Research Article

Monitoring and Forecasting Air Pollution Levels by Exploiting Satellite, Ground-Based, and Synoptic Data, Elaborated with Regression Models

Silas Michaelides,¹ Dimitris Paronis,² Adrianos Retalis,³ and Filippos Tymvios^{1,4}

¹The Cyprus Institute, 2121 Aglantzia, Nicosia, Cyprus

²Institute for Astronomy, Astrophysics, Space Application & Remote Sensing, National Observatory of Athens, 15236 Athens, Greece

³Institute for Environmental Research and Sustainable Development, National Observatory of Athens, 15236 Athens, Greece

⁴Department of Meteorology, 1086 Nicosia, Cyprus

Correspondence should be addressed to Silas Michaelides; silas.michaelides@gmail.com

Received 1 January 2017; Accepted 10 April 2017; Published 7 May 2017

Academic Editor: Gabriele Curci

Copyright © 2017 Silas Michaelides et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper presents some of the results of a project that aimed at the design and implementation of a system for the spatial mapping and forecasting the temporal evolution of air pollution from dust transport from the Sahara Desert into the eastern Mediterranean and secondarily from anthropogenic sources, focusing over Cyprus. Monitoring air pollution (aerosols) in near real-time is accomplished by using spaceborne and in situ platforms. The results of the development of a system for forecasting pollution levels in terms of particulate matter concentrations are presented. The aim of the present study is to utilize the recorded PM₁₀ (particulate matter with aerodynamic diameter less than 10 μm) ground measurements, Aerosol Optical Depth retrievals from satellite, and the prevailing synoptic conditions established by Artificial Neural Networks, in order to develop regression models that will be able to predict the spatial and temporal variability of PM₁₀ in Cyprus. The core of the forecasting system comprises an appropriately designed neural classification system which clusters synoptic maps, Aerosol Optical Depth data from the Aqua satellite, and ground measurements of particulate matter. By exploiting the above resources, statistical models for forecasting pollution levels were developed.

1. Introduction

Air quality has attained considerable scientific interest during the last decades because it directly or indirectly impacts a large number of social and economic activities but above all it adversely affects human health and well-being. For example, Middleton et al. [1] have estimated that, in Cyprus, for every 10 $\mu\text{g m}^{-3}$ increase in daily average PM₁₀ concentrations (particulate matter with diameter less than 10 μm), there was a 0.9% increase in all-cause and 1.2% increase in cardiovascular admissions to medical establishments. However, contrary to what might have been expected in this country, recent findings indicate that air pollution does not appear to have a significant effect on the temperature–mortality relation [2].

Numerous studies deal with aerosol monitoring and assessment both for naturally and anthropogenically induced aerosol particles. To this end, in situ measurements and satellite remote sensing are widely used for air quality monitoring on the regional and global scale. Conventional techniques of particulate matter measurements define regional air quality on a rather local sense; hence, dense networks of particulate matter recording ground stations are required to assess the spatial variations of surface air quality. Conversely, satellite remote sensing, in terms of Aerosol Optical Depth (AOD) measurements, could provide information with the required spatial coverage for the areas under examination (a list of symbols and abbreviations is given in Abbreviations). AOD, which is a quantitative measure of the amount of depletion that a beam of solar radiation undergoes as it passes through

the atmosphere, is one of the most important properties of atmospheric aerosols.

Several studies have focused on the use of satellite retrieved AOD as the main input parameter in models (simple linear regressive models, nonlinear and multilinear models) to predict the ambient particulate matter concentrations [3–16].

Attention has also been given to the influence of the prevailing synoptic conditions (in terms of synoptic patterns on weather charts) on the relationship between AOD and PM_{10} [17, 18].

Monitoring of air quality is an important application of remote sensing. In recent years, special emphasis is given to the application of methods for the detection and monitoring of air pollution from anthropogenic sources (e.g., industrial activity, road traffic in big cities). The international literature reports methods of measuring AOD over urban areas applicable to high spatial resolution satellite data [19–22], medium spatial resolution [23], and low spatial resolution [24–26]. Another component of air pollution, that of natural origin, which is particularly important in countries of the southeast Mediterranean, such as Cyprus which this study focuses on, is the transfer of dust particles from the Sahara Desert. Satellite remote sensing has been used extensively to study this phenomenon mainly above sea level [27].

In Cyprus, observations of precipitation and temperature have started more than a century ago; however, systematic recording, collecting, and archiving of a wider range of meteorological elements (temperature, humidity, pressure, visibility, cloud, etc.) started around the middle of the last century. In contrast, the recording of some pollutants from land-based air pollution monitoring systems is a more recent practice, having started about twenty years ago, initially in a rudimentary form but gradually evolving into a network of pollution measuring stations, covering mostly urban and industrial areas. Despite its great local significance for Cyprus, the systematic study of dust transportation from the surrounding deserts essentially started with the European Union funded MEDUSE (Mediterranean Dust Experiment) project in the mid 1990s; over the duration of that project, dust episodes were continuously recorded and documented by using synoptic analyses, satellite imagery, and dust deposition measurements [28]; also, the identification of synoptic situations leading to such dust transfer outbreaks was attempted. Within the same framework, an atmospheric model which included a subsystem (module) for dust transport was installed and operated [29]. More recently, several more focused studies have shed light on various aspects of dust transportation over the island of Cyprus [30–33].

The knowledge regarding the characteristics of a relationship between synoptic conditions and dust transport episodes is generally incomplete, since the complex nature of these episodes is not fully understood. However, recent studies focus on the analysis of synoptic conditions which are associated with dust storms in various geographical regions that are prone to such phenomena. Such studies have been those of Barkan and Alpert [34] who have analysed the synoptic events leading to dusty conditions in the Sahara. Fattahi et al. [35] classified the synoptic patterns associated

FIGURE 1: Locations of the four PM_{10} measuring stations.

with dust storms in Iran, while Al-Jumaily and Ibrahim [36] performed a similar analysis in Iraq. The association of dust events in the Middle East with synoptic patterns associated with dust was studied by Hamidi et al. [37], while Awad and Mashat [38] scrutinized the synoptic characteristics associated with widespread dust events in Saudi Arabia.

In the present study, the space-based monitoring of atmospheric pollution is based on the AOD retrievals by using the data from the MODIS (MODerate resolution Imaging Spectroradiometer) sensor onboard the Aqua satellite; the in situ system consists of a network of PM_{10} measuring stations. The forecasting system is based on the development of multivariate regression relationships which integrates the classes of the prevailing synoptic situation (as these are established by using Artificial Neural Networks, abbreviated hereafter as ANN), remotely sensed AOD data, and in situ ground measurements of PM_{10} ; these regression relationships can be used for estimating the air pollution levels as soon as the required predictors are available.

2. Data

The present study makes use of three types of data. The first type comprises measurements of PM_{10} at a number of ground recording stations; these in situ observations are recorded at regular times using appropriate monitoring equipment (for details of equipment used and observation procedures, see [39]). The second type of data consists of AOD measurements which are derived from the MODIS sensor onboard the Aqua satellite. Finally, the third set of data which are used in this study in order to determine the prevailing synoptic condition are based on the 500 hPa NCEP (National Centers for Environmental Prediction, USA) analyses. The three types of data are discussed below.

