary Material show visual images taken from the Aqua satellite during the events depicted in Figs. 2e and f. Fig. 6 presents the length distribution of all of the dust events. The most prevalent are events of between four and eight hours but very short events, as short as three hours (eg Fig. 2c), are not rare. The longest events are more than four days long (eg Fig. 2f). It is noteworthy that the intensity of dust events is only weakly associated with their lengths. The Pearson correlation between the event lengths and the 50th, 75th, 90th percentiles and maximum PM_{10} concentration during the events are 0.19, 0.25, 0.33 and 0.44, respectively.

Dust events are unevenly distributed around the year as can be seen in Fig. 7. The months June–September are almost completely devoid of dust presence in the study area. The number of events and dust hours significantly increase from October, peak around the end of winter and decrease throughout the spring. In February and March there are relatively more dust hours compared to the corresponding mean number of dust events. This suggests that the winter events are on average longer than events during the rest of the year. The annual dust variability can be seen in Fig. 8, which shows the annual number of dust events and the annual percentage of dust–related time points. Unlike the well correlated calendar month distributions of dust events and dust hours (Fig. 7), there is a weak correlation (0.37) between the two curves in Fig 8. This demonstrates that analyses based on different measures of the dust phenomenon may yield very different long–term assessments of its trend and variability.

5.3. Spatial analysis

Based on the dust presence flagging scheme and the dust event definition (see Section 4), a possible measure of the spatial extent of a dust event is the number of stations simultaneously reporting dust at any time points of its duration. The existence of missing
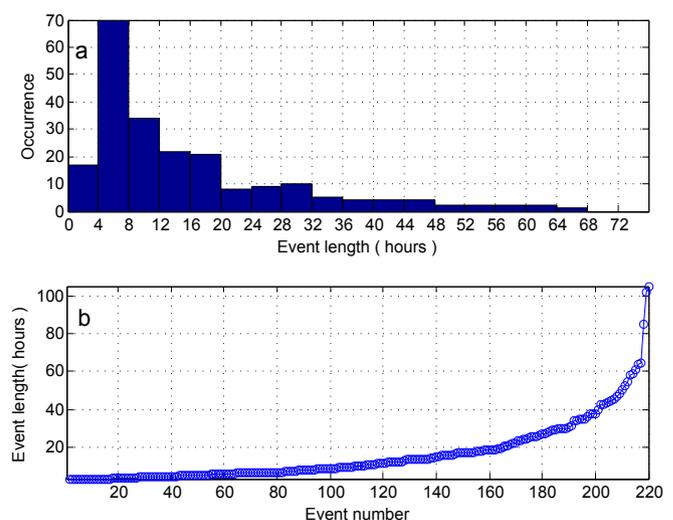

Fig. 5. The number of 12:00 UTC time points which were classified to each of the synoptic system groups based on the data from station 9. (a) The sample of normal time points. (b) The sample of time points flagged as associated with dust. The synoptic systems groups are: Red Sea Troughs (RST), Persian Trough (PT), High to the West (HW), Siberian Highs (SH), Winter Lows (WL) and Sharav Low (SL). See details in Alpert et al. (2004).

