

Article

Contributing towards Representative PM Data Coverage by Utilizing Artificial Neural Networks

Chris G. Tzanis *  and Anastasios Alimissis

Climate and Climatic Change Group, Section of Environmental Physics and Meteorology, Department of Physics, National and Kapodistrian University of Athens, 15784 Athens, Greece; alimiss@phys.uoa.gr

* Correspondence: chtzanis@phys.uoa.gr

Abstract: Atmospheric aerosol particles have a significant impact on both the climatic conditions and human health, especially in densely populated urban areas, where the particle concentrations in several cases can be extremely threatening (increased anthropogenic emissions). Most large cities located in high-income countries have stations responsible for measuring particulate matter and various other parameters, collectively forming an operating monitoring network, which is essential for the purposes of environmental control. In the city of Athens, which is characterized by high population density and accumulates a large number of economic activities, the currently operating monitoring network is responsible, among others, for PM₁₀ and PM_{2.5} measurements. The need for satisfactory data availability though can be supported by using machine learning methods, such as artificial neural networks. The methodology presented in this study uses a neural network model to provide spatiotemporal estimations of PM₁₀ and PM_{2.5} concentrations by utilizing the existing PM data in combination with other climatic parameters that affect them. The overall performance of the predictive neural network models' scheme is enhanced when meteorological parameters (wind speed and temperature) are included in the training process, lowering the error values of the predicted versus the observed time series' concentrations. Furthermore, this work includes the calculation of the contribution of each predictor, in order to provide a clearer understanding of the relationship between the model's output and input. The results of this procedure showcase that all PM input stations' concentrations have an important impact on the estimations. Considering the meteorological variables, the results for PM_{2.5} seem to be affected more than those for PM₁₀, although when examining PM₁₀ and PM_{2.5} individually, the wind speed and temperature contribution is on a similar level with the corresponding contribution of the available PM concentrations of the neighbouring stations.

Keywords: artificial neural networks; feed-forward networks; spatiotemporal predictions; particulate matter; climatic parameters; machine learning



Citation: Tzanis, C.G.; Alimissis, A. Contributing towards Representative PM Data Coverage by Utilizing Artificial Neural Networks. *Appl. Sci.* **2021**, *11*, 8431. <https://doi.org/10.3390/app11188431>

Academic Editor: Artur Badyda

Received: 11 July 2021

Accepted: 8 September 2021

Published: 11 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Advances in the field of air quality estimations have been rapid, particularly during the last few decades, demonstrating an increasing interest and attention in both the research community and authorities responsible for the impact assessment of air quality in modern communities. Many cities worldwide are struggling with poor air quality conditions and subsequently with increased mortality and hospital admission rates, mainly due to cardiovascular and respiratory illnesses [1,2]. This is mostly evident for cities with limited access to clean energy, resulting in an increased need for electric power generation and oil/gas extraction, both procedures responsible for emissions amplification, thus citing air pollution levels as an indicator of sustainable development goals [3]. However, all modern socioeconomic centers, where the majority of human activities transpire, need to carefully monitor and evaluate outdoor and indoor pollutants [4] which, in combination with global climate change (global warming), can lead to higher mortality rates [5]. The connection

between high air quality parameters concentrations and health-related effects has been further established in various studies [6–11]. Apart from the health aspect, high pollutant concentration values are associated with non-health-related effects in crucial fields, such as agriculture, building materials, objects of cultural heritage, forest ecosystems, etc., and are contributing to a more severe deterioration of overall air quality and consequently of the environment [12–17].

A portion of air pollution of significant importance, which, in contrast with other pollutants, is responsible not only for long-term but also short-term effects on human health, consists of particles of small diameters and various compositions [18,19]. Particulate Matter (PM), due to its long-time suspension and the ability to travel far distances in the atmosphere, is characterized as one of the leading causes of worldwide mortality and is associated with direct and indirect effects on the climate system [20,21], as well as a wide range of health problems, mostly depending on the size of the particles [22–27]. The most studied and adequately monitored categories of PM are the PM₁₀ and PM_{2.5} [28,29]. Zanobetti and Schwartz performed a study at a national level in the US, analyzing the critical effects of PM₁₀ and PM_{2.5}, and concluded that both are associated with increased rates of mortality [30]. Janssen et al. found a close relationship between these two PM size fractions and all-cause and cause-specific mortality, by using data from Statistics Netherlands during the 2008–2009 time period and for the entire Dutch population [31].

The need for spatially continuous data of air quality parameters can be satisfied by using interpolation methods at locations where there are no available observations. Akkala et al. (2010) provided a thorough review of commonly used interpolation techniques in which they presented their basic principle as well as their most important advantages and disadvantages and, additionally, they included scenarios in which these methods would be at their best potential [32]. The process of spatial interpolation eventually was supported by utilizing more advanced machine learning methodologies, such as Artificial Neural Networks (ANNs), in cases where radioactive gases concentrations data were needed [33]. Neural Network approaches have been continuously modified regarding their structure, leading to models with enhanced generalization ability and extrapolation capability and, most importantly, improved accuracy [34,35]. Gummadi et al. (2014) introduced two more regression techniques for modelling, specifically, a Random Forest Regression and a Support Vector Regression (SVR), and evaluated their usefulness and limitations in comparison with conventional methods and ANN-based approaches [36].

In the field of PM modeling, spatial and temporal estimations of PM₁₀ and PM_{2.5} concentrations can be performed by utilizing ANN statistical models, which have been proven to effectively simulate PM pollution fields in previous studies [37–40]. In general, the importance of using advanced techniques (e.g., auto-regressive models, tensor-based approaches, deep neural networks, etc.) for spatiotemporal predictive modeling, by combining data from different locations, has been demonstrated in relevant works [41–45], as these techniques consistently outperform more conventional methods. Regarding ANNs applications, the input parameters used for developing the models can be different in many cases, utilizing air quality concentrations from ground stations, satellite data and/or values from numerical climate models [39,46–49]. Adding parameters as inputs may be beneficial for estimation purposes as more information is inserted to the networks, however, the latter can become more complex and time-consuming. An essential task in this context is to carefully compare different scenarios of input parameters which will help choose the optimum input set which can be used effectively to provide accurate estimations.

This work proposes a framework that can be applied in urban environments, characterized by topographically complex terrain and high variability regarding climatic conditions, at points of interest where PM pollution measurements are needed. At these points of interest, by carrying out an experimental campaign for a short period, the results can be utilized to train ANN models. However, for the latter, meteorological predictors are also of extreme importance due to the heavy influence of climatic conditions on PM pollution distribution fields. The overall methodology which is presented examines and evaluates

how both PM concentrations and meteorological values can support PM concentrations' point estimations. Specifically, a Feed-Forward Neural Networks (FFNNs) approach was used, in order to make spatial point estimations of PM₁₀ and PM_{2.5} concentrations, aiming to develop a simple yet effective scheme which has the ability to provide representative PM datasets for stations with data gaps or to expand the available data. The spatial estimation of PM concentrations by using FFNNs and data from neighboring stations has been performed before successfully, when compared with other schemes [38]. However, this study evaluates additionally the incorporation of crucial meteorological parameters, such as the surface temperature and wind speed, and how these additions affect the performance of the networks. The methodology utilizes data from ground-based observations, obtained from monitoring stations located in the city of Athens, Greece, which is a densely populated metropolitan area, characterized by regional variability considering the type of each subsidiary area that is part of the city. Furthermore, an important part of the presented methodology is to provide an approach for understanding the contribution of each model input to the output by utilizing an approach proposed by Garson [50]. This approach can contribute to addressing the lack of explanatory power, which is a common problem associated with ANNs. and provide insight on the structure of the function being approximated, which associates input and output parameters.