2.1. In Situ PM_{10} Measurements. The ground measurements of the concentration of aerosols (PM_{10}) were provided by the monitoring stations in Nicosia (city-traffic), Larnaca (city-traffic), Zygi (industrial), and Ayia Marina (background), the operating and maintenance responsibility of which is the responsibility of the Department of Labour Inspection of the Ministry of Labour and Social Insurance of Cyprus. These measurements were used as a reference in order to reflect the differences in pollution levels at different types of stations. Figure 1 shows the locations of these four stations.

It is worth mentioning at this point that the Ayia Marina monitoring station is located in an area which has relatively

FIGURE 2: Time series of the average daily values of PM_{10} concentrations for each of the four measuring stations (gaps in the time series are shown as zeros).

low local pollution sources; thus it is considered as an EMEP (i.e., operating within the framework of the protocol of the European Monitoring and Evaluation Program) Background Representative Station; for this reason, a large proportion of the PM_{10} measured at this station can be attributed to external sources (e.g., dust transportation).

In Figure 2, the time series of the daily average of PM_{10} concentration values are given for the period 2003–2005, for

each of the four stations of the ground network. Measuring PM_{10} is very important for the determination of dust events. In this context, a dust event is considered as an extreme PM_{10} concentration day, that is, a day when the average PM_{10} measurement exceeds the threshold of $50 \mu g m^{-3}$. In the three-year period mentioned above, 85 such dust deposition events were recorded (out of a total of 1096 days). It has been established that there exists a seasonal preference for

TABLE 1: Average value, standard deviation, and maximum value of PM_{10} concentrations calculated on the basis of hourly values in the period 2003–2005.

Station name	Average (2003–2005)	Standard deviation (2003–2005)	Minimum (2003–2005)	Maximum (2003–2005)
Ayia Marina	30	30	5	513
Larnaca	56	24	19	305
Nicosia	56	31	16	664
Zygi	45	29	8	479

dust events to occur [40]. Indeed, evidence supports the perception that Spring and Autumn are the two seasonal periods favoring dust episodes, whereas Summer appears to be suppressing these events; dust episodes are rather rare in Winter.

From Figure 2 we observe that, generally, the values corresponding to the measurements at the station of Ayia Marina are significantly less than the others, throughout the course of the three years under investigation. The stations of Nicosia and Larnaca have similar time series with no significant difference between them, while the station at Zygi also produces generally high values, being located in an industrial zone on the south coast of the island. We also note the existence of daily maxima with values exceeding $500 \mu g m^{-3}$ recorded at all stations and which are related to particular dust transfer events from the Sahara Desert. In Table 1, for each of the above stations, the respective average values are given together with the standard deviations, minima and maxima, derived from the hourly values over the three-year period of study. The values in this table reflect the lower average pollutant levels at the background station (Ayia Marina). The highest value is noted at Nicosia which is basically an urban station with intense road traffic.

2.2. Satellite Derived Aerosol Optical Depth Measurements.

The MODIS sensor is onboard the Earth Observing System (EOS) Terra and Aqua satellites. The available data are archived since April 2000, for Terra satellite, and since May 2002 for the Aqua satellite. The AOD measurements are acquired both above the ocean surface [41] and over land [42], through two independent algorithms. For the recovery of aerosols over oceans, the measured radiation values are used in six spectral regions (550–2100 nm) with a spatial resolution of 500 m. Aggregation from 500 m to 10 km resolution product allows the rigorous screening of clouds, avoiding gaps between the data and it produces a valuable end product. The accuracy of MODIS AOD data over land is expected to be $\Delta AOD = \pm 0.05 \pm 0.15$ AOD [42]. Especially, when in the composition of the aerosol dust is dominating, an additional deviation of the order of +10% can be noted.

For the needs of the present study, AOD MODIS level-2 data (MYD04_L2, Collection 6) with a resolution of $10 km \times 10 km$ for the years 2003–2005 from the Aqua satellite were used (for details see [43]). From these data, which included, on average, one image per day, AOD values were extracted for each station from the Optical_Depth_Land_And_Ocean Scientific Dataset. Spatial mean values were also calculated for the entire region of Cyprus for the period 2003–2005.

FIGURE 3: Mean AOD and PM_{10} values over the four stations (Ayia Marina, Larnaca, Nicosia, and Zygi) for the period 2003–2005 calculated from the common pairs of the collocated datasets.

In Figure 3, the mean values of both Aqua AOD and PM_{10} are given for the period 2003–2005, based on the two collocated datasets and the available data pairs. It is noticed that at Ayia Marina (EMEP, background station) both the AOD and PM_{10} concentration values are overall lower than the respective values at the other three sites. We also observe that, for the 2003–2005 period, the spatial patterns for AOD and PM_{10} data are similar as regards the monthly average values.

Figure 4 displays the monthly average values of AOD in the 2003–2005 period, over the four ground measuring stations. Spatial mean and standard deviation AOD values were calculated over Cyprus for the period 2003–2005 (see Figure 5).

The AOD values of Figure 5 show a peak in April and a secondary peak in September. The observed maxima are due to dust transport phenomena mostly from the Sahara which have a greater incidence in the period spring-summer. During the period November to February, all values are smaller than the values during the remaining months of the year. It is noteworthy that the average minimum values of the AOD for this period are around 0.15 or less.

2.3. *Synoptic Map Data.* It is very important for forecasters to identify on these maps specific geometric configurations which can be used to characterize the synoptic patterns of the atmosphere. An initial attempt of clustering the synoptic conditions was made by Lamb [44], while in the literature there are numerous classification methods [45]. One of the

FIGURE 4: Monthly average values of Aqua AOD in the 2003–2005 period corresponding to each of the four ground measuring stations.

FIGURE 5: Spatial mean and standard deviation of Aqua AOD values over Cyprus for the period 2003–2005.

purposes of this paper is to present a relatively new classification methodology of synoptic conditions by using ANN. More specifically, ANN of Kohonen type (see [46] for details) were used for the classification of the geodynamic height distributions at 500 hPa. As a result of this classification, average synoptic patterns were defined for each class by averaging the spatial distributions corresponding to the days in each class.

The data used for the establishment of the classification and subsequent clustering of upper-air synoptic patterns are the grid values of the 500 hPa isobaric level at 1200 UTC of each day, for the period from 1 January 1980 to 31 December 2005, as they are archived by NCEP. The geographical area adopted covers the European continent, the Middle East, and North Africa, as this is determined by latitude circles 20°N and 60°N and meridians 20°W and 40°E. The area is covered by a rectangular grid with each grid box having dimensions $2.5^\circ \times 2.5^\circ$; hence, the synoptic pattern at each time (i.e., 1200 UTC, every day) is defined by $7 \times 25 = 425$ grid points.

The classification of synoptic patterns established by using Kohonen's Self Organized Map algorithm has been used previously to relate specific patterns with extreme rainfall events [47] and heat events [48]. In the present study, it is shown how this neural network classification approach can be used for relating specific patterns at the 500 hPa isobaric surface with levels of PM_{10} in the Cyprus.

3. Methodologies

The methodology adopted in the present research aims at developing appropriate but simple regression models that can be used to predict today's PM_{10} concentration levels. Such predictions produce an estimate of the air quality (in terms of PM_{10} levels) which can be useful to local decision makers responsible for the day-to-day operational application of related legal frameworks, including issuance of warnings for pollution exceedance levels and adoption of appropriate measures.