Fig. 6. Dust events lengths based on the dust flagging scheme using the data from station 9. (a) Histogram of dust event lengths up to 72 h. (b) All the dust event lengths sorted by their length.
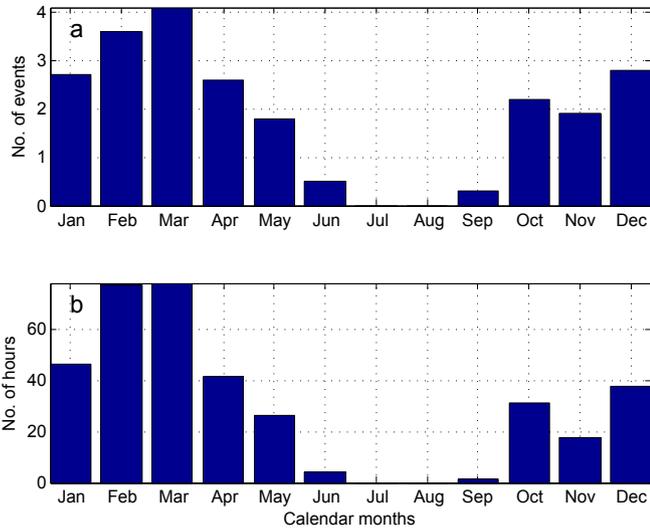
Fig. 7. The annual cycle of dust presence based on the dust flagging scheme using the data from station 9. (a) The mean number of dust events per calendar month. (b) The mean number of dust-related hours per calendar month.
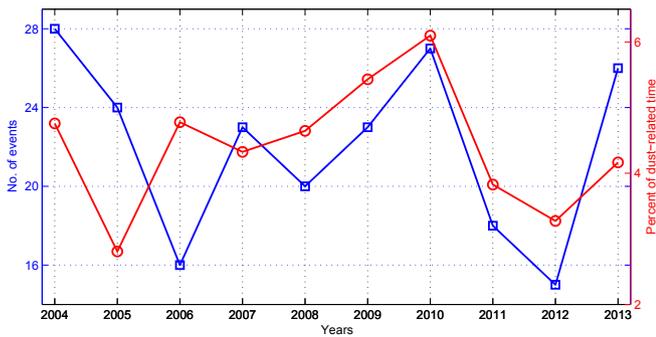

Fig. 8. The inter-annual variability of the dust phenomenon, depicted by the number of dust events and the percentage of dust−related time points.
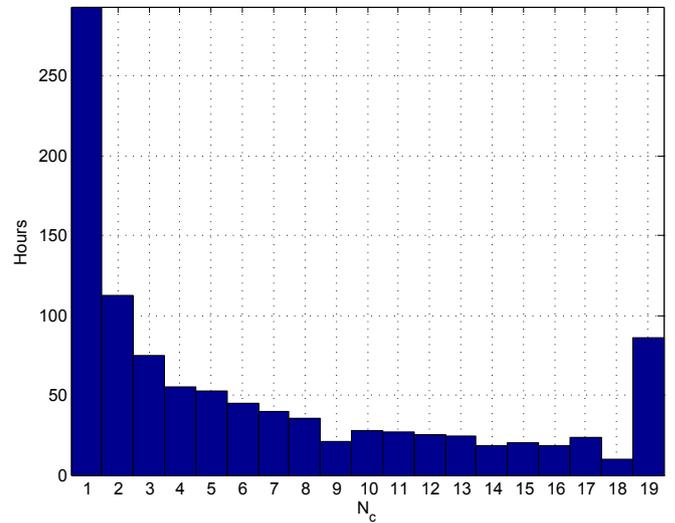

Fig. 9. The empirical probability distribution of the mean annual number of PM_{10} stations simultaneously reporting dust during the study period.

values in the PM_{10} data series might introduce bias, for example, the true spatial extent during two time points when 15 out of the 19 stations in the study area reported dust presence is probably not similar if in one of them only 15 stations observed valid data (ie 100% dust presence in the available observations) while in the other all 19 stations observed valid data (79% dust presence). During the study period, only in 27.7% of the time points all the stations reported valid values. The corresponding numbers for 18, 17, 16 and 15 stations reporting valid values are 31.6%, 22.4%, 10.7% and 4.5% (together above 96% of the time points). In the following, results related to the number of stations simultaneously reporting dust presence will consider the number N_c calculated as

$$N_c = floor\{mp/n\}, \qquad (2)$$

where p is the number of stations reporting dust presence at the time point, n is the corresponding number of valid observations, $m = 19$ is the number of stations in the monitoring array and *floor* $\{y\}$ is the mathematical operator yielding the largest integer not greater than y. For example, if at five different time points the numbers of dust−reporting stations are p = [3 12 15 16 18] while the corresponding numbers of valid observations are n = [17 18 16 18 19], then Nc = [3 12 17 16 18].