2. Materials and Methods

2.1. Data

The area of study is metropolitan Athens, which is part of the Attica region in Greece. Important characteristics of the functional urban area of Athens are the considerably high population density, the complexity of its meteorological and geophysical features and the agglomeration of the majority of economic activities in Greece, which are being associated with various PM pollution sources (vehicular traffic, domestic fuel burning, natural dust and salt, industrial activities, etc.). The Athens basin is defined by four major mountain ranges. These are Mounts Parnitha, Pentelikon, Hymmetus and Aigaleo, which are natural borders at the north, northeast, east central and west respectively, and they affect the air pollutants dispersion and transportation mechanisms. Additionally, the city lies on the north coast of the Saronic Gulf and the west coast of the Euboean Gulf and thus is affected by the sea breeze and other flows. Subsequently, the complex topography and fluctuation of climatic conditions are associated with complex PM₁₀ and PM_{2.5} concentration profiles, characterized by spatial variability even for stations at close proximity [51,52]. The area of study and the locations of each monitoring station are presented in Figure 1.

The importance of the area, considering the PM pollution fields, over the last few years is additionally connected to the post-2010 time period and the economic crisis that affected Greece, during which particle concentrations in all major cities, and especially Athens, increased significantly due to the residential extensive burn of low-cost biomass as an alternative source of fuel for heating [53,54].

For this study, PM₁₀ and PM_{2.5} hourly data were obtained from the air quality monitoring network operated by the Hellenic Ministry of the Environment, Energy and Climate Change (MEE), which has operated in the Attica region since 1984. More specifically, data from nine (AGP, ARI, ELE, THR, KOR, LYK, MAR, PIR and PER) and six stations (AGP, ARI, ELE, THR, LYK and PIR) for PM₁₀ ($\mu\text{g}/\text{m}^3$) and PM_{2.5} ($\mu\text{g}/\text{m}^3$), respectively, were used in order to create the PM database, and they are presented in Table 1.



Figure 1. The Attic peninsula and the locations of the air quality monitoring stations.

Table 1. Air quality monitoring stations, their code names, and important characteristics.

Station	Code	Lon	Lat	Altitude (m-a.s.l.)	Type
Ag. Paraskevi	AGP	23°49′09″	37°59′42″	290	Suburban/Background
Aristotelous	ARI	23°43′39″	37°59′16″	95	Urban/Traffic
Elefsina	ELE	23°32′18″	38°03′04″	20	Suburban/Industrial
Thrakomakedones	THR	23°45′29″	38°08′36″	550	Suburban/Background
Koropi	KOR	23°52′44″	37°54′04″	140	Suburban/Background
Lykovrisi	LYK	23°47′19″	38°04′04″	234	Suburban/Background
Marousi	MAR	23°47′14″	38°01′51″	170	Urban/Background
Piraeus	PIR	23°38′42″	37°56′40″	4	Urban/Traffic
Peristeri	PER	23°41′18″	38°01′14″	80	Urban/Background

The selection of the stations was mostly based on data availability in each case. Additionally, daily data for two meteorological parameters (wind speed in km/h and temperature in °C) were obtained for the target station (AGP) from the automatic weather stations NOAA (National Observatory of Athens Automated Network) network of the National Observatory of Athens (NOA). The methodology used in this study could be applied for a different target station in the area. However, the AGP station was selected among the six common stations of PM₁₀ and PM_{2.5} due to the temperature (T) and wind speed (WS) values' availability, which can help in better supporting the methodology. Ultimately, the analysis covers a three-year time period (2016–2018) for both the PM and the meteorological parameters. Figure 2 depicts the average monthly evolution for both PM₁₀ and PM_{2.5} and for the 2016 to 2018 time period at the AGP station. All three years'

monthly averaged concentrations for each pollutant are presented in the same diagram for comparative purposes.

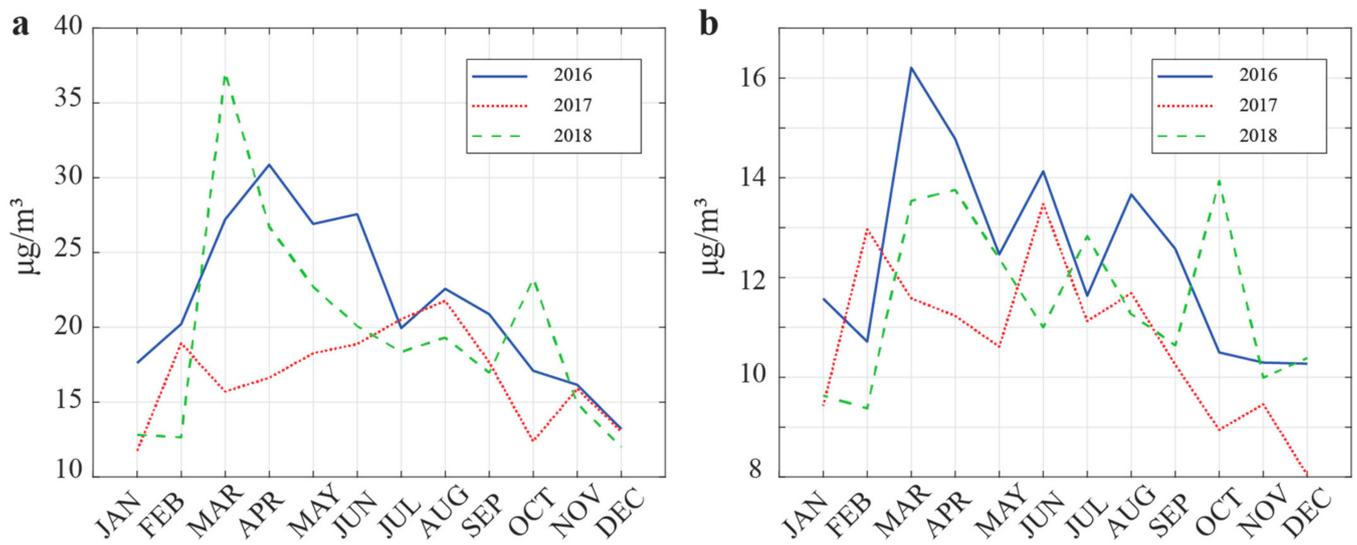


Figure 2. Monthly averaged concentrations for PM₁₀ and PM_{2.5} ((a,b) respectively) and for the three-year time period (2016–2018) at AGP target station.

