Review

Soft Computing Applications in Air Quality Modeling: Past, Present, and Future

Muhammad Muhitur Rahman ^{1,*} , Md Shafiullah ² , Syed Masiur Rahman ³,
Abu Nasser Khondaker ³, Abduljamiu Amao ⁴ and Md. Hasan Zahir ²

¹ Department of Civil and Environmental Engineering, College of Engineering, King Faisal University, Al-Ahsa 31982, Saudi Arabia

² Center of Research Excellence in Renewable Energy, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia; shafiullah@kfupm.edu.sa (M.S.); hzahir@kfupm.edu.sa (M.H.Z.)

³ Center for Environment & Water, Research Institute, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia; smrahman@kfupm.edu.sa (S.M.R.); nasserk@kfupm.edu.sa (A.N.K.)

⁴ Center for Integrative Petroleum Research, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia; amao@kfupm.edu.sa

* Correspondence: mrahman@kfu.edu.sa

Received: 24 March 2020; Accepted: 7 May 2020; Published: 14 May 2020

Abstract: Air quality models simulate the atmospheric environment systems and provide increased domain knowledge and reliable forecasting. They provide early warnings to the population and reduce the number of measuring stations. Due to the complexity and non-linear behavior associated with air quality data, soft computing models became popular in air quality modeling (AQM). This study critically investigates, analyses, and summarizes the existing soft computing modeling approaches. Among the many soft computing techniques in AQM, this article reviews and discusses artificial neural network (ANN), support vector machine (SVM), evolutionary ANN and SVM, the fuzzy logic model, neuro-fuzzy systems, the deep learning model, ensemble, and other hybrid models. Besides, it sheds light on employed input variables, data processing approaches, and targeted objective functions during modeling. It was observed that many advanced, reliable, and self-organized soft computing models like functional network, genetic programming, type-2 fuzzy logic, genetic fuzzy, genetic neuro-fuzzy, and case-based reasoning are rarely explored in AQM. Therefore, the partially explored and unexplored soft computing techniques can be appropriate choices for research in the field of air quality modeling. The discussion in this paper will help to determine the suitability and appropriateness of a particular model for a specific modeling context.

Keywords: Adaptive neuro-fuzzy inference system; artificial neural networks; air quality model; deep learning; ensemble model; evolutionary techniques; fuzzy logic model; review; soft computing model; support vector machine

1. Introduction

Air pollutants cause widespread detrimental effects on physical, biological, and economic systems. The prominent and dangerous pollutants are carbon oxides (CO_x), nitrogen oxides (NO_x), Sulphur oxides (SO_x), ozone (O₃), lead (Pb), respirable suspended particles or particulate matters (PM_{2.5} and PM₁₀), and total volatile organic compounds (TVOC) [1–4]. Exposure to such pollutants causes many diseases, including respiratory diseases, type 2 diabetes, asthma, allergies, and cancer [5–9]. Typically, the environmental regulatory agencies regulate the atmospheric quality against high air pollutant concentrations to protect human health by minimizing different detrimental effects. Air quality models play a vital role in assessing the quality of the atmosphere. They can stimulate the health of

the atmosphere and provide increasing domain knowledge and reliable forecasting. Air pollutant prediction models can provide early warnings, and their effective utilization can reduce significant numbers of measuring and data acquisition stations [10–12]. Therefore, scientists and researchers have explored many predictive tools, such as autoregressive integrated moving average (ARIMA) [13], bias adjustment [14], linear unbiased estimator [15], principal component regression (PCR) [16], non-parametric regression [17], machine learning [18], and hybrid [19] techniques for AQM.

Artificial Intelligence-based (AI) models, especially artificial neural networks (ANN) or the universal approximation approach of any non-linear function, provide flexible, efficient, less assumption dependent, and adaptive modeling techniques for environmental systems. The capability of different forms of neural networks have been investigated for modeling many air pollutants, including CO_x [20], NO_x [20,21], O₃ [22], PM_{2.5}, and PM₁₀ [23], and SO_x [21] in different locations throughout the world. Another machine-learning tool, namely the support vector machine (SVM) follows the structural risk minimization principle. It performs better than the empirical risk minimization process in minimizing the error of the training data. It maps the training data into a high dimensional feature space (Hilbert space) and turns the nonlinear regression function into a linear one [24–26]. Researchers have also employed this tool for the prediction of air pollutant concentrations [27–31]. Besides, evolutionary computational algorithm-driven ANN and SVM models have also been explored in the field of AQM, for instance a genetic-algorithm-tuned ANN modeled the hourly PM₁₀ concentrations in Greece [31].