Fig. 9 shows the occurrence distribution of N_c during the study

period. Total coverage of the study area (ie $N_c = 19$) occurs during 85 h a year on average. It is more than the number of yearly hours for all cases apart from $N_c = 1,2$. This is expected though given that many dust events have large spatial extent, way beyond the spatial scale of the study area. With the exception of $N_c = 19$ case, occurrences of N_c tend to decrease as N_c increases. Very conspicuous is the number of occurrences of $N_c = 1$, which transpired on average 293 h a year. An important question is whether the cases of small N_c values are indeed dust−related events, locally high PM_{10} concentrations due to non−dust PM sources, or artefacts of the clustering scheme. Fig. 10 shows the 25th, 50th, 75th and 90th percentiles of the PM_{10} concentrations as a function of N_c. The statistics were calculated based on the PM_{10} concentrations at all the monitoring stations (whether they reported dust or not) for all the time points assigned with a certain N_c value. The most important and conspicuous feature of the curves in Fig. 10 is that the PM_{10} concentrations increase almost monotonically with N_c. The larger the
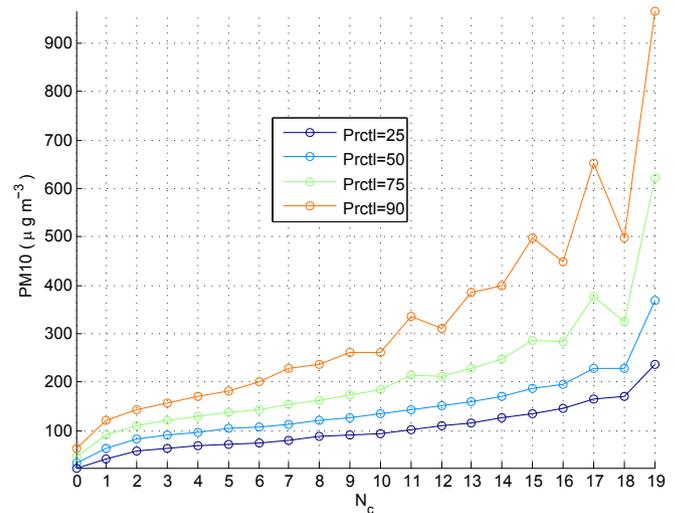

Fig. 10. Statistics of the PM_{10} concentrations as a function of N_c, the number of stations simultaneously reporting dust. The statistics for each N_c value were calculated using the data from all the stations at the time points assigned that N_c value.

spatial extent of a dust event, the higher are the associated PM_{10} concentrations. The correlation between N_c and the concentration statistics is around 0.9. A second important observation is the very steep increase of the PM_{10} concentration statistics between $N_c = 0$ and $N_c = 1$. The doubling of the statistics suggests that the $N_c = 1$ cases are not completely isolated occurrences of high PM_{10} concentrations. It is interesting to note that the contributions to the cases of $N_c = 1$ are not equally distributed among the stations. Station 4 contributed 89 h to $N_c = 1$, by far more than any other, followed by stations 19, 1, and 17 which contributed 34, 29 and 20 h, respectively (see Fig. SM7a for the contributions of all the other stations). Whereas no latitudinal trend in the distribution of the number of $N_c = 1$ cases is evident (eg station 1 is the most southern while stations 17 and 19 are among the most northern ones), we did infer a general weak tendency for decreasing dust presence from south to north using the correlations of the 95th and 99th PM_{10} concentration percentiles in the 10-year records (−0.47 and −0.61 with p−values of 0.0419 and 0.0056, for the 95th and 99th percentiles, respectively). Interestingly, the correlation between the stations' contributions to the N_c=1 cases and their elevation above sea level is high and very significant (0.87, p−value < 0.00001). This may result from atmospheric stratification and the presence of elevated dust levels aloft. The correlation of the station elevations and high PM_{10} concentrations is weak (0.34 and 0.62 for the 90th and 95th PM_{10} concentration percentiles of the 10-year records, respectively) so the dust presence scheme seems to be a sharper tool to study this issue.