The regression approach adopted here is a simple to develop objective forecasting method that lessens potential biases which may arise from human subjectivity. The required computational effort can be reduced by using standard software and widely available computing resources. Once the data on which the regression relationships will be built upon are collected and archived, several regression relationships may be scrutinized and an adequate formulation can be selected.

The data upon which the regression relationships are to be constructed are subject to an appropriate clustering, bearing in mind the prevailing synoptic conditions over the wider Mediterranean area. This was considered as an important ingredient in the development of the forecasting method, because experience has indicated that, in Cyprus, increased levels of pollution resulting from the transfer of particulate matter from the adjacent desert areas are greatly influenced by the prevailing weather patterns, as explained above.

Bearing in mind the above, the methodology adopted and used in this paper comprises a three-step procedure:

- (i) Development of (spatial) synoptic classification algorithms using ANN;
- (ii) Clustering of days in the period for which the regression will be constructed on the basis of the above synoptic classification;
- (iii) Construction of a linear regression with suitable predictors and today's PM_{10} as the predictand, for each cluster of days. Such suitable predictors are considered to be the measured in situ particulates (PM_{10}) and MODIS AOD data.

Apparently, the above steps are integrated into a forecasting system that makes use of both spaceborne and in situ measurements, on the one hand, and the prevailing synoptic conditions, on the other hand.

The proposed system uses the existing infrastructures of institutions in Cyprus. The benefits from its development are important, not only for the scientific and research community, but mainly for its operational level. The early notification

or warning of air pollution levels could contribute to the diffusion of information to the influenced targets (public, authorities, industries, etc.) so that the associated effects can be minimized and any necessary measures can be taken timely. Furthermore, the proposed approach is expected to contribute to the integration of our knowledge on the issue of large scale dust transport phenomena, aiming at the improvement of the authorities' abilities to take timely action and effectively deal with such episodes.

The automation of the processes of data transmission, diffusion, and mapping derived from the measurements of the atmospheric pollution network, along with the simultaneous mapping of other relative data, could provide to the final end-user an additional tool of modern technology to adjust his existing infrastructure.

3.1. Objective Classification and Clustering of Synoptic Conditions. The systematic use of synoptic charts maps dates back to the beginnings of modern meteorological practice. At regular intervals, surface and upper-air synoptic stations, scattered around the world, generate meteorological observations of specific meteorological parameters, at the Earth's surface and at the upper atmospheric levels, respectively. Normally, surface stations reports are made at three-hour intervals; the upper-air monitoring stations reporting at six-hour intervals record geodynamic height, wind speed and direction, temperature, and humidity at specific isobaric levels. Traditionally, for the analysis and determination of the synoptic situation, weather forecasters use contour charts at various atmospheric levels, but very often the 500 hPa level is adopted (roughly determining the level that splits the atmospheric mass into two halves and also representing the level nondivergence for mid latitudes).

Meteorologists often identify on these charts specific configurations, or patterns, which can be used to characterize the synoptic state of the atmosphere. In synoptic climatology, relationships between synoptic-scale atmospheric circulations and local-scale surface environmental variables are sought [49]. Such relationships can be studied by considering classification-based synoptic climatologies in which the atmospheric circulation conditions are stratified into a set of discrete synoptic chart patterns: days belonging to a given class share similar contour patterns. Thus, a very large number of circulation patterns can be communicated by a smaller number of descriptive classes.

Synoptic chart classifications can be produced either manually or by an automated technique. On the one hand, in manually determined classifications, an expert subjectively decides how the synoptic charts are grouped but the main effort in this exercise is to ensure that the classes that are established have meteorological significance. On the other hand, automated techniques are objective; such automated (unsupervised) chart classification methods, although less time consuming to generate and straightforwardly replicable, may fail to recognize important links between synoptic-scale circulation conditions and environmental variables at the surface.

Principal Component Analysis has been widely used by various authors to classify synoptic or mesoscale patterns

[50–52], while in the literature there are numerous objective classification methods [53].

In this paper, a relatively new objective classification methodology of synoptic states which makes use of ANN is implemented. Details of the concept and mathematical aspects of such a neural network approach are given by Michaelides et al. [54, 55]; here a brief description of the methodology which is adopted to generate the synoptic classes, namely, the Kohonen Self Organized Map, is given. The number of output elements (i.e., classes) was set to various “realistic” values. The selection of the number of classes is subjective, with the decision dependent, to a large extent, on experience gained in the field [56]. The input vector is formed by the values at the 425 grid points, as explained above. The (gridded) table two input dimensions for each geometric formation height are transformed into one-dimensional vector and supplies the neural network. The input vector is that formed by the values in the 425 grid points. Since the exact number of classes is not a priori known, for the present study four identical systems were designed with the ability to categorize the inputs (i.e., the synoptic patterns) into 20, 24, 30, and 35 classes. A detailed description of how the neural network is implemented can be found in the literature [54, 55]; hence, there is no need to duplicate it here. In the present paper, only the implementation with 24 classes is presented.

As a result of the above described classification, the synoptic analyses in the 1980–2005 period have been clustered and the average spatial distribution of all the members in each class is subsequently used to determine the average patterns for each class, as shown in Figure 6.

The frequency with which each of the 24 classes appears in the 1980–2005 period is shown in Figure 7. It is noted that classes 9 and 24 are the two most frequent, whereas classes 1 and 15 are the two least frequent.

Based on the 24-class classification, each day in the three-year period 2003–2005 has been assigned to a particular class. The assignment of a unique class to each of the days in this three-year period is shown in Figure 8. This diagram reveals a quasi-cyclic distribution of the classes which is interpreted as a seasonal signal which is very strong and it actually masks the synoptic-scale signal [57].

It can be noted that certain classes exhibit such a preference during particular seasons of the year, whereas some other classes do not seem to appear at all during the same seasons. This is more clearly seen in Figure 9.

3.2. Linear Regression Models. The main objective of the paper is to develop linear regression models for predicting particulate matter concentration levels with a diameter less than $10\ \mu\text{m}$ (PM_{10}) from MODIS Aqua AOD data and measurements of PM_{10} , taking into account the prevailing synoptic condition. Considering the graphs of Figure 1, it is obvious that the time series of the PM_{10} measurements have a strong persistence component. Hence, measurements of PM_{10} from previous days should comprise suitable candidates as predictors. Several combinations of time-dependent AOD and PM_{10} predictors were tested in this study.

FIGURE 6: Continued.

FIGURE 6: Continued.

FIGURE 6: Average spatial distribution of the 500 hPa isobaric surface, associated with each of the 24 classes. Contours are drawn for every 80 geopotential meters.

FIGURE 7: Frequency of appearance of the 24 synoptic classes in the 1980–2005 period.

4. Results

As part of the research effort, various regression models have been tested for each station and for each cluster of days (i.e., days belonging to the same class), separately, and for all cases (i.e., irrespective of the cluster that each case belongs to). The predictand (dependent variable) in all of these regression relationships is the PM_{10} concentration at the time of the MODIS (Aqua) overpass over the island of Cyprus.

Two of these regression models are described below, followed by a discussion on applying them to the available data.