The fuzzy logic model (FLM) is a powerful tool to deal with the complex engineering problems that are difficult to solve using traditional mathematical models. It can capture the impreciseness of linguistic terms in statements of natural language [32–34]. Jorquera et al. [35] have illustrated the effectiveness of FLM in predicting daily maximum O₃ levels. Besides, the combination of neural networks and fuzzy logic models (neuro-fuzzy model) has also received widespread attention in AQM due to their adaptiveness and better generalization performance. For instance, the neuro-fuzzy systems and FLM were employed to model concentrations of O₃ and PM₁₀ [36] and short-term hydrocarbon (HC) concentrations [37]. Moreover, deep learning and deep multi-output long short-term memory (LSTM) neural networks (deep learning) modeled air pollutant concentrations in the work of [38] and [39], respectively. Likewise, the ensemble models are also popular for classification and regression problems that use few learning machines to learn partial solutions for a given problem and combine the solutions to obtain a more generalized but complete solution to the original problem. A model for nitrogen dioxide (NO₂) concentrations in Spain was developed using an ensemble technique [40]. Furthermore, hybrid models combining data processing techniques and various above-mentioned soft computing structures were also investigated for modeling air pollutant concentrations [41–45]. Other models, including GMDH [46], HMM [47], and random forest partitioned (RFP) technique [48], were also explored for AQM in a few cases.

Different kinds of soft computing approaches have been explored in air quality modeling throughout the world. The growing number of studies can create confusion in the selection of appropriate soft computing models amongst the available models. Therefore, there is a continuous need to review articles to guide decision-makers, academics, and researchers. However, only a few review articles are available that discuss soft computing techniques in air quality modeling [11,18,49–51]. They are mostly limited to either artificial neural networks or deep learning techniques. Articles covering the whole spectrum of the available soft computing techniques can rarely be found. Besides, in recent years, other AI-based models such as functional networks, genetic programming, type-2 fuzzy logic, genetic fuzzy, genetic neuro-fuzzy, and case-based reasoning models have been successfully used to solve many complex non-linear regression problems, including oil-yield [52], particle size in a fluidized bed reactor [53], flood flows [54], and power system stability [55]. However, such models have not yet been explored significantly in air quality modeling for the prediction of the air pollutant concentration levels. Besides, evolutionary optimization (EO) technique-driven ANN, SVM, extreme learning machines (ELM), and neuro-fuzzy models are also obtaining popularity for their accuracy in predicting problems without EO-driven models. Hence, similar EO-driven techniques can also be

explored for air quality modeling. Moreover, the ELM, GMDH, HMM, and RFP models are rarely observed in the current literature for air quality prediction. Furthermore, different hybrid models can also be explored for AQM. Therefore, considering the importance of the AQM issue throughout the world, this study critically reviews and analyzes the existing soft computing models and provides guidelines for future research directions on the exploration of many soft computing models.

This paper is divided into five main sections. The second section will present the analysis of the input selection approaches for soft computing models in air quality modeling. The third section critically evaluates the proposed machine learning models in the literature. The fourth section investigates the potential soft computing models that are assumed to be suitable for air quality modeling. The fifth section reports the conclusions of this study.

2. Input and Output Selection Approaches

Researchers throughout the world mostly use meteorological and air pollutant data for modeling air quality, employing soft computing techniques as inputs. In a few cases, the influence of temporal data, traffic data, satellite data, and geographic data have also been employed. However, researchers rarely used direct industrial emission data, power plant emission data, and other types of data. Table 1 provides the components of different types of data used as inputs for the soft computing techniques in AQM. It also includes the list of modeled air pollutants.