2.2. Methodology

Initially, as mentioned above, the AGP station was selected as a target station for which all the steps of the methodology were performed. This station had high percentages of data availability (>90%) for both pollutants and for all three years (2016–2018), which was important for the evaluation of the results. Accordingly, yearly averaged, maximum and minimum concentrations for PM₁₀ and PM_{2.5} were calculated in AGP. These descriptive statistics are helpful during the discussion of the results and act as an initial description of the 2016–2018 PM conditions for this specific location. The selected machine learning scheme which was used in this study is an FFNN model designed for spatial point interpolation. According to Hornik et al., this type of architecture can effectively simulate the relationship between input and output to various degrees of accuracy, based on several parameters that are part of the networks structure (Figure 3) [55].

The FFNN is a multilayer perceptron and the information flow follows one direction, advancing from the input to the output without looping [56]. The equation through which the output of a neuron can be calculated is the following:

$$y = f\left(\sum_{i=1}^M x_i w_i + b\right) \quad (1)$$

where f is the activation function, x_i the inputs, w_i the synaptic weights and b the bias. The synaptic weights are the internal connections among the neurons of the network (Figure 3), and through adjustments of their values, the strength of the connections is modified [57]. The PM₁₀ and PM_{2.5} concentrations were estimated by using AGP as a target station and the remaining stations' concentrations as inputs. The number of input stations is different for each pollutant (eight and five for PM₁₀ and PM_{2.5}, respectively). Three stations for PM₁₀ are not available for PM_{2.5} due to limited data, and they were excluded. Additionally, the daily temperature and wind speed values at AGP were used as predictors in the model. Four different models were developed in order to compare their performance. For the first model, the predictors were only the data of the input stations. For the second, third and fourth model, the number of predictors/inputs increased by adding the temperature values, the wind speed values and both wind speed and temperature values, respectively. In all

four models, the output was the AGP PM concentrations. Eventually, eight models were created in total (four for PM₁₀ and four for PM_{2.5}). The aim of this additive process was to investigate how much the meteorological parameters affect the accuracy of the estimations. Initially, the datasets (PM₁₀, PM_{2.5}, Temperature and Wind Speed) were randomly divided into the training (70%), validation (15%) and test (15%) subsets. While the pollutants and meteorological data points for these datasets were randomly selected from the 2016–2018 time period, they were common for all inputs of each individual network development. When the network used a data point for a random hour from a monitoring station, the same hour was selected for the remaining stations. This procedure was followed so as to retain the daily variability and avoid mixing seasons and even days due to the short-term fluctuations of the PM concentrations.

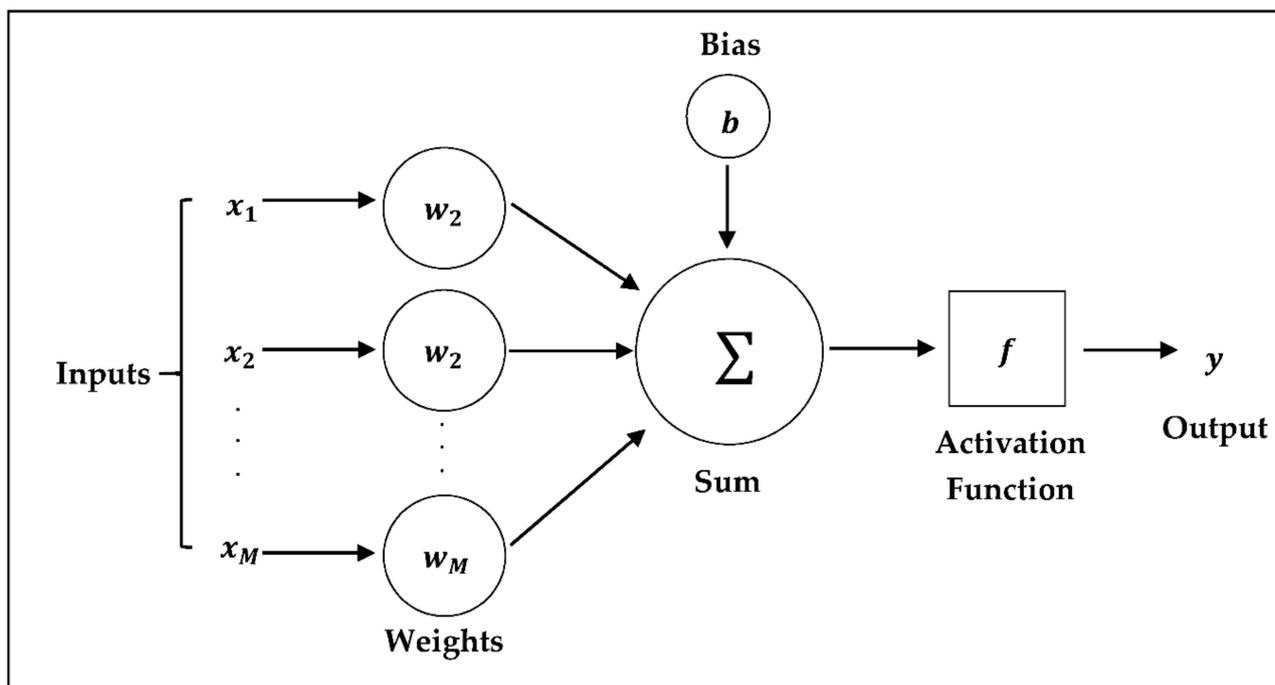


Figure 3. An example of an artificial neuron and its basic characteristics.

The next step involved the selection of the optimum number of neurons in the hidden layer. The FFNN consists of three layers, the input, hidden and output layers [58]. The number of neurons in the input and output layers is completely determined by the inputs and outputs. The hidden layer size is an important part of the network architecture. In order to select the optimum architecture, the criterion that was followed was the minimization of the Mean Absolute Error (MAE) on the validation subset [59]. Lower MAE values correspond to a better performing network in relation to a lower degree of complexity. Different FFNN configurations were tested for multiple runs (10 repetitions), due to the initial weights of the neurons of the model being randomly established, and thus, the average result of these runs guaranteed the randomness of the process. The number of hidden neurons tested in all cases ranged from one to forty. To avoid pattern exploitation in the training subset (overfitting), the early stopping approach was used [60], which, according to the validation subset error (when it started to increase), stopped the training process. The final networks that were developed were evaluated for their estimation accuracy by applying two difference and correlation statistical measures, the MAE and the

coefficient of determination (R^2), [38,59,61,62] on the test subset of the output vector. These criteria are calculated by using the following equations:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |E_i - A_i| \quad (2)$$

$$R^2 = \left(\frac{\sum_{i=1}^n (A_i - \bar{A})(E_i - \bar{E})}{\sqrt{\sum_{i=1}^n (A_i - \bar{A})^2} \sqrt{\sum_{i=1}^n (E_i - \bar{E})^2}} \right)^2 \quad (3)$$

where n is the number of data points, E the estimated and A the observed concentrations. The best-performing models are associated with lower MAE and higher R^2 values and are evaluated based on the results of both statistical parameters. For the MAE metric, the standard deviation (SD) was also calculated to indicate the dispersion of the estimated concentrations from the MAE value. Additionally, the FFNN models' results were compared with the corresponding estimations of a multiple linear regression model (MLR) [38,63] in order to further establish the superior predictive ability of the FFNNs. Finally, the accuracy of the FFNN models is also examined by plotting scatter diagrams which additionally contribute towards an easier comparison among the models. The scatter diagrams provide information considering the relationship between the observed and estimated values at high, medium and low concentration levels.