4.1. Simple Linear Regression. In the simplest approach, the AOD recorded on day d (i.e., the current day) is used as the

FIGURE 8: Daily synoptic classification with 24 synoptic classes in the period 2003–2005.

independent variable in a simple linear regression model. The dependent variable is the PM_{10} measured on the same day d . The form of this relationship is

$$PM_{10}(d) = A_s + B_s \times AOD(d), \quad (1)$$

where A_s and B_s are the simple regression coefficients. The coefficient of determination (R^2) for this simple linear model is shown in Figure 10, for the Ayia Marina, Larnaca, Nicosia, and Zygi stations and for each synoptic condition (1–24); it is also shown for all cases, irrespective of synoptic condition (denoted as All in this figure). For some synoptic conditions the values of the coefficient of determination are not shown

FIGURE 9: Frequency of appearance of classes in seasons for the 2003–2005 period.

FIGURE 10: Coefficient of determination for the simple linear model for the Ayia Marina, Larnaca, Nicosia, and Zygi stations. It is shown separately for each of the synoptic conditions (1–24) and for all cases irrespective of the prevailing synoptic condition (shown as All).

due to the availability of only limited data for establishing a reliable relation.

It is noticed that when all cases are considered irrespective of the synoptic condition, the coefficient of determination is lower than 0.60. However, the coefficient of determination varies when considering separately the classes of the synoptic condition (classes 1–24). It is noticed that the values are significantly higher when the synoptic condition is considered in some cases (e.g., $R^2 = 0.8$, for Ayia Marina and synoptic condition 4), implying that the prevailing synoptic condition has an impact on the goodness of fit of the above PM_{10} –AOD relation.

4.2. Multivariate Linear Regression. The next model is a multivariate linear regression that considers as dependent variable the PM_{10} measured on day d (i.e., the current day) and as independent variables the PM_{10} of the previous day ($d-1$) and the AOD of the day d . The form of this multivariate regression is

$$PM_{10}(d) = A_m + B_{m1} \times PM_{10}(d-1) + B_{m2} \times AOD(d), \quad (2)$$

FIGURE 11: The coefficient of determination for the multivariate linear regression, for the Ayia Marina, Larnaca, Nicosia, and Zygi stations. It is shown separately for each of the synoptic conditions (1–24) and for all cases irrespective of the prevailing synoptic condition (shown as All).

where A_m , B_{m1} , and B_{m2} are the multiple linear regression coefficients.

The coefficient of determination (R^2) for the above multivariate linear regression, for the Ayia Marina, Larnaca, Nicosia, and Zygi stations, is shown in Figure 11, separately for each of the synoptic conditions (1–24) and for all cases, irrespective of the prevailing synoptic condition (shown as All). Values equal to one were found for synoptic conditions where the number of available cases was equal to three, while where values of the coefficient of determination are not shown this is due to limited data availability for establishing a reliable relation. It is noticed that, under certain synoptic conditions, the coefficient of determination exhibits a better goodness of fit, in contrast to the results presented in Figure 10. For example, for Ayia Marina station and synoptic condition 8, the value increases from 0.65 to 0.83.

The overall weaker coefficients of determination observed in Figures 10 and 11 when the regressions are used irrespective of the synoptic condition indicate that the goodness of fit can generally be improved when the prevailing synoptic condition is taken into account.

In the following, selected results are presented only for the Ayia Marina (EMEP) station. The coefficients of determination corresponding to the two linear models discussed above, namely, the simple and multiple regressions, are jointly summarized in Figure 12 for this station. It is noticed that the multiple regression model leads to an increase of the values of coefficient of determination, compared to the simple regression model, implying that the multiple regression yields a better goodness of fit than the simple regression. Therefore, the discussion that follows focuses on the multiple regression relationship only.

Figure 13 depicts the coefficient of determination for the Ayia Marina station in terms of the synoptic conditions ranked from the maximum (left) to the minimum (right) R^2 value for the multivariate regression model described by (2). The number shown above each bar in this plot indicates the number of the available data pairs (i.e., for which simultaneous PM_{10} and AOD measurements exist) for the specific synoptic condition. It should be noted that

FIGURE 12: Coefficient of determination for the two regressions for the Ayia Marina station in terms of the prevailing synoptic condition (1–24) and for all cases irrespective of the synoptic condition (shown as All).

FIGURE 13: Coefficient of determination for the multiple regression model for the Ayia Marina station, for each of the prevailing synoptic conditions (1–24) and for all cases irrespective of the synoptic condition (shown as All). The synoptic conditions are ranked on the basis of the value of the coefficient of determination, from the maximum (left) to the minimum (right) value. The number given above each bar is the number of the available data pairs for the respective synoptic condition.

the first two cases (namely, 3 and 7) comprise three pairs only and thus the value of the coefficient is biased to one (hence, these two cases are omitted from the discussion in the following paragraphs). As it can be observed, the synoptic conditions 11, 4, 8, and 17 exhibit the highest coefficients of determination, namely, 0.93, 0.88, 0.84, and 0.78, respectively. The coefficient of determination is also given for all the 461 available cases, irrespective of the synoptic condition (shown as All). Synoptic condition 5 is not shown in this figure because it is not associated with any available cases.

4.3. Model Performance Metrics and Statistical Scores. To measure the performance of the PM_{10} forecasting system (focusing on the multiple linear regression, as explained above), the following three performance metrics have been used, namely, the Index of Agreement (IA being basically, a standardized measure of the degree of model prediction error and varies between 0 and 1; a value of 1 indicates a perfect match, whereas 0 indicates no agreement at all), the Root

Mean Squared Error (RMSE), and the Mean Absolute Error (MAE), defined as follows:

$$IA = 1$$

$$= \frac{\sum_1^n (PM_{10}^m - PM_{10}^p)^2}{\sum_1^n (|PM_{10}^m - (1/n) \sum_1^n PM_{10}^m| + |PM_{10}^p - (1/n) \sum_1^n PM_{10}^p|)^2}, \quad (3)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_1^n (PM_{10}^m - PM_{10}^p)^2},$$

$$MAE = \frac{1}{n} \sum_1^n |PM_{10}^m - PM_{10}^p|,$$

where the superscripts m and p refer to the measured and predicted PM_{10} , respectively, and n is the total number of pairs used.

In Figure 14, selected scatterplots between measured PM_{10} (PM_{10}^m) and predicted PM_{10} (PM_{10}^p) are presented for the four synoptic conditions having the highest coefficient of determination, namely, 11, 4, 8, and 17 (ignoring 3 and 7, as explained above). A fitted linear relationship between the two quantities is also shown, together with the respective coefficient of determination.

The measured and predicted PM_{10} pairs were also used to construct a 2×2 contingency table (see [58]) as shown in Table 2, where the threshold of the daily average of $50 \mu g m^{-3}$ mentioned in Section 2 above was adopted. This is the threshold referred to in the relevant European Union Directive on ambient air quality and cleaner air for Europe (see Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008). The aim is to investigate how well the predicted extreme PM_{10} events correspond to the measured extreme PM_{10} events. Based on the definitions of a , b , c , and d in this table, the following three statistical scores (namely, POD, FAR, and BIAS) are calculated.

The probability of detection (POD) is defined as

$$POD = \frac{a}{a + c} \quad (4)$$

which equals the number of hits (number of cases with both the predicted and measured PM_{10} greater than the selected threshold) divided by the total number of cases with measured PM_{10} greater than the threshold. The POD ranges from 0 to 1, with the perfect score being 1; however, often high detection is associated with high number of false alarms.

The False Alarm Ratio (FAR) is expressed as

$$FAR = \frac{b}{a + b} \quad (5)$$

which is equal to the number of false alarms (number of cases for which the predicted PM_{10} is greater than the threshold but the measured PM_{10} is equal or less) divided by the total number of cases with predicted PM_{10} equal to or less than the threshold set. The FAR also ranges from 0 to 1; the perfect score is equal to 0.