The spatial variability of the aerosol concentrations during dust events is depicted in Fig. 11. A measure of the spatial variability of the concentrations for a certain value of N_c is the mean value of the absolute concentrations differences between pairs of observation stations, divided by the distance between them,

$$V(N_c) = \frac{1}{ML} \sum_{i=1}^{M} \sum_{j=1}^{L} C_{ij}/D_{ij}, \qquad (3)$$
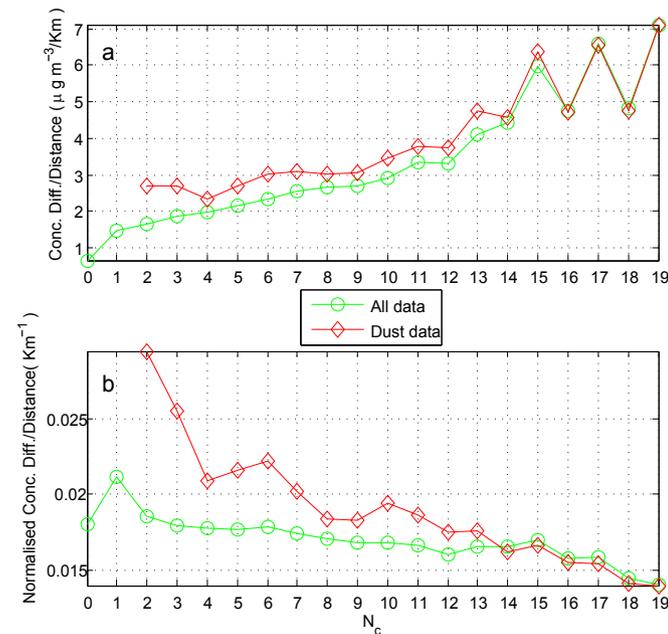
where C_{ij} is the absolute value of the PM_{10} concentration difference between each of the L pairs of observations at the *ith* out of M time points with N_c stations reporting dust, and D_{ij} is the corresponding distance. The quantity V as a function of N_c is shown in Fig. 11a. It is almost monotonically increasing with N_c, which means increasing nominal spatial variability in the PM_{10} concentrations with N_c. However, there usually is a positive correlation between the magnitudes of observations and their variability. Thus the increase in V with increasing N_c may be just a result of the fact that dust events with a large spatial extent (ie large N_c) are characterised by larger loads of PM (see Fig. 10). Fig. 11b shows $V(N_c)$ normalised by the mean PM_{10} concentrations. It is clear from Fig. 11b that in fact, the larger the extent of a dust event the more homogeneous are the PM_{10} concentrations across the region.

6. Summary and discussion

This paper explores the small scale spatiotemporal variability of naturally−occurring dust presence. The analyses are based on a rigorous dust detection scheme which minimises the spatial detection bias. The dust classification is performed using half−hourly temporal resolution and enables studying the fine temporal features of the dust phenomenon. The arbitrariness of the thresholds defining dust may have had some impact on the dust presence time and dust event lengths but the sensitivity of the main results to these thresholds is very weak.

Dust presence is clearly characterised by very distinct distribution of PM_{10} concentration magnitudes, $PM_{2.5}/PM_{10}$ concentration ratio, temperature, relative humidity, insolation, and by the associated synoptic systems. It occurs in the study area almost exclusively in between October to May, with a peak season in February and March. Dust occurrence is in between 2 and 6% of the time and about 15−30 dust events per annum. The correlation between these two measures of dust prevalence is weak due to the very wide range of dust event lengths, between three hours to more than four days, and the relative rarity of the dust phenomenon.

Dust covers the whole study area only about 85 h a year. Most of the time when dust is present in the region the coverage is only partial, with the number of dust hours increasing as the spatial extent of the dust presence decreases. The events of large spatial extent tend to be more intense and more spatially homogeneous in terms of PM concentrations, however, the length of dust events is only weakly correlated with their intensity. The SH and WL synoptic system groups, both typical winter phenomenon, are the ones most often associated with dust events of large spatial extent.