The last part of the methodology includes an analysis regarding the distinguishing of the significance of every input variable to the output, for all the FFNN models that were developed, utilizing an algorithm proposed by Garson [50]. This methodology is based on recognizing the associations that the synaptic weights reveal considering the inputs and output relationship and was also used in other studies in the field of air quality, to quantify the importance of each station's data (inputs) to the estimated values for the target station [38,63,64]. The Relative Importance (RI) percentage is calculated with the use of Equation (4),

$$RI_{IK}(\%) = \frac{\sum_{j=1}^h \left| \frac{w_{ji}w_{kj}}{\sum_{i=1}^n |w_{ji}|} \right|}{\sum_{i=1}^n \sum_{j=1}^h \left| \frac{|w_{kj}| |w_{ji}|}{\sum_{i=1}^n |w_{ji}|} \right|} \quad (4)$$

where w_{ij} , w_{kj} are the connection weights between the i -th input and j -th hidden neuron, and between the j -th hidden and k -th output neuron, respectively. In general, ANNs provide little explanatory insight into the individual contribution of the input variables in the estimation procedure. The RI method addresses this issue and can be used as a variable selection technique for similar problems.

3. Results and Discussion

As aforementioned, the results presented in this section are for the AGP station. Descriptive statistics for the 2016–2018 period in the AGP monitoring station are presented in Table 2. This table includes yearly mean, max and min concentrations for each year individually and the corresponding values for the three years in total. Both PM_{10} and $\text{PM}_{2.5}$ are measured in $\mu\text{g}/\text{m}^3$, and the monitoring methodology is based on beta radiation absorption.

Table 3 includes the number of data points that were used for each subset during the development of the models (input data). There are more available data points for PM_{10} due to the increased number of monitoring stations that were used as inputs. In both pollutant cases, when the meteorological parameters are added, they qualify as an additional predictor that has the same number of data points with the input stations' concentrations. Thus, the scenario with the most inputs, i.e., where both WS and T are incorporated, has a higher number of data points available for the training, validation and test subsets. In all cases, the architecture of the ANNs, following the experimental

design of this work, defines the number of data points included in the input and output vectors. The size of the training-validation-test subset for the output vector is based on the 70-15-15 percentages which were introduced in the previous Section, and the resulting data points are 10,339-2215-2215 for PM₁₀ and 11,090-2376-2376 for PM_{2.5}. The data points of the output vector are the same for all four scenarios, as the output is always the PM concentrations at AGP.

The architecture of the models is presented in Table 4. The number of inputs is the total number of predictor stations, and one (T, WS) or two meteorological parameters inputs (T and WS) are added according to the model used in the second, third and fourth row of the table. It is evident that the number of hidden neurons in the models for both PMs is lower when the meteorological data are not included in the inputs (16 and 13 hidden neurons for PM₁₀ and PM_{2.5}, respectively). The same number ranges from 26 to 30 for the remaining six schemes. This difference can be associated with the increased complexity of these networks. As more inputs with different characteristics are added to the network, the latter needs additional hidden neurons to simulate the relationship between input and target data.

Table 2. Yearly Mean, Max and Min values of PM₁₀ and PM_{2.5} concentrations in the AGP monitoring station for the period 2016–2018.

Year	PM ₁₀			PM _{2.5}		
	Mean	Max	Min	Mean	Max	Min
2016	21.70	714	1	12.30	208	0
2017	16.83	136	1	10.73	67	0
2018	19.85	530	1	11.60	118	0
Total	19.43	714	1	11.54	208	0

Table 3. Number of data points in each subset for PM₁₀ and PM_{2.5} and the four different input cases.

	Training	PM ₁₀ Validation	Test	Training	PM _{2.5} Validation	Test
PM	82,712	17,720	17,720	55,450	11,880	11,880
PM + T	93,051	19,935	19,935	66,540	14,256	14,256
PM + WS	93,051	19,935	19,935	66,540	14,256	14,256
PM + T + WS	103,390	22,150	22,150	77,630	16,632	16,632

Table 4. Number of neurons in each layer of the models for PM₁₀ and PM_{2.5} and the four different input cases.

	Input	PM ₁₀ Hidden	Output	Input	PM _{2.5} Hidden	Output
PM	8	16	1	5	13	1
PM + T	9	30	1	6	30	1
PM + WS	9	27	1	6	28	1
PM + T + WS	10	26	1	7	28	1

Tables 5 and 6 show the MAE, SD of MAE and R^2 values for each of the eight models and the corresponding cases for the MLR method. These values are the result of applying the abovementioned metrics on the test subset of the output vector. When comparing the two methodologies, it is evident that the FFNN models outperform the MLR scheme for both PM₁₀ and PM_{2.5} and all input scenarios. They display lower MAE and SD of MAE and higher R^2 values, indicating that the FFNNs methodology simulates more effectively the nonlinear relationship between the input and output parameters. Considering individually

the results of the FFNNs method, Table 5 provides some interesting findings. In general, the PM₁₀ models are associated with low error and high correlation values, providing satisfactory results regardless of the input data that were used. The PM_{2.5} models' results showcase higher MAE error values (considering that the MAE error is higher when compared with the average PM_{2.5} values) and lower R^2 values (which can possibly be attributed to the lower number of input data, and subsequently, less information during the training process). However, on average, in both cases, the schemes that include T and WS values give lower MAE and higher R^2 values. This is evident especially for the two models that include both T and WS, where the lowest MAE (3.67 $\mu\text{g}/\text{m}^3$ and 2.39 $\mu\text{g}/\text{m}^3$) and highest R^2 (0.94 and 0.75) values are produced. Although the MAE value average for the models with PM₁₀ inputs (3.85 $\mu\text{g}/\text{m}^3$) is higher than the corresponding value for PM_{2.5} (2.44 $\mu\text{g}/\text{m}^3$), the MAE statistical metric uses the same scale as the data being measured and is not suitable for comparison between PM₁₀ and PM_{2.5} in contrast to R^2 , which illustrates better results for the PM₁₀ cases. A conclusion of significant importance can be drawn by comparing the MAE values with the yearly mean, maximum and minimum concentrations, which are presented in Table 2. While the FFNNs that include as predictors the surface temperature and wind speed correspond to better performance statistics (lower MAE and higher R^2), the differences between the models are small regarding the Table 2 values. This fact illustrates the effectiveness, in this case, of the models that are using only concentrations from neighbouring stations. However, adding more parameters or changing the networks configuration (i.e., selecting the subsets data by chronological order and not randomly, using different approaches to avoid overfitting etc.) can further improve the results.