Table 1. Inputs and outputs of the soft computing models.

Inputs			Outputs
<i>Meteorological data</i>	<i>Air pollutant data</i>	<i>Geographical data</i>	API or AQI
General weather condition	AQI	Altitude	CO _x
Temperatures	CO _x	Longitude	NO _x
Wind speed	NH ₃	Latitude	O ₃
Wind direction	NO _x	<i>Sustainability and economic parameters</i>	Pb
Wind bearing	O ₃		PM _{2.5}
Atmospheric pressure	Pb		PM ₁₀ (TSP/RSP)
Relative humidity	PM _{2.5}		SO _x
Solar radiation	PM ₁₀ (TSP or RSP)		TVOC
Sunshine duration	SO _x		
Precipitation/rain	TVOC		
Air mass origin			
Moisture content			
Dew point			
Urban heat island			
Visibility			
Cloud cover			
Stability class			
Mixing height			
Planetary layer height			
Solar elevation			
Friction velocity			
<i>Temporal data</i>	<i>Traffic data</i>	<i>Other data types</i>	
Hour of the day	Vehicle movement	Direct industrial and thermal power	
Day of the week and month	Vehicle volume	plant data	
Month of the year	Vehicle emission	Satellite data (aerosol optical depth)	
Weekday or weekend day	Vehicle type (two- or	Daily fire pixel observations	
Sine and cosine of the hour	three-wheelers, diesel- or	Drilling diameter	
Sine of the weekday	gasoline-powered)	Moisture and slit content	
		Rock mass density	
		Rebound hardness number	

To reduce the volume of the input space by selecting most dominant variables, many different techniques, including cross-correlation analysis (CCA), principal component analysis (PCA), random forest (RF), learning vector quantization (LVQ), and rough set theory (RST), have been explored in literature. Among the mentioned techniques, CCA is the simplest in analyzing the relationship of the multiple time series data. This linear data analysis approach can also be employed to select the

inputs of non-linear models [56]. The deployment of CCA reduced the number of input variables and provided better generalization performance while predicting O_3 and PM_{10} concentrations in Milan, Italy [57]. The PCA is another modern and popular tool in data analysis for extracting relevant information from complex datasets [58–60]. This simple and non-parametric method reduces the dimensionality of a given complex dataset by identifying the meaningful underlying variables known as principal components. It provides the optimal information preserving transformations within the class of linear methods [61]. However, typical PCA is suitable for two-dimensional data collected from a steady-state process that contains linear relationships between the variables. Still, in reality, the data are multidimensional and contain nonlinearity [62]. To address such limitations, several extensions, including hierarchical PCA [63], dynamic PCA [64], nonlinear PCA [65], and online adaptive PCA [66], are often employed. Researchers have employed PCA techniques while predicting PM_{10} and $PM_{2.5}$ [41], O_3 [42], RSP [67], an air quality index (AQI), or an air pollution index (API) [68].

The RF is a straightforward technique to compute the contributions of each variable to the decision-making in a high dimensional dataset. It creates multiple trees using classification and regression trees (CART) that grow fully without pruning [69–71]. Researchers have employed RF to model air pollutant concentrations including $PM_{2.5}$ [71], SO_2 [72,73], NO_2 [72], CO [72], PM_{10} [72], and O_3 [72]. The RST is another easily understandable and mature mathematical tool that successfully determines the smallest attribute subset that represents similar knowledge of the original attribute set [74]. It has already made tremendous achievements in many fields, including pattern recognition, medical diagnostics, machine learning, and knowledge discovery in the database [75]. Lei et al. [76] simplified the sample structure by reducing unnecessary attributes using RST to improve the training speed of the wavelet neural network while modeling AQI. Besides, the LVQ is an adaptive nearest-neighbor pattern classifier technique. It consists of four layers (input, competitive, linear, and output) that are based on the competitive learning of training data. Details of the LVQ processes can be found in the work of [77]. Jiang et al. [78] employed the LVQ techniques to reduce the number of input variables while modeling $PM_{2.5}$ concentrations using a neuro-fuzzy system. Other data processing techniques including wavelet decomposition [47,79], complete ensemble empirical mode decomposition with adaptive noise and variational mode decomposition (CEEMD-ANVMD) [80], fuzzy k-means (FKM), fuzzy c-means (FCM) [81], and others (logarithmic, scaling, and normalization) [82], were also employed to select input variables. Furthermore, the normalization and standardization of the data is a regularly used data pre-processing technique [11].