The BIAS is computed as

$$BIAS = \frac{(a + b)}{(a + c)} \quad (6)$$

FIGURE 14: Scatterplots between measured and predicted PM₁₀ values for the four synoptic conditions with the highest coefficient of determination.

TABLE 2: Contingency table used for calculating the statistical scores.

		Measured PM ₁₀	
		>50 μg m ⁻³ (event observed)	≤50 μg m ⁻³ (event not observed)
Predicted PM ₁₀	>50 μg m ⁻³ (event predicted)	<i>a</i> (hits)	<i>b</i> (false alarms)
	≤50 μg m ⁻³ (event not predicted)	<i>c</i> (misses)	<i>d</i> (correct negatives)

representing the total number of cases with predicted PM₁₀ equal to or less than the threshold set, divided by the total number of cases with measured PM₁₀ greater than the threshold. It has a range of values from 0 to infinity. The desired value is 1 which indicates that the event (threshold in our case) is predicted exactly as often as it is observed.

As shown in Figure 13, the coefficient of determination varies with the prevailing synoptic condition. This observation points to the conclusion that the evaluation of the model as regards its potential to predict extreme PM₁₀ events (e.g., greater than 50 μg m⁻³) must take into the account this variability among the various synoptic conditions. For this reason, all performance metrics and statistical scores have been calculated following the ranking adopted in Figure 13.

Figure 15 summarizes the statistical scores for the Ayia Marina station, using all the measured and predicted values for PM₁₀ during the three-year period. In this figure, the statistical scores have been calculated for each one of the 24 synoptic conditions. Missing values (except for synoptic

condition 5 for which no data existed) indicate that the respective parameters are undefined, since the denominators in (4), (5), and (6) are equal to zero. This happens when no extreme events (PM₁₀ > 50 μg m⁻³) were recorded (when $a = 0$ and $c = 0$, POD and BIAS are undefined) and if, in addition, no false alarms have been produced ($b = 0$, FAR is undefined). This figure reveals that the probability of detection is higher than 0.5 in most cases. For example, for synoptic conditions 4, 8, and 17, the respective values are 0.75, 1.00, and 0.67 with corresponding FAR values equal to 0.40, 0.00, and 0.50. It is pointed out that the POD and FAR values estimated for all cases irrespective of the synoptic condition (shown as All) are 0.17 and 0.55, respectively.

Figure 16 summarizes the performance metrics for each one of the 24 synoptic conditions for the Ayia Marina station, using all the measured and predicted values for PM₁₀ during the three-year period. IA decreases as the coefficient of determination decreases (its value becomes less than 0.8 for synoptic condition 19 and to the right of this). The values of the RMSE and MAE range from 0 to 28.6 μg m⁻³ and from

FIGURE 15: Statistical scores for each synoptic condition (1–24) and considering all cases irrespective of the synoptic condition (shown as All). Synoptic conditions are ranked as explained in Figure 13.

FIGURE 16: Performance metrics for each synoptic condition (1–24) and considering all cases irrespective of the synoptic condition (shown as All). Synoptic conditions are ranked as explained in Figure 13.

0 to $17.9 \mu\text{g m}^{-3}$, respectively. The maximum value for these errors is found for the synoptic condition 4. This relatively high value can be explained by the fact that the maximum PM_{10} concentration ($293.7 \mu\text{g m}^{-3}$) of the entire dataset is found in relation to this synoptic condition. Thus, the high PM_{10} values associated with this synoptic condition bias the estimated errors to high values. Except for this high error, the RMSE remains lower than $15 \mu\text{g m}^{-3}$ for sixteen out of the twenty-four synoptic conditions. The respective value when considering all the available cases used is $20.4 \mu\text{g m}^{-3}$.

5. Concluding Remarks

Bearing in mind the available evidence [17], the main objective of the work undertaken in the present project is to develop a predictive system for estimating the particulate matter concentration levels with a diameter of less than $10 \mu\text{m}$ (PM_{10}) from MODIS AOD data and in situ measurements of PM_{10} , taking into account the prevailing synoptic condition. To achieve this goal, multivariate linear regression was applied. The results have shown that considering the

prevailing synoptic condition is an important ingredient of an air pollution forecasting methodology, based on linear regression. The resources used in the present research are easily accessible; they can be combined into a predictive system which may provide quite accurate predictions of imminent dust episodes.

The expected benefits from the development of such a system will be important not only to science and research, but primarily in operational applications [59, 60]. Timely issuance of warnings to both the public and the competent authorities of the state will minimize the impact of episodes of dust pollution through the adoption of appropriate measures [61, 62]. It is expected that the analysis adopted in the present paper can contribute to the filling of gaps in our knowledge about the dust transport phenomenon and to using this knowledge to improve the ability of competent authorities in dealing promptly and effectively with such incidents. Automating the transmission process, diffusion, and visualization of data from the air pollution measurement network and simultaneously displaying them with other relevant data, such as the prevailing weather conditions and AOD measurements, will be an additional tool of modern technology in the service of the citizen.

Multiple linear regression (e.g., [63–65]) and ANN (e.g., [66–70]) have extensively been employed in forecasting particulate matter concentration levels; in both distinct approaches, combinations of time-lagged particulate matter concentrations and meteorological parameters have been used as input to the forecasting models that were developed. The comparative advantages of such linear regression and ANN models have even been investigated by some researchers (e.g., [71–73]). In general, there is an agreement that ANN models are more accurate in predicting particulate matter levels. The superiority of ANN over regression is partly ascribed to the inherent nonlinearity between the predictand and the predictors used. Indeed, the ANN's ability to simulate nonlinear relationships involving heterogeneous variables is generally considered as an advantage.

Bearing in mind the above reference to the adoption of linear regression and ANN models in forecasting particulate matter concentration levels, the novelty of the present study is that the forecasting model that was developed follows a distinct hybrid approach, appropriately mixing ANN and linear regression modelling: ANN have been used in objectively classifying the synoptic conditions first; subsequently, multiple linear regression that utilizes satellite and in situ parameters as predictors was subsequently considered in conjunction with the prevailing synoptic condition.

Quality of a model is considered as its fitness for the purpose it was created, and obviously it is an issue that needs to be considered in any predictive scheme [74]. In this respect, Quality Assurance (QA) of a predictive model refers to an integrated system of activities involving several tasks, including its evaluation in order to ensure that the model is of the expected quality. The adoption of a complete QA scheme can be quite laborious, involving a great deal of systematic work (such as planning, documentation, and implementation), especially when the model is to be used operationally. Martinez et al. [75] have set four questions that

the issue of model evaluation is required to answer: (1) How well does the model predict maximum pollution levels? (2) How well does the model predict the number of exceedance of the relevant air quality standard? (3) How well do the fluctuations in predictions reproduce the fluctuations of the observed pollutant levels in time and space? (4) How close are computed concentrations and measurements? In this paper, it has been attempted to give answers to some of the above questions through a comprehensive presentation of performance metrics and scores.

The results of the evaluation of the proposed methodology have been presented in terms of the coefficient of determination (R^2), three statistical scores (POD, FAR, and BIAS), and three performance metrics (IA, RMSE, and MAE). In general, it seems that considering the prevailing synoptic condition which has objectively been determined (by the ANN subsystem) adds to the potential of the forecasting methodology presented.