Table 5. MAE, SD of MAE and R^2 values that correspond to the FFNN methodology, for PM₁₀ and PM_{2.5} and the four different input cases.

	PM ₁₀			PM _{2.5}		
	MAE ($\mu\text{g}/\text{m}^3$)	SD of MAE ($\mu\text{g}/\text{m}^3$)	R^2	MAE ($\mu\text{g}/\text{m}^3$)	SD of MAE ($\mu\text{g}/\text{m}^3$)	R^2
PM	3.99	4.85	0.83	2.45	2.39	0.70
PM + T	3.89	4.15	0.91	2.44	2.38	0.70
PM + WS	3.85	4.40	0.90	2.47	2.12	0.65
PM + T + WS	3.67	3.86	0.94	2.39	2.16	0.75

Table 6. MAE, SD of MAE and R^2 values that correspond to the MLR methodology, for PM₁₀ and PM_{2.5} and the four different input cases.

	PM ₁₀			PM _{2.5}		
	MAE ($\mu\text{g}/\text{m}^3$)	SD of MAE ($\mu\text{g}/\text{m}^3$)	R^2	MAE ($\mu\text{g}/\text{m}^3$)	SD of MAE ($\mu\text{g}/\text{m}^3$)	R^2
PM	4.96	5.4	0.83	2.99	3.18	0.55
PM + T	4.97	5.39	0.83	2.97	3.15	0.56
PM + WS	4.93	5.35	0.83	2.99	3.18	0.55
PM + T + WS	4.93	5.34	0.84	2.97	3.15	0.56

An additional evaluation of the FFNN models is performed by plotting scatter diagrams of the predicted versus the observed values, as presented in Figure 4.

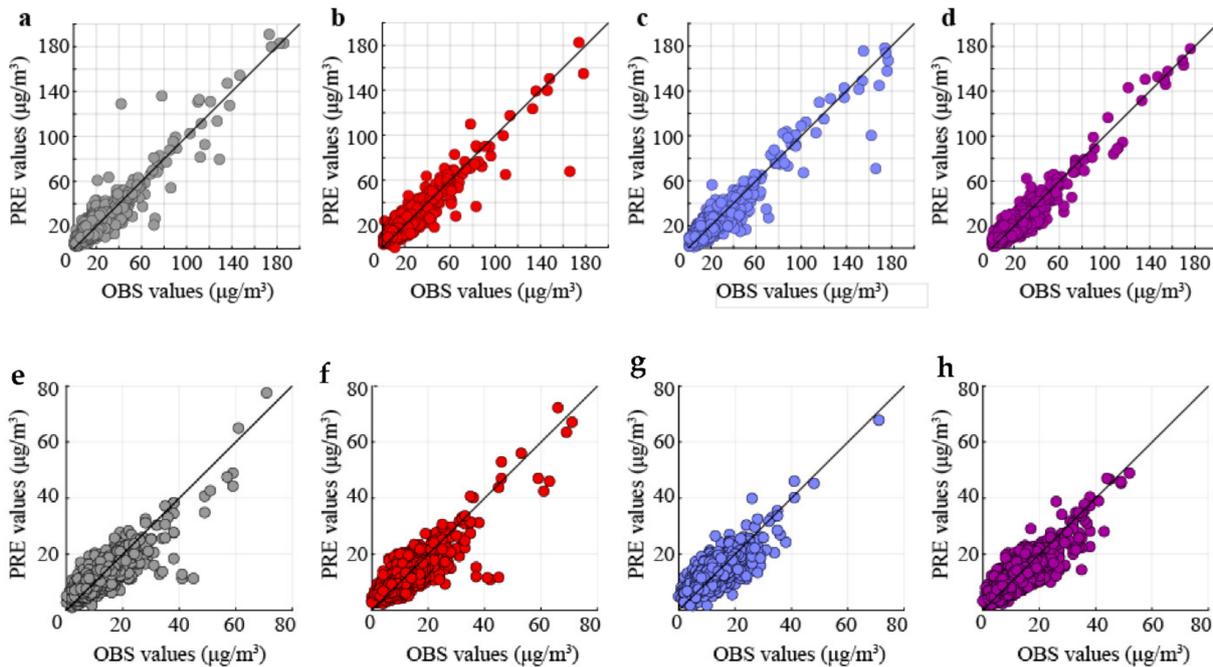


Figure 4. PM₁₀ and PM_{2.5} scatter diagrams of the predicted versus the observed concentrations for the different input cases: (a,e) PM only, (b,f) PM and T, (c,g) PM and WS and (d,h) PM, T and WS.

The scatter diagrams for PM₁₀ and PM_{2.5} (Figure 4) are consistent with the MAE and R² performance statistics (Table 5). The degree of dispersion for the PM₁₀ (Figure 4a–d) is lower compared to PM_{2.5} (Figure 4e–h). This can be explained by the lower number of inputs provided in order to train the models. Specifically, during the training process, the number of input stations is eight for PM₁₀ and five for PM_{2.5}, meaning that the air quality network density for the latter was lower. On the contrary, there are no notable differences when the diagrams are compared based on the different inputs. According to the MAE values, the performance of all the models, when studied separately for each figure, reveals that there is not a scheme that identifies as substantially superior. However, a closer examination reveals that the models which include both meteorological parameters (T and WS) produce scatter diagrams with lower dispersion across the line of optimum agreement. This is especially evident regarding the higher concentration values (upper right) for the PM₁₀ models, where the markers are closer to the diagonal.

Finally, the results of the Garson methodology are presented in Table 7. The percentage of contribution for the meteorological parameters is nearly half when compared with the monitoring stations concentrations in the case of the PM₁₀ models. For PM_{2.5}, the corresponding percentages are at a similar level (~15%). Additionally, the monitoring stations which are of the same type (Suburban/Background) as AGP (KOR, LYK and THR), and those which are at proximity (MAR, LYK and ARI), are expected to contribute more to the AGP concentrations estimations. However, Table 7 reveals that all stations have a significant importance for the models.

Table 7. Relative importance (%) of the input data for the two cases of PM₁₀ and PM_{2.5} FFNN models, where both PM concentrations and meteorological parameters (T, WS) are used.

	ARI	ELE	THR	KOR	LYK	MAR	PIR	PER	T	WS
PM ₁₀	8.38	12.06	13.85	10.02	8.93	13.61	11.07	10.45	5.07	6.56
PM _{2.5}	14.14	11.92	14.70		13.54		15.42		16.69	13.59

4. Conclusions

This study used an FFNN application for estimating PM₁₀ and PM_{2.5} concentrations. ANN approaches, in general, have the advantage to be able to model effectively nonlinear relationships compared to other methodologies. An important aspect is the evaluation of the developed models during different scenarios of input parameters. In nearly all cases, the MAE and R^2 values were lower and higher, respectively, when the meteorological values were added during the training process. The models that showcased a better performance were those who had as additional inputs both T and WS, although there were not crucial differences noticed among the schemes of the four different scenarios. Regarding the comparison between PM₁₀ and PM_{2.5}, the estimations for the latter had a higher degree of dispersion in the scatter diagrams of the observed versus the estimated values. This can be explained due to the more limited information provided during training (more input stations for PM₁₀). The Garson methodology results reveal that all monitoring stations in the Attica region, which were involved in the FFNN development process, are important for the PM estimations. Future work can extend this methodology to include more target stations with different characteristics and/or add more climate parameters. These additions, considering their impact and usefulness for the models, can be further analyzed and supported by applying suitable feature selection and feature ranking techniques [65–67]. Finally, ANN ensemble approaches [68] can be examined, aiming to reduce the variance of predictions and the generalization error by combining the results of multiple models.