The majority of dust events (about 56%, see Fig. 6b) are not longer than 12 h. We thus believe that our strategy of employing a high temporal resolution of dust identification enables a better specification of the dust phenomenon. The very common use of daily resolution dust specification (eg Ganor at el., 2010; Krasnov et al., 2014) might result in misclassification and reduced power of the statistical analyses. For example, Ganor et al. (2010) reported elevated number of dust days during April occurrences of PT synoptic systems, not known as associated with dust. It is very probable that the dust in those cases was actually associated with a Sharav Low, a synoptic system which has a sharp peak of prevalence in April but which may be very short in duration and thus might not be registered as the synoptic system on its day of occurrence (Alpert et al., 2004). Another example is the daily−based dust presence definition of Krasnov et al. (2014) which was used by Vodonos et al. (2014) in a study of daily hospitalisations and their possible association with exacerbation of chronic obstructive pulmonary disease. Their daily mean 71 μgm^{-3} PM_{10} concentration threshold to define a day as associated with dust may clearly lead to exposure misclassification due to the prevalence of many short dust

Fig. 11. (a) The mean PM_{10} concentrations difference between pairs of stations, divided by their distance as a function of N_c. (b) The values of plate (a) normalised by the mean PM_{10} concentrations for each N_c value.

events. In Fig. 2c we depict such an event which straddles the days of 26-05-2005 and 27-05-2005. The mean PM$_{10}$ concentration in both days was above 71 $\mu g m^{-3}$. However, the exposure of the population was probably minimal as the event took place in between 23:00 and 02:00, and above average concentrations were observed only after 21:30 and before 06:00, hours when most people are at home and are less exposed to ambient PM.

Our results reveal some insights that have not been reported thus far and debunk some common claims. We have found no clear basis to the assumption made by Krasnov et al. (2014) that the majority of natural dust storms take place mainly during the daytime. Fig. SM8, produced using our high resolution dust detection scheme does not reveal any significant association of dust presence and the hour of the day. Any such association is haphazard and vary from station to station. We did confirm the claim in the literature of an elevated dust presence in the south compared to the north in Israel (eg Ganor and Foner, 2001; Krasnov et al., 2014). This association is weak though and we noted many instances of dust present in northern Israel but not in its southern parts. For example, during the study period 2004–2013, during 75 half–hourly time points all the southern and central stations (1–12) reported dust while none of the northern ones (13–19) did. However, the opposite case (stations 13–19 reporting dust but none of stations 1–12 does) also occurred during 41 time points. The clear presence of dust only in part of the study area and not at all elsewhere is intriguing. Dust is transported by synoptic phenomena with a spatial scale of about 1000 km but we notice that intra–regional variations on the tens of km spatial scale exist. The small scale variability in the dust plumes can be visually seen in Fig. SM6 in the Supplementary Material. A detailed and interesting glimpse into the conditions during various types of dust events is given in Figs. SM9 and SM10 of the Supplementary Material. The dust–detection methodology used in this paper enables studying the phenomena depicted in Figs. SM8, SM9 and SM10 quantitatively. Additional finding is the frequent prevalence of local dust events which can be detected only in a small number of stations. The strong positive correlation which we found between the stations' elevation above sea level and the number of such occurrences may serve as a hint about the processes by which the dust plumes arrive at the surface. As only twice–daily atmospheric stratification data were available, we could not follow this lead. We hope that future studies using higher resolution stratification data (eg Uzan and Alpert, 2012) will arrive at deeper understanding.

The Middle East and North Africa region is a large source of dust and it is also one of the most affected by dust storms. Much work has been carried out to study the issue of dust from the western Maghreb (eg Ozer et al., 2006) through the Arabian peninsula (Maghrabi et al., 2011) to the Persian Gulf (eg Hamidi et al., 2013). In most cases the studies concentrated on large dust events and the synoptic conditions that bring them about. Our work contributes to the understanding of the smaller scale variability in time and space and we hope that our methodology will be used by future studies of dust in the region.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://

dx.doi.org/10.1016/j.atmosenv.2015.08.058.