It is envisaged that the adoption of more complex regression models for the prediction of the PM_{10} values will further enhance the proposed methodology. For example, linear-mixed effects models proposed in the literature [12, 76, 77] can be used for incorporating the synoptic condition as a variable in the model.

The research is intended to be further extended covering a more extensive database of input data so that the final product would have greater reliability in estimating air pollution levels in Cyprus. Also, verification based on independent data is very important in further developing the approach presented in the present study because it will permit a more robust estimation of the models' competency to predict PM_{10} (c.f., Section 4.3).

Abbreviations

A_m, B_{m1}, B_{m2} :	Multiple linear regression coefficients
A_s, B_s :	Simple linear regression coefficients
ANN:	Artificial Neural Networks
AOD:	Aerosol Optical Depth
BIAS:	Bias
d :	Current day
EMEP:	European Monitoring and Evaluation Program
EOS:	Earth Observing System
FAR:	False Alarm Ratio
IA:	Index of Agreement
MAE:	Mean Absolute Error
MODIS:	MODerate resolution Imaging Spectroradiometer
NCEP:	National Centers for Environmental Prediction, USA
PM_{10}^m :	Measured PM_{10}
PM_{10}^p :	Predicted PM_{10}
PM_{10} :	Particulate matter with aerodynamic diameter less than $10\ \mu m$
POD:	Probability of Detection
QA:	Quality Assurance
RSME:	Root Mean Square Error
R^2 :	Coefficient of determination.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The in situ data of particulate matter were provided by the Department of Labour Inspection of the Ministry of Labour and Social Insurance of Cyprus; the geopotential fields were provided by the National Centers for Environmental Prediction (NCEP) of the United States; the Aerosol Optical Depth data were provided by The Level-1 and Atmosphere Archive & Distribution System, Distributed Active Archive Center.

References

- [1] N. Middleton, P. Yiallourous, S. Kleanthous et al., "A 10-year time-series analysis of respiratory and cardiovascular morbidity in Nicosia, Cyprus: the effect of short-term changes in air pollution and dust storms," *Environmental Health: A Global Access Science Source*, vol. 7, article 39, 2008.
- [2] H. Tsangari, A. Paschalidou, S. Vardoulakis et al., "Human mortality in Cyprus: the role of temperature and particulate air pollution," *Regional Environmental Change*, vol. 16, pp. 1905–1913, 2016.
- [3] D. A. Chu, Y. J. Kaufman, G. Zibordi et al., "Global monitoring of air pollution over land from the Earth Observing System-Terra Moderate Resolution Imaging Spectroradiometer (MODIS)," *Journal of Geophysical Research: Atmospheres*, vol. 108, 4661, no. D21, 4661 pages, 2003.
- [4] S. Grgurić, J. Križan, G. Gašparac et al., "Relationship between MODIS based aerosol optical depth and PM_{10} over Croatia," *Central European Journal of Geosciences*, vol. 6, no. 1, pp. 2–16, 2014.
- [5] P. Gupta, S. A. Christopher, J. Wang, R. Gehrig, Y. Lee, and N. Kumar, "Satellite remote sensing of particulate matter and air quality assessment over global cities," *Atmospheric Environment*, vol. 40, no. 30, pp. 5880–5892, 2006.
- [6] P. Gupta and S. A. Christopher, "Seven year particulate matter air quality assessment from surface and satellite measurements," *Atmospheric Chemistry and Physics*, vol. 8, no. 12, pp. 3311–3324, 2008.
- [7] R. M. Hoff and S. A. Christopher, "Remote sensing of particulate pollution from space: have we reached the promised land?" *Journal of the Air & Waste Management Association*, vol. 59, no. 6, pp. 645–675, 2009.
- [8] L. Kong, J. Xin, W. Zhang, and Y. Wang, "The empirical correlations between $PM_{2.5}$, PM_{10} and AOD in the Beijing metropolitan region and the $PM_{2.5}$, PM_{10} distributions retrieved by MODIS," *Environmental Pollution*, vol. 216, pp. 350–360, 2016.
- [9] N. Kumar, A. Chu, and A. Foster, "An empirical relationship between $PM_{2.5}$ and aerosol optical depth in Delhi Metropolitan," *Atmospheric Environment*, vol. 41, no. 21, pp. 4492–4503, 2007.
- [10] N. Kumar, A. Chu, and A. Foster, "Remote sensing of ambient particles in Delhi and its environs: estimation and validation," *International Journal of Remote Sensing*, vol. 29, no. 12, pp. 3383–3405, 2008.