Author Contributions: C.G.T. and A.A. were involved in the investigation, conceptualization, writing—original draft preparation and writing—review and editing of this work, while, individually, C.G.T. was responsible for the data curation, validation of the results and supervised the whole procedure. Both C.G.T. and A.A. performed the various steps of the methodology, processed the data and developed the neural network models. Both authors were involved in the discussion of the results and commented on the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The air quality and meteorological datasets generated and/or analyzed during the current study are publicly available in the Ministry of Environment and Energy repository, (ypen.gov.gr) (accessed on 30 June 2020) and the National Observatory of Athens repository (<https://meteosearch.meteo.gr/>) (accessed on 30 June 2020) respectively.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hoek, G.; Krishnan, R.M.; Beelen, R.; Peters, A.; Ostro, B.; Brunekreef, B.; Kaufman, J.D. Long-term air pollution exposure and cardio-respiratory mortality: A review. *Environ. Health* **2013**, *12*, 43. [[CrossRef](#)]
2. Cairncross, E.K.; John, J.; Zunckel, M. A novel air pollution index based on the relative risk of daily mortality associated with short-term exposure to common air pollutants. *Atmos. Environ.* **2007**, *41*, 8442–8454. [[CrossRef](#)]
3. WHO. *Guidelines for Air Quality*; World Health Organization: Geneva, Switzerland, 2000.
4. Gurjar, B.R.; Jain, A.; Sharma, A.; Agarwal, A.; Gupta, P.; Nagpure, A.; Lelieveld, J. Human health risks in megacities due to air pollution. *Atmos. Environ.* **2010**, *44*, 4606–4613. [[CrossRef](#)]
5. Willers, S.M.; Jonker, M.F.; Klok, L.; Keuken, M.P.; Odink, J.; Elshout, S.V.D.; Sabel, C.E.; Mackenbach, J.P.; Burdorf, A. High resolution exposure modelling of heat and air pollution and the impact on mortality. *Environ. Int.* **2016**, *89–90*, 102–109. [[CrossRef](#)] [[PubMed](#)]
6. Orru, H.; Ebi, K.L.; Forsberg, B. The Interplay of Climate Change and Air Pollution on Health. *Curr. Environ. Health Rep.* **2017**, *4*, 504–513. [[CrossRef](#)] [[PubMed](#)]
7. Héroux, M.-E.; Anderson, H.R.; Atkinson, R.; Brunekreef, B.; Cohen, A.; Forastiere, F.; Hurley, F.; Katsouyanni, K.; Krewski, D.; Krzyzanowski, M.; et al. Quantifying the health impacts of ambient air pollutants: Recommendations of a WHO/Europe project. *Int. J. Public Health* **2015**, *60*, 619–627. [[CrossRef](#)]
8. Varotsos, C.; Christodoulakis, J.; Tzani, C.; Cracknell, A.P. Signature of tropospheric ozone and nitrogen dioxide from space: A case study for Athens, Greece. *Atmos. Environ.* **2014**, *89*, 721–730. [[CrossRef](#)]