However, missing data is a regularly encountered phenomenon in AQM that hampers the overall processes and accuracy of the modeling techniques. This can occur due to a wide range of factors, such as errors in the measurement processes, faults in data acquisition, and insufficient sampling. Researchers have recovered the missing data in several well-established techniques in AQM, including linear interpolation [40,47,81], multivariate imputation by chained equations [83], and the expectation maximization (EM) imputation method [47,84]. In many cases, they also followed list-wise and pair-wise deletion approaches in dealing with the missing data [11]. Finally, the accuracy and precision of the AQM techniques are justified using several statistical performance indices (SPI). The SPI includes bias ratio (BR), correlation coefficient (R), determination coefficient (R^2), index of agreement (IA), mean absolute error (MAE), mean absolute percentage error (MAPE), mean bias error (MBE), the proportion of systematic error (PSE), root mean squared error (RMSE), time taken (tt), and variance ratio (VR) [85].

3. Analysis of Available Soft Computing Models

Primarily, artificial neural networks (ANN), support vector machines (SVM), evolutionary ANN and SVM, and fuzzy logic and neuro-fuzzy systems are widely used in AQM. Besides, the hybrid techniques combining several soft computing methods are also receiving widespread attention in AQM. The deep learning machine, ensemble model, and other types of soft computing techniques have been employed in a few cases for modeling purposes. An extensive number of papers that use soft

computing approaches are summarized in Table 2. This section reviews and analyses such modeling approaches thoroughly.

3.1. Artificial Neural Networks Models

ANNs are very popular in modeling complex and non-linear engineering problems as they are capable of parallel computing abilities and adaptive to external disturbances. In general, they consist of input, hidden, and output layers. The hidden layers process the input variables employing different squashing functions and send them to the output layer. They are also flexible, less assumption dependent, and adaptive in modeling environmental issues, making them popular in AQM [86]. Naturally, they have salient advantages over traditional statistical models in air quality forecasting [87]. The tested ANN structures for air pollutants prediction include multilayer perceptron neural networks (MLP-NN), radial basis function neural networks (RBF-NN), square multilayer perceptron neural networks (SMLP-NN), ward neural networks (W-NN), pruned neural networks (P-NN), recursive neural networks (R-NN), general regression neural networks (GR-NN), graph convolutional neural network (GC-NN), and backpropagation neural networks (BP-NN) (Table 2). This subsection provides critical reviews of neural network-based approaches in air quality modeling throughout the world.

Hrust et al. [20] have employed ANN to forecast four air pollutants (NO_2 , O_3 , CO, and PM_{10}) in Zagreb, Croatia as a function of meteorological variables and concentrations of the respective pollutants. Biancofiore et al. [23] predicted daily average PM_{10} and $\text{PM}_{2.5}$ concentrations 1-day to 3-days ahead in Pescara, Italy employing R-NN using meteorological data and air pollutant (i.e., PM_{10} and CO) concentrations. Radojević et al. [21] employed ward and general regression neural networks (W-NN and GR-NN) to estimate the daily average concentrations of SO_2 and NO_x in Belgrade, Serbia using meteorological variables and periodic parameters (hour of the day, day of the week, and month of the year) as inputs. Pawlak et al. [22] developed an ANN model to predict the maximum hourly mean of surface O_3 concentrations for the next day in rural and urban areas in central Poland. The model used six input variables, including four forecasted meteorological parameters, recorded the O_3 concentration of the previous day, and the month. Corani [57] compared the performance of the FF-NN model with the pruned neural network (P-NN) and lazy learning for predicting O_3 and PM_{10} concentrations in Milan, Italy. The author used air pollutant data (O_3 , NO, and NO_2) and meteorological data (pressure, temperature, wind speed, solar radiation, rain, and humidity) as the inputs for the prediction of O_3 concentrations. For PM_{10} concentration, the predictions used only two air pollutant data (PM_{10} and SO_2) and two meteorological data (temperature and pressure). The lazy learning model outperformed other models in terms of correlation and mean absolute error.