References

Asaf, D., Pedersen, D., Peleg, M., Matveev, V., Luria, M., 2008. Evaluation of background levels of air pollutants over Israel. Atmos. Environ. 42, 8453–8463.

Alpert, P., Osetinsky, I., Ziv, B., Shafir, H., 2004. Semi–objective classification for daily synoptic system: application to the eastern Mediterranean climate change. Int. J. Climatol. 24, 1001–1011.

Dadvand, P., Basagana, X., Figueras, F., Amoly, E., Tobias, A., de Nazelle, A., Querol, X., Sunyer, J., Nieuwenhuijsen, M.J., 2011. Saharan dust episodes and pregnancy. J. Environ. Monit. 13, 3222–3228.

Escudero, M., Castillo, S., Querol, X., Avila, A., Alarcon, M., Viana, M.M., Alastuey, A., Cuevas, E., Rodríguez, S., 2005. Wet and dry African dust episodes over Eastern Spain. J. Geophys. Res. 110 http://dx.doi.org/10.1029/2004JD004731.

Escudero, M., Querol, X., Avila, A., Cuevas, E., 2007a. Origin of the exceedances of the European daily PM limit value in regional background areas of Spain. Atmos. Environ. 41, 730744.

Escudero, M., Querol, X., Pey, J., Alastuey, A., Pérez, N., Ferreira, F., Cuevas, E., Rodríguez, S., Alonso, S., 2007b. A methodology for the quantification of the net African dust load in air quality monitoring networks. Atmos. Environ. 41, 55165524.

Evan, A.T., Heidinger, A.K., Pavolonis, M.J., 2006. Development of a new over–water advanced very high resolution radiometer dust detection algorithm. Int. J. Remote Sens. 27, 3903–3924.

Ganor, E., Foner, H.A., 2001. Mineral dust concentrations, deposition fluxes and deposition velocities in dust episodes over Israel. J. Geophys. Res. 106, 1843118437.

Ganor, E., Stupp, A., Alpert, P., 2009. A method to determine the effect of mineral dust aerosols on air quality. Atmos. Environ. 43, 54635468.

Ganor, E., Osetinsky, I., Stupp, A., Alpert, P., 2010. Increasing trend of African dust, over 49 years, in the eastern Mediterranean. J. Geophys. Res. 115, D07201. http://dx.doi.org/10.1029/2009JD012500.

Goudie, A.S., Middleton, N.J., 2006. Desert Dust in the Global System. Springer Verlag, Heidelberg, p. 280.

Goudie, A.S., 2014. Desert dust and human health disorders. Environ. Int. 63, 101–113.

Hamidi, M., Kavianpour, M.R., Shao, Y., 2013. Synoptic analysis of dust storms in the Middle East. Asia–Pacific J. Atmos. Sci. 49, 279–286.

Hastie, Tibshirani, Friedman, J., 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer, New York, p. 745.

Israelevich, P., Ganor, E., Alpert, P., Kishcha, P., Stupp, A., 2012. Predominant transport paths of Saharan dust over the Mediterranean Sea to Europe. J. Geophys. Reaserch 117, D02205. http://dx.doi.org/10.1029/2011JD016482.

Jiménez-Guerrero, P., Pérez, C., Jorba, O., Baldasano, J., 2008. Contribution of Saharan dust in an integrated air quality system and its on–line assessment. Geophys. Res. Lett. 35, L03814.

Karanasiou, A., Moreno, N., Moreno, T., Viana, M., de Leeuw, F., Querol, X., 2012. Health effects from Sahara dust episodes in Europe: literature review and research gaps. Environ. Int. 47, 10714.

Koçak, M., Mihalopoulos, N., Kubilay, N., 2007. Contributions of natural sources to high PM$_{10}$ and PM$_{2.5}$ events in the eastern Mediterranean. Atmos. Environ. 41, 38063818.