9. Fang, Y.; Naik, V.; Horowitz, L.W.; Mauzerall, D.L. Air pollution and associated human mortality: The role of air pollutant emissions, climate change and methane concentration increases from the preindustrial period to present. *Atmos. Chem. Phys. Discuss.* **2013**, *13*, 1377–1394. [[CrossRef](#)]
10. Pascal, M.; Corso, M.; Chanel, O.; Declercq, C.; Badaloni, C.; Cesaroni, G.; Henschel, S.; Meister, K.; Haluza, D.; Martin-Olmedo, P.; et al. Assessing the public health impacts of urban air pollution in 25 European cities: Results of the Aphekom project. *Sci. Total Environ.* **2013**, *449*, 390–400. [[CrossRef](#)] [[PubMed](#)]
11. Curtis, L.; Rea, W.; Smith-Willis, P.; Fenyves, E.; Pan, Y. Adverse health effects of outdoor air pollutants. *Environ. Int.* **2016**, *32*, 815–830. [[CrossRef](#)]
12. Grøntoft, T. Estimation of Damage Cost to Building Façades per kilo Emission of Air Pollution in Norway. *Atmosphere* **2020**, *11*, 686. [[CrossRef](#)]
13. Locosselli, G.M.; de Camargo, E.P.; Moreira, T.; Todesco, E.; Andrade, M.D.F.; de André, C.D.S.; de André, P.A.; Singer, J.; Ferreira, L.S.; Saldiva, P.; et al. The role of air pollution and climate on the growth of urban trees. *Sci. Total Environ.* **2019**, *666*, 652–661. [[CrossRef](#)]
14. Di Turo, F.; Proietti, C.; Screpanti, A.; Fornasier, M.F.; Cionni, I.; Favero, G.; De Marco, A. Impacts of air pollution on cultural heritage corrosion at European level: What has been achieved and what are the future scenarios. *Environ. Pollut.* **2016**, *218*, 586–594. [[CrossRef](#)] [[PubMed](#)]
15. Barca, D.; Comite, V.; Belfiore, C.M.; Bonazza, A.; La Russa, M.F.; Ruffolo, S.A.; Crisci, G.M.; Pezzino, A.; Sabbioni, C. Impact of air pollution in deterioration of carbonate building materials in Italian urban environments. *Appl. Geochem.* **2014**, *48*, 122–131. [[CrossRef](#)]
16. De La Fuente, D.; Vega, J.M.; Viejo, F.; Díaz, I.; Morcillo, M. City scale assessment model for air pollution effects on the cultural heritage. *Atmos. Environ.* **2011**, *45*, 1242–1250. [[CrossRef](#)]
17. Paoletti, E.; Bytnerowicz, A.; Andersen, C.; Augustaitis, A.; Ferretti, M.; Grulke, N.; Günthardt-Goerg, M.S.; Innes, J.; Johnson, D.; Karnosky, D.; et al. Impacts of Air Pollution and Climate Change on Forest Ecosystems—Emerging Research Needs. *Sci. World J.* **2007**, *7* (Suppl. 1), 1–8. [[CrossRef](#)]
18. Amanollahi, J.; Tzani, C.; Abdullah, A.M.; Ramli, M.F.; Pirasteh, S. Development of the models to estimate particulate matter from thermal infrared band of Landsat Enhanced Thematic Mapper. *Int. J. Environ. Sci. Technol.* **2013**, *10*, 1245–1254. [[CrossRef](#)]
19. Anderson, J.O.; Thundiyil, J.G.; Stolbach, A. Clearing the Air: A Review of the Effects of Particulate Matter Air Pollution on Human Health. *J. Med. Toxicol.* **2012**, *8*, 166–175. [[CrossRef](#)]
20. Varotsos, C.A.; Melnikova, I.N.; Cracknell, A.P.; Tzani, C.; Vasilyev, A.V. New spectral functions of the near-ground albedo derived from aircraft diffraction spectrometer observations. *Atmos. Chem. Phys.* **2014**, *14*, 6953–6965. [[CrossRef](#)]
21. Tzani, C.; Varotsos, C.A. Tropospheric aerosol forcing of climate: A case study for the greater area of Greece. *Int. J. Remote Sens.* **2008**, *29*, 2507–2517. [[CrossRef](#)]
22. Hamanaka, R.B.; Mutlu, G.M. Particulate Matter Air Pollution: Effects on the Cardiovascular System. *Front. Endocrinol.* **2018**, *9*, 680. [[CrossRef](#)] [[PubMed](#)]
23. Raaschou-Nielsen, O.; Beelen, R.; Wang, M.; Hoek, G.; Andersen, Z.J.; Hoffmann, B.; Stafoggia, M.; Samoli, E.; Weinmayr, G.; Dimakopoulou, K.; et al. Particulate matter air pollution components and risk for lung cancer. *Environ. Int.* **2016**, *87*, 66–73. [[CrossRef](#)]
24. Kim, K.-H.; Kabir, E.; Kabir, S. A review on the human health impact of airborne particulate matter. *Environ. Int.* **2015**, *74*, 136–143. [[CrossRef](#)]
25. Chen, R.; Kan, H.; Chen, B.; Huang, W.; Bai, Z.; Song, G.; Pan, G. Association of Particulate Air Pollution with Daily Mortality: The China Air Pollution and Health Effects Study. *Am. J. Epidemiol.* **2012**, *175*, 1173–1181. [[CrossRef](#)] [[PubMed](#)]
26. Laumbach, R.J.; Kipen, H.M. Respiratory health effects of air pollution: Update on biomass smoke and traffic pollution. *J. Allergy Clin. Immunol.* **2012**, *129*, 3–11. [[CrossRef](#)] [[PubMed](#)]
27. Dockery, D.W. Health Effects of Particulate Air Pollution. *Ann. Epidemiol.* **2009**, *19*, 257–263. [[CrossRef](#)] [[PubMed](#)]
28. Milinevsky, G.; Miatselskaya, N.; Grytsai, A.; Danylevsky, V.; Bril, A.; Chaikovskiy, A.; Yukhymchuk, Y.; Wang, Y.; Liptuga, A.; Kyslyi, V.; et al. Atmospheric Aerosol Distribution in 2016–2017 over the Eastern European Region Based on the GEOS-Chem Model. *Atmosphere* **2020**, *11*, 722. [[CrossRef](#)]
29. Suh, H.H. Particulate matter. *Expo. Assess. Occup. Environ. Epidemiol.* **2003**, *1*, 221–236. [[CrossRef](#)]
30. Zanobetti, A.; Schwartz, J. The Effect of Fine and Coarse Particulate Air Pollution on Mortality: A National Analysis. *Environ. Health Perspect.* **2009**, *117*, 898–903. [[CrossRef](#)]
31. Janssen, N.; Fischer, P.; Marra, M.; Ameling, C.; Cassee, F. Short-term effects of PM_{2.5}, PM₁₀ and PM_{2.5–10} on daily mortality in the Netherlands. *Sci. Total Environ.* **2013**, *463–464*, 20–26. [[CrossRef](#)]
32. Akkala, A.; Devabhaktuni, V.; Kumar, A. Interpolation techniques and associated software for environmental data. *Environ. Prog. Sustain. Energy* **2010**, *29*, 134–141. [[CrossRef](#)]
33. Akkala, A.; Devabhaktuni, V.; Kumar, A.; Bhatt, D. Development of an ANN interpolation scheme for estimating missing radon concentrations in Ohio. *Open Environ. Biol. Monit. J.* **2011**, *4*, 21–30. [[CrossRef](#)]
34. Akkala, A.; Bhatt, D.; Devabhaktuni, V.; Kumar, A. Knowledge-based neural network approaches for modeling and estimating radon concentrations. *Environ. Prog. Sustain. Energy* **2012**, *32*, 355–364. [[CrossRef](#)]