Anderetta et al. [87] proposed an MLP-NN model with a Bayesian learning scheme to forecast hourly SO_2 concentration levels in the industrial area of Ravenna, Italy, emphasizing the high levels of SO_2 occurring during relatively rare episodes. The authors employed historical meteorological data and SO_2 concentrations of current and previous hours as inputs. Biancofiore et al. [88] modeled hourly O_3 concentration 1-h, 3-h, 6-h, 12-h, 24-h, and 48-h ahead using meteorological data, a photochemical parameter (measured UVA radiation), and air pollutant concentrations (O_3 and NO_2 concentration). The authors used R-NN architecture and measured data from central Italy. Some of the modeling exercises have placed emphasis on emission sources. Gualtieri et al. [89] have demonstrated the capability of ANN in predicting short-term hourly PM_{10} concentrations in Brescia, Italy. The authors used hourly atmospheric pollutant (NO_x , SO_2 , $\text{PM}_{2.5}$, and PM_{10}) concentrations, meteorological parameters, and road traffic counts (municipality boundary and city center traffic volumes) as the inputs.

Table 2. Summary of the soft computing approaches employed for air quality modeling.

Year	Ref.	Study Region	Model Name	Outputs	Year	Ref.	Study Region	Model Name	Outputs
ANN Models					Fuzzy Logic and Neuro-Fuzzy Models				
1993	[90]	Šoštanj, Slovenia	MLP-NN	SO ₂	1995	[91]	Japan	Neuro-fuzzy system	CO
1994	[92]	Mexico	ANN	O ₃	1998	[35]	Santiago, Chile	Fuzzy logic model	O ₃
1999	[93]	UK	MLP-NN	O ₃	2003	[37]	Villa San Giovanni, Italy	Neuro-fuzzy system	HC
1999	[94]	Central London, UK	MLP-NN	NO _x	2004	[95]	Seoul, Korea	Neuro-fuzzy system	O ₃
2000	[96]	Santiago, Chile	MLP-NN	PM _{2.5}	2006	[97]	Zonguldak, Turkey	ANFIS	SO ₂ and TSP
2001	[98]	Santiago, Chile	ANN	PM _{2.5}	2007	[99]	Mexico City, Mexico	Genetic fuzzy System	O ₃
2002	[100]	Delhi, India	LM-NN	SO ₂	2009	[36]	Northern Italy	Neuro-fuzzy system	O ₃ and PM ₁₀
2002	[101]	Kuwait	ANN	O ₃	2010	[102]	Delhi, India	Neuro-fuzzy system	CO
2005	[103]	Belgium	ANN	PM ₁₀	2012	[104]	Dukla, Czech Republic	Fuzzy logic model	O ₃
2005	[57]	Milan, Italy	FF-NN	O ₃ and PM ₁₀	2012	[105]	Bogotá, Colombia	Genetic fuzzy System	PM ₁₀
2005	[106]	US–Mexico border	MLP-NN, RBF-NN, and SMLP-NN	PM _{2.5}	2013	[107]	Johor Bahru, Malaysia	Fuzzy time series	API
2006	[108]	Carcaixent, Spain	MLP-NN	O ₃	2013	[109]	Tehran, Iran	ANFIS	O ₃
2006	[110]	Santiago, Chile	MLP-NN	PM ₁₀	2013	[111]	Bor, Serbia	ANFIS	SO ₂
2008	[112]	Ontario, Canada	FF-NN and B-NN	O ₃	2015	[113]	Delhi, India	Neuro-fuzzy system	PM _{2.5}
2008	[114]	Istanbul, Turkey	FF-NN	SO ₂ , PM ₁₀ , and CO	2015	[115]	Romania	ANFIS	PM ₁₀
2009	[20]	Zagreb, Croatia	ANN	NO ₂ , O ₃ , CO, and PM ₁₀	2016	[116]	Howrah, India	ANFIS	SO ₂ , NO ₂ , CO, O ₃ , and PM ₁₀
2010	[117]	Istanbul, Turkey	FF-NN	SO ₂ , PM ₁₀ , and CO	2017	[12]	Jeddah, Saudi Arabia	ANFIS	O ₃
2013	[118]	New South Wales, Australia	RBF-NN	O ₃	2017	[78]	China	Neuro-fuzzy system	PM _{2.5}

Table 2. Cont.