- [11] N. Kumar, A. D. Chu, A. D. Foster, T. Peters, and R. Willis, "Satellite remote sensing for developing time and space resolved estimates of ambient particulate in Cleveland, OH," *Aerosol Science and Technology*, vol. 45, no. 9, pp. 1090–1108, 2011.
- [12] H. J. Lee, Y. Liu, B. A. Coull, J. Schwartz, and P. Koutrakis, "A novel calibration approach of MODIS AOD data to predict PM_{2.5} concentrations," *Atmospheric Chemistry and Physics*, vol. 11, no. 15, pp. 7991–8002, 2011.
- [13] S. Seo, J. Kim, H. Lee et al., "Estimation of PM₁₀ concentrations over Seoul using multiple empirical models with AERONET and MODIS data collected during the DRAGON-Asia campaign," *Atmospheric Chemistry and Physics*, vol. 15, no. 1, pp. 319–334, 2015.
- [14] A. van Donkelaar, R. V. Martin, M. Brauer et al., "Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: development and application," *Environmental Health Perspectives*, vol. 118, no. 6, pp. 847–855, 2010.
- [15] J. Wang and S. A. Christopher, "Intercomparison between satellite derived aerosol optical thickness and pm_{2.5} mass: implications for air quality studies," *Geophysical Research Letters*, vol. 30, 2095, 2003.
- [16] X. Q. Yap and M. Hashim, "A robust calibration approach for PM₁₀ prediction from MODIS aerosol optical depth," *Atmospheric Chemistry and Physics*, vol. 13, no. 6, pp. 3517–3526, 2013.
- [17] M. Zeeshan and N. T. Kim Oanh, "Assessment of the relationship between satellite AOD and ground PM₁₀ measurement data considering synoptic meteorological patterns and Lidar data," *Science of the Total Environment*, vol. 473–474, pp. 609–618, 2014.
- [18] M. Schaap, A. Apituley, R. M. A. Timmermans, R. B. A. Koelemeijer, and G. de Leeuw, "Exploring the relation between aerosol optical depth and PM_{2.5} at Cabauw, the Netherlands," *Atmospheric Chemistry and Physics*, vol. 9, no. 3, pp. 909–925, 2009.
- [19] T. H. Lin, A. J. Chen, G. R. Liu, and T. H. Kuo, "Monitoring the atmospheric aerosol optical depth with SPOT data in complex terrain," *International Journal of Remote Sensing*, vol. 23, no. 4, pp. 647–659, 2002.
- [20] G.-R. Liu, T.-H. Lin, and A. J. Chen, "Estimation of Atmospheric Aerosol Depth with SPOT Satellite Data," in *In Proceedings of the 20th Asian Conference on Remote Sensing (ACRS '99)*, Hong Kong, China, 1999, <http://a-a-r-s.org/aars/proceeding/ACRS1999/Papers/ENV99-2.htm>.
- [21] A. Retalis, C. Cartalis, and E. Athanassiou, "Assessment of the distribution of aerosols in the area of Athens with the use of Landsat Thematic Mapper data," *International Journal of Remote Sensing*, vol. 20, no. 5, pp. 939–945, 1999.
- [22] D. G. Hadjimitsis, A. Retalis, and C. R. I. Clayton, "The assessment of atmospheric pollution using satellite remote sensing technology in large cities in the vicinity of airports," *Water, Air, Soil Pollution: Focus*, vol. 2, pp. 631–640, 2002.
- [23] Y. J. Kaufman, N. Gobron, B. Pinty, J.-L. Widlowski, and M. M. Verstraete, "Relationship between surface reflectance in the visible and mid-IR used in MODIS aerosol algorithm—Theory," *Geophysical Research Letters*, vol. 29, no. 23, pp. 31-1–31-4, 2002.
- [24] F. Dulac, D. Tanre, G. Bergametti, P. Buat-Menard, M. Desbois, and D. Sutton, "Assessment of the African airborne dust mass over the western Mediterranean Sea using Meteosat data," *Journal of Geophysical Research*, vol. 97, no. 2, pp. 2489–2506, 1992.
- [25] M. J. Costa, M. Cervino, E. Cattani et al., "Aerosol characterization and optical thickness retrievals using GOME and METEOSAT satellite data," *Meteorology and Atmospheric Physics*, vol. 81, no. 3–4, pp. 289–298, 2002.
- [26] A. Retalis, N. Sifakis, N. Grosso, D. Paronis, and D. Sarigiannis, "Aerosol optical thickness retrieval from AVHRR images over the Athens urban area," in *Proceedings of 2003 IEEE International Geoscience & Remote Sensing Symposium (IGARSS) '03*, pp. 2182–2184, Toulouse, France, July 2003.
- [27] C. Moulin, F. Guillard, F. Dulac, and C. E. Lambert, "Long-term daily monitoring of Saharan dust load over ocean using Meteosat ISCCP-B2 data 1. Methodology and preliminary results for 1983–1994 in the Mediterranean," *Journal of Geophysical Research Atmospheres*, vol. 102, no. 14, pp. 16947–16958, 1997.
- [28] S. Michaelides, P. Evripidou, and G. Kallos, "Monitoring and predicting Saharan Desert dust events in the eastern Mediterranean," *Weather*, vol. 54, no. 11, pp. 359–365, 1999.
- [29] A. Lachanas, P. Evripidou, and S. Michaelides, "Regional weather prediction on small network of workstations," in *Proceedings of the 24th EUROMICRO Conference*, pp. 1053–1060, Vasteras, Sweden, 1998.
- [30] A. Retalis and S. C. Michaelides, "Synergetic use of TERRA/MODIS imagery and meteorological data for studying aerosol dust events in Cyprus," *International Journal of Environment and Pollution*, vol. 36, no. 1–3, pp. 139–150, 2009.
- [31] A. Retalis, D. G. Hadjimitsis, S. Michaelides et al., "Comparison of aerosol optical thickness with in situ visibility data over Cyprus," *Natural Hazards and Earth System Science*, vol. 10, no. 3, pp. 421–428, 2010.
- [32] A. Nisantzi, D. G. Hadjimitsis, E. Akylas et al., "Study of air pollution with the use of modis data, lidar and sun photometers in Cyprus," in *Advances in Meteorology, Climatology and Atmospheric Physics*, Springer Atmospheric Sciences, pp. 1133–1139, Springer Berlin Heidelberg, Berlin, Germany, 2013.
- [33] D. G. Hadjimitsis, R.-E. Mamouri, A. Nisantzi et al., "Air pollution from space," in *Remote Sensing of Environment—Integrated Approaches*, D. G. Hadjimitsis, Ed., pp. 97–130, InTech, Rijeka, Croatia, 2013.
- [34] J. Barkan and P. Alpert, "Synoptic patterns associated with dusty and non-dusty seasons in the Sahara," *Theoretical and Applied Climatology*, vol. 94, no. 3–4, pp. 153–162, 2008.
- [35] E. Fattahi, K. Noohi, and H. Shiravand, "Study of dust storm synoptical patterns in southwest of Iran," *Desert*, vol. 17, pp. 49–55, 2012.
- [36] K. J. Al-Jumaily and M. K. Ibrahim, "Analysis of synoptic situation for dust storms in Iraq," *International Journal of Energy and Environment*, vol. 4, no. 5, pp. 851–858, 2013.
- [37] M. Hamidi, M. R. Kavianpour, and Y. Shao, "Synoptic analysis of dust storms in the Middle East," *Asia-Pacific Journal of Atmospheric Sciences*, vol. 49, no. 3, pp. 279–286, 2013.
- [38] A. Awad and A.-W. Mashat, "The synoptic patterns associated with spring widespread dusty days in central and eastern Saudi Arabia," *Atmosphere*, vol. 5, no. 4, pp. 889–913, 2014.
- [39] S. Achilleos, J. S. Evans, P. K. Yiallourou, S. Kleanthous, J. Schwartz, and P. Koutrakis, "PM₁₀ concentration levels at an urban and background site in Cyprus: The impact of urban sources and dust storms," *Journal of the Air and Waste Management Association*, vol. 64, no. 12, pp. 1352–1360, 2014.
- [40] S. Michaelides, F. Tymvios, S. Athanasatos, and M. Papadakis, "Trends of dust transport episodes in cyprus using a classification of synoptic types established with artificial neural networks," *Journal of Climatology*, vol. 2013, pp. 1–8, 2013.