35. Yerrabolu, P.; Mareddy, L.; Bhatt, D.; Aggarwal, P.; Kumar, A.; Devabhaktuni, V. Correction Model-Based ANN Modeling Approach for the Estimation of Radon Concentrations in Ohio. *Environ. Prog. Sustain. Energy* **2012**, *32*, 1223–1233. [[CrossRef](#)]
36. Gummadi, J.; Bhatt, D.; Adusumilli, S.; Devabhaktuni, V.; Acosta, W.; Kumar, A. Interpolation techniques for modeling and estimating indoor radon concentrations in Ohio: Comparative study. *Environ. Prog. Sustain. Energy* **2014**, *34*, 169–177. [[CrossRef](#)]
37. Mirzaei, M.; Amanollahi, J.; Tzani, C.G. Evaluation of linear, nonlinear, and hybrid models for predicting PM_{2.5} based on a GTWR model and MODIS AOD data. *Air Qual. Atmos. Health* **2019**, *12*, 1215–1224. [[CrossRef](#)]
38. Tzani, C.G.; Alimissis, A.; Philippopoulos, K.; Deligiorgi, D. Applying linear and nonlinear models for the estimation of particulate matter variability. *Environ. Pollut.* **2019**, *246*, 89–98. [[CrossRef](#)] [[PubMed](#)]
39. Fernando, H.; Mammarella, M.; Grandoni, G.; Fedele, P.; Di Marco, R.; Dimitrova, R.; Hyde, P. Forecasting PM₁₀ in metropolitan areas: Efficacy of neural networks. *Environ. Pollut.* **2012**, *163*, 62–67. [[CrossRef](#)] [[PubMed](#)]
40. Hooyberghs, J.; Mensink, C.; Dumont, G.; Fierens, F.; Brasseur, O. A neural network forecast for daily average PM₁₀ concentrations in Belgium. *Atmos. Environ.* **2005**, *39*, 3279–3289. [[CrossRef](#)]
41. Corizzo, R.; Ceci, M.; Fanaee, H.; Gama, J. Multi-aspect renewable energy forecasting. *Inf. Sci.* **2021**, *546*, 701–722. [[CrossRef](#)]
42. Bessa, R.; Trindade, A.; Silva, C.S.; Miranda, V. Probabilistic solar power forecasting in smart grids using distributed information. *Int. J. Electr. Power Energy Syst.* **2015**, *72*, 16–23. [[CrossRef](#)]
43. Akay, B.; Ragni, D.; Ferreira, C.S.; Van Bussel, G. Experimental investigation of the root flow in a horizontal axis wind turbine. *Wind. Energy* **2014**, *17*, 1093–1109. [[CrossRef](#)]
44. Kalajdjieski, J.; Zdravevski, E.; Corizzo, R.; Lameski, P.; Kalajdziski, S.; Pires, I.; Garcia, N.; Trajkovic, V. Air Pollution Prediction with Multi-Modal Data and Deep Neural Networks. *Remote Sens.* **2020**, *12*, 4142. [[CrossRef](#)]
45. Arsov, M.; Zdravevski, E.; Lameski, P.; Corizzo, R.; Koteli, N.; Gramatikov, S.; Mitreski, K.; Trajkovic, V. Multi-Horizon Air Pollution Forecasting with Deep Neural Networks. *Sensors* **2021**, *21*, 1235. [[CrossRef](#)] [[PubMed](#)]
46. Li, T.; Shen, H.; Zeng, C.; Yuan, Q.; Zhang, L. Point-surface fusion of station measurements and satellite observations for mapping PM_{2.5} distribution in China: Methods and assessment. *Atmos. Environ.* **2017**, *152*, 477–489. [[CrossRef](#)]
47. Chellali, M.R.; Abderrahim, H.; Hamou, A.; Nebatti, A.; Janovec, J. Artificial neural network models for prediction of daily fine particulate matter concentrations in Algiers. *Environ. Sci. Pollut. Res.* **2016**, *23*, 14008–14017. [[CrossRef](#)]
48. Gupta, P.; Christopher, S.A. Particulate matter air quality assessment using integrated surface, satellite, and meteorological products: A neural network approach. *J. Geophys. Res. Space Phys.* **2009**, *114*, 1–14. [[CrossRef](#)]
49. Liu, Y.; Sarnat, J.A.; Kilaru, V.; Jacob, D.J.; Koutrakis, P. Estimating Ground-Level PM_{2.5} in the Eastern United States Using Satellite Remote Sensing. *Environ. Sci. Technol.* **2005**, *39*, 3269–3278. [[CrossRef](#)]
50. Garson, G.D. Interpreting neural-network connection weights. *AI Expert* **1991**, *6*, 47–51.
51. Mavrakou, T.; Philippopoulos, K.; Deligiorgi, D. The impact of sea breeze under different synoptic patterns on air pollution within Athens basin. *Sci. Total Environ.* **2012**, *433*, 31–43. [[CrossRef](#)]
52. Tzani, C.G.; Koutsogiannis, I.; Philippopoulos, K.; Deligiorgi, D. Recent climate trends over Greece. *Atmos. Res.* **2019**, *230*, 104623. [[CrossRef](#)]
53. Taghvaei, S.; Sowlat, M.H.; Mousavi, A.; Hassanvand, M.S.; Yunesian, M.; Naddafi, K.; Sioutas, C. Source apportionment of ambient PM_{2.5} in two locations in central Tehran using the Positive Matrix Factorization (PMF) model. *Sci. Total Environ.* **2018**, *628–629*, 672–686. [[CrossRef](#)]
54. Argyropoulos, G.; Besis, A.; Voutsas, D.; Samara, C.; Sowlat, M.H.; Hasheminassab, S.; Sioutas, C. Source apportionment of the redox activity of urban quasi-ultrafine particles (PM_{0.49}) in Thessaloniki following the increased biomass burning due to the economic crisis in Greece. *Sci. Total Environ.* **2016**, *568*, 124–136. [[CrossRef](#)]
55. Hornik, K.; Stinchcombe, M.; White, H. Multilayer feedforward networks are universal approximators. *Neural Netw.* **1989**, *2*, 359–366. [[CrossRef](#)]
56. Kulluk, S.; Ozbakir, L.; Baykasoglu, A. Training neural networks with harmony search algorithms for classification problems. *Eng. Appl. Artif. Intell.* **2012**, *25*, 11–19. [[CrossRef](#)]
57. Haykin, S. *Neural Networks and Learning Machines*, 3rd ed.; Pearson Education Inc.: Hoboken, NJ, USA, 2009.
58. Chattopadhyay, S. Feed forward Artificial Neural Network model to predict the average summer-monsoon rainfall in India. *Acta Geophys.* **2007**, *55*, 369–382. [[CrossRef](#)]
59. Tzani, C.; Alimissis, A.; Koutsogiannis, I. Addressing Missing Environmental Data via a Machine Learning Scheme. *Atmosphere* **2021**, *12*, 499. [[CrossRef](#)]
60. Piotrowski, A.P.; Napiorkowski, J. A comparison of methods to avoid overfitting in neural networks training in the case of catchment runoff modelling. *J. Hydrol.* **2013**, *476*, 97–111. [[CrossRef](#)]
61. Fallahi, S.; Amanollahi, J.; Tzani, C.G.; Ramli, M.F. Estimating solar radiation using NOAA/AVHRR and ground measurement data. *Atmos. Res.* **2018**, *199*, 93–102. [[CrossRef](#)]
62. Rahimpour, A.; Amanollahi, J.; Tzani, C.G. Air quality data series estimation based on machine learning approaches for urban environments. *Air Qual. Atmos. Health* **2021**, *14*, 191–201. [[CrossRef](#)]
63. Alimissis, A.; Philippopoulos, K.; Tzani, C.G.; Deligiorgi, D. Spatial estimation of urban air pollution with the use of artificial neural network models. *Atmos. Environ.* **2018**, *191*, 205–213. [[CrossRef](#)]
64. Adams, M.D.; Kanaroglou, P.S. Mapping real-time air pollution health risk for environmental management: Combining mobile and stationary air pollution monitoring with neural network models. *J. Environ. Manag.* **2016**, *168*, 133–141. [[CrossRef](#)]

65. González-Enrique, J.; Turias, I.J.; Ruiz-Aguilar, J.J.; Moscoso-López, J.A.; Franco, L. Spatial and meteorological relevance in NO₂ estimations: A case study in the Bay of Algeciras (Spain). *Stoch. Environ. Res. Risk Assess.* **2019**, *33*, 801–815. [[CrossRef](#)]
66. González-Enrique, J.; Ruiz-Aguilar, J.J.; Moscoso-López, J.A.; Urda, D.; Turias, I.J. A comparison of ranking filter methods applied to the estimation of NO₂ concentrations in the Bay of Algeciras (Spain). *Stoch. Environ. Res. Risk Assess.* **2021**, *4*, 1–21. [[CrossRef](#)]
67. González-Enrique, J.; Ruiz-Aguilar, J.; Moscoso-López, J.; Urda, D.; Deka, L.; Turias, I. Artificial Neural Networks, Sequence-to-Sequence LSTMs, and Exogenous Variables as Analytical Tools for NO₂ (Air Pollution) Forecasting: A Case Study in the Bay of Algeciras (Spain). *Sensors* **2021**, *21*, 1770. [[CrossRef](#)]
68. Van Roode, S.; Ruiz-Aguilar, J.J.; González-Enrique, J.; Turias, I.J. An artificial neural network ensemble approach to generate air pollution maps. *Environ. Monit. Assess.* **2019**, *191*, 727. [[CrossRef](#)]