Krasnov, H., Katra, I., Koutrakis, P., Friger, M.D., 2014. Contribution of dust storms to PM$_{10}$ levels in an urban arid environment. J. Air & Waste Manag. 64, 89–94.

Liao, T.W., 2005. Clustering of time series data–a survey. Pattern Recognit. 38, 1857–1874.

Maghrabi, A., Alharbi, B., Tapper, N., 2011. Impact of the March 2009 dust event in Saudi Arabia on aerosol optical properties, meteorological parameters, sky temperature and emissivity. Atmos. Environ. 45, 21642173.

Mallone, S., Stafoggia, M., Faustini, A., Gobbi, G.P., Matcom, A., Forastiere, F., 2011. Saharan dust and associations between particulate matter and daily mortality in Rome, Italy. Environ. Health Perspect. 119, 1409–1414.

Matlab 8.1.0.604 (R2013a), 2013. The MathWorks Inc., Natick, Massachusetts.

Mitsakou, C., Kallos, G., Papantoniou, N., Spyrou, C., Solomos, S., Astitha, M., Housiadas, C., 2008. Saharan dust levels in Greece and received inhalation doses. Atmos. Chem. Phys. 8, 71817192.

Ozer, P., Laghdaf, M.B.O.M., Lemine, S.O.M., Gassani, J., 2006. Estimation of air quality degradation due to Saharan dust at Nouakchott, Mauritania, from horizontal visibility data. Water Air Soil Pollut. 178, 79–87.

Pey, J., Querol, X., Alastuey, A., Forastiere, F., Stafoggia, M., 2013. African dust outbreaks over the Mediterranean Basin during 20012011: PM$_{10}$ concentrations, phenomenology and trends, and its relation with synoptic and mesoscale meteorology. Atmos. Chem. Phys. 13, 1395–1410.

Prospero, J.M., Ginoux, P., Torres, O., Nicholson, S.E., Gill, T.E., 2002. Environmental characterization of global sources of atmospheric soil dust identified with the nimbus 7 total ozone mapping spectrometer (TOMS) absorbing aerosol product. Rev. Geophys. 40, 2–31.

Querol, X., Pey, J., Pandolfi, M., Alastuey, A., Cusack, M., Pérez, N., Moreno, T., Viana, M., Mihalopoulos, N., Kallos, G., Kleanthous, S., 2009. African dust contributions to mean ambient PM$_{10}$ mass–levels across the Mediterranean Basin. Atmos. Environ. 43, 4266–4277.

Rodríguez, S., Querol, X., Alastuey, A., Plana, F., 2002. Sources and processes affecting levels and composition of atmospheric aerosol in the Western

Mediterranean. J. Geophys. Res. 107 (D24), 4777.

Uzan, L., Alpert, P., 2012. The coastal boundary layer and air pollution–a high temporal resolution analysis in the East Mediterranean coast. open Atmos. Sci. J. 6, 9–18.

Viana, M., Salvador, P., Artinano, B., Querol, X., Alastuey, A., Pey, J., Latz, A.J., Cabanas, M., Moreno, T., Dos Santos, S.G., Herce, M.D., Hernandez, P.D., Garcia, D.R., Fernandez-Patier, R., 2010. Assessing the performance of methods to detect and quantify African dust in airborne particulates. Environ. Sci.

Technol. 44, 88148820. http://dx.doi.org/10.1021/es1022625.

Vodonos, A., Friger, M., Katra, I., Avnon, L., Krasnov, H., Koutrakis, P., Schwartz, J., Lior, O., Novack, V., 2014. The impact of desert dust exposures on hospitalizations due to exacerbation of chronic obstructive pulmonary disease. Air Qual. Atmos. Health 7, 433–439.

Zhang, P., Lu, N., Hu, X., Dong, C., 2006. Identification and physical retrieval of dust storm using three MODIS thermal IR channels. Glob. Planet. Change 52, 197–206.