- [41] D. Tanré, Y. J. Kaufman, M. Herman, and S. Mattoo, "Remote sensing of aerosol properties over oceans using the MODIS/EOS spectral radiances," *Journal of Geophysical Research D: Atmospheres*, vol. 102, no. 14, pp. 16971–16988, 1997.
- [42] L. A. Remer, Y. J. Kaufman, D. Tanré et al., "The MODIS aerosol algorithm, products, and validation," *Journal of the Atmospheric Sciences*, vol. 62, no. 4, pp. 947–973, 2005.
- [43] R. Levy, C. Hsu et al., *MODIS Atmosphere L2 Aerosol Product. NASA MODIS Adaptive Processing System*, Goddard Space Flight Center, MD, USA, 2015, http://dx.doi.org/10.5067/MODIS/MOD04_L2.006.
- [44] H. H. Lamb, "Types and spells of weather around the year in the British Isles: annual trends, seasonal structure of the year, singularities," *Quarterly Journal of the Royal Meteorological Society*, vol. 76, no. 330, pp. 393–429, 1950.
- [45] B. C. Hewitson and R. G. Crane, "Self-organizing maps: applications to synoptic climatology," *Climate Research*, vol. 22, no. 1, pp. 13–26, 2002.
- [46] T. Kohonen, "The self-organizing map," *Proceedings of the IEEE*, vol. 78, no. 9, pp. 1464–1480, 1990.
- [47] F. Tymvios, K. Savvidou, and S. C. Michaelides, "Association of geopotential height patterns with heavy rainfall events in Cyprus," *Advances in Geosciences*, vol. 23, pp. 73–78, 2010.
- [48] F. Tymvios, S. Michaelides, D. Charalambous et al., "Connection of heat events in Cyprus with synoptic upper air patterns," in *Advances in Meteorology, Climatology and Atmospheric Physics*, Springer Atmospheric Sciences, pp. 787–792, Springer Berlin Heidelberg, Berlin, Germany, 2013.
- [49] R. G. Barry and A. H. Perry, "Synoptic climatology. methods and applications," Methuen, London, 1973.
- [50] A. J. Cannon, P. H. Whitfield, and E. R. Lord, "Synoptic map-pattern classification using recursive partitioning and principal component analysis," *Monthly Weather Review*, vol. 130, no. 5, pp. 1187–1206, 2002.
- [51] E. García-Ortega, L. López, and J. L. Sánchez, "Atmospheric patterns associated with hailstorm days in the Ebro Valley, Spain," *Atmospheric Research*, vol. 100, no. 4, pp. 401–427, 2011.
- [52] A. Merino, E. García-Ortega, L. López, J. L. Sánchez, and A. M. Guerrero-Higueras, "Synoptic environment, mesoscale configurations and forecast parameters for hailstorms in Southwestern Europe," *Atmospheric Research*, vol. 122, pp. 183–198, 2013.
- [53] A. Philipp, J. Bartholy, C. Beck et al., "Cost733cat—a database of weather and circulation type classifications," *Physics and Chemistry of the Earth*, vol. 35, no. 9–12, pp. 360–373, 2010.
- [54] S. C. Michaelides, C. S. Pattichis, and G. Kleovoulou, "Classification of rainfall variability by using artificial neural networks," *International Journal of Climatology*, vol. 21, no. 11, pp. 1401–1414, 2001.
- [55] S. C. Michaelides, F. Liassidou, and C. N. Schizas, "Synoptic classification and establishment of analogues with artificial neural networks," *Pure and Applied Geophysics*, vol. 164, no. 6–7, pp. 1347–1364, 2007.
- [56] M. B. Richman, "On the application of cluster analysis to growing season precipitation data in North America east of the Rockies," *Journal of Climate*, vol. 8, no. 4, pp. 897–931, 1995.
- [57] B. Yarnal, *Synoptic climatology in environmental analysis*, Bellhaven Press, London, UK, 1993.
- [58] D. S. Wilks, *Statistical Methods in the Atmospheric Sciences, 2nd Edition*, Academic Press, Second Edition, 2006.
- [59] U. Dayan, J. Heffter, J. Miller, and G. Gutman, "Dust intrusion events into the Mediterranean basin," *Journal of Applied Meteorology*, vol. 30, no. 8, pp. 1185–1199, 1991.
- [60] T. Slini, A. Kaprara, K. Karatzas, and N. Moussiopoulos, "PM₁₀ forecasting for Thessaloniki, Greece," *Environmental Modelling and Software*, vol. 21, no. 4, pp. 559–565, 2006.
- [61] S. A. Morman and G. S. Plumlee, "Dust and human health," in *Mineral dust: A Key Player in the Earth System*, P. Knippertz and J. B. W. Stuut, Eds., pp. 385–409, Springer Science + Business, Berlin, Germany, 2014.
- [62] A. Namdeo and C. Stringer, "Investigating the relationship between air pollution, health and social deprivation in Leeds, UK," *Environment International*, vol. 34, pp. 585–591, 2008.
- [63] P. Goyal, A. T. Chan, and N. Jaiswal, "Statistical models for the prediction of respirable suspended particulate matter in urban cities," *Atmospheric Environment*, vol. 40, no. 11, pp. 2068–2077, 2006.
- [64] E. Stadlober, S. Hörmann, and B. Pfeiler, "Quality and performance of a PM₁₀ daily forecasting model," *Atmospheric Environment*, vol. 42, no. 5, pp. 1098–1109, 2008.
- [65] A. Vlachogianni, P. Kassomenos, A. Karppinen, S. Karakitsios, and J. Kukkonen, "Evaluation of a multiple regression model for the forecasting of the concentrations of NO_x and PM₁₀ in Athens and Helsinki," *Science of the Total Environment*, vol. 409, no. 8, pp. 1559–1571, 2011.
- [66] A. Kurt, B. Gulbagci, F. Karaca, and O. Alagha, "An online air pollution forecasting system using neural networks," *Environment International*, vol. 34, no. 5, pp. 592–598, 2008.
- [67] J. Hooyberghs, C. Mensink, G. Dumont, F. Fierens, and O. Brasseur, "A neural network forecast for daily average PM₁₀ concentrations in Belgium," *Atmospheric Environment*, vol. 39, no. 18, pp. 3279–3289, 2005.
- [68] G. Grivas and A. Chaloulakou, "Artificial neural network models for prediction of PM₁₀ hourly concentrations, in the Greater Area of Athens, Greece," *Atmospheric Environment*, vol. 40, no. 7, pp. 1216–1229, 2006.
- [69] P. Perez and J. Reyes, "An integrated neural network model for PM₁₀ forecasting," *Atmospheric Environment*, vol. 40, no. 16, pp. 2845–2851, 2006.
- [70] D. Jiang, Y. Zhang, X. Hu, Y. Zeng, J. Tan, and D. Shao, "Progress in developing an ANN model for air pollution index forecast," *Atmospheric Environment*, vol. 38, no. 40, pp. 7055–7064, 2004.
- [71] A. Chaloulakou, G. Grivas, and N. Spyrellis, "Neural network and multiple regression models for PM₁₀ prediction in Athens: a comparative assessment," *Journal of the Air and Waste Management Association*, vol. 53, no. 10, pp. 1183–1190, 2003.
- [72] A. Z. Ul-Saufie, A. S. Yahya, N. A. Ramli, and H. A. Hamid, "Comparison between multiple linear regression and feed forward back propagation neural network models for predicting PM₁₀ concentration level based on gaseous and meteorological parameters," *International Journal of Applied Science and Technology*, vol. 1, pp. 42–49, 2011.
- [73] A. K. Paschalidou, S. Karakitsios, S. Kleanthous, and P. A. Kassomenos, "Forecasting hourly PM₁₀ concentration in Cyprus through artificial neural networks and multiple regression models: Implications to local environmental management," *Environmental Science and Pollution Research*, vol. 18, no. 2, pp. 316–327, 2011.
- [74] C. Borrego, M. Schatzmann, and S. Galmarini, "Quality assurance of air pollution models," in *Air Quality in Cities—SATURN EUROTRAC-2 Subproject Final Report*, N. Moussiopoulos, Ed., pp. 155–183, Springer, Berlin, Germany, 2003.
- [75] J. R. Martinez, H. S. Javitz, R. E. Ruff, A. Valdes, K. C. Nitz, and W. F. Dabberdt, *Methodology for Evaluation Highway Air Pollution Dispersion Models*, Transportation Research

Board, Washington, DC, USA, 1981, http://onlinepubs.trb.org/Onlinepubs/nchrp/nchrp_rpt_245.pdf.

- [76] M. Sorek-Hamer, I. Kloog, P. Koutrakis et al., “Assessment of $PM_{2.5}$ concentrations over bright surfaces using MODIS satellite observations,” *Remote Sensing of Environment*, vol. 163, pp. 180–185, 2015.
- [77] A. Retalis, D. Paronis, and D. Katsanos, “Intercomparison between MODIS 3km aerosol optical depth product and ground PM_{10} measurements over Athens, Greece,” in *Proceedings of SPIE of Remote Sensing of Clouds and the Atmosphere XX*, Toulouse, France, September 2015.