Year	Ref.	Study Region	Model Name	Outputs	Year	Ref.	Study Region	Model Name	Outputs
2013	[109]	Tehran, Iran	ANN	O ₃	2017	[119]	Queensland, Australia	ANFIS	PM _{2.5}
2013	[120]	Northeast Spain	MLP-NN	PM ₁₀	2018	[121]	Queensland, Australia	ANFIS	NO ₂
2013	[107]	Johor Bahru, Malaysia	MLP-NN	API	2019	[122]	Mid-southern Taiwan	Fuzzy inference system	PM _{2.5} and Pb
2014	[123]	Kocaeli, Turkey	ANN	PM ₁₀	2020	[124]	Delhi, India	ANFIS	PM _{2.5}
2014	[125]	Athens, Greece	ANN	O ₃	2020	[126]	Tehran, Iran	ANFIS	CO, SO ₂ , O ₃ , and NO ₂
2014	[127]	Nagercoil, India	MPL and RBF NN	O ₃	Ensembles Models				
2015	[128]	Australia	RBF-NN	O ₃	2012	[129]	Europe and North America	Ensemble model	O ₃
2015	[130]	Lisbon, Portugal	BP-NN	PM ₁₀	2015	[131]	Mexico City, Mexico	Two ensemble techniques	O ₃
2016	[132]	Beijing City, China	ANN	PM _{2.5}	2016	[133]	Harbin and Chongqing, China	Decomposition-ensemble method	PM _{2.5}
2016	[134]	Algiers, Algeria	ANN	PM ₁₀	2017	[135]	Guangzhou and Lanzhou, China	Decomposition-ensemble learning paradigm	PM _{2.5}
2016	[136]	Santiago, Chile	FF-NN	PM _{2.5}	2017	[137]	Shandong Province, China	Ensemble model	PM _{2.5}
2016	[138]	Seoul, Korea	ANN	PM ₁₀	2018	[139]	Beijing, China	Deep spatial-temporal ensemble model	PM _{2.5}
2017	[140]	Canada	MLP-NN and ELM	O ₃ , PM _{2.5} , and NO ₂	2018	[79]	Shenyang and Chengdu, China	Secondary-decomposition-ensemble	PM _{2.5}
2017	[141]	Tabriz, Iran	MLP-NN	NO _x	2018	[142]	Beijing, China	Stacked ensemble method	PM _{2.5}
2018	[89]	Brescia, Italy	ANN	PM ₁₀	2019	[143]	Italy	Ensemble model	PM ₁₀ and PM _{2.5}

Table 2. Cont.

Year	Ref.	Study Region	Model Name	Outputs	Year	Ref.	Study Region	Model Name	Outputs
2018	[144]	Northern China	ANN	O ₃	2019	[145]	China	Ensemble model	AQI
2019	[21]	Belgrade, Serbia	W-NN and GR-NN	SO _x and NO _x	2019	[146]	United States	Ensemble model	PM _{2.5}
2019	[22]	Central Poland	ANN	O ₃	2019	[40]	Madrid, Spain	Ensemble neural networks	NO ₂
2019	[147]	Andalusia, Spain	BP-NN and RBF-NN	PM ₁₀	2019	[148]	London, UK	Ensemble neural networks	O ₃ and PM ₁₀
2019	[149]	São Carlos-SP, Brazil	MLP-NN and NARX-NN	PM ₁₀	2019	[150]	London, UK	Ensemble learning technique	SO ₂
2019	[151]	Rio de Janeiro, Brazil	ANN	PM _{2.5}	2019	[152]	London, UK	Ensemble data mining	NO ₂
2019	[153]	Shiraz, Iran	NARX-NN	CO	2019	[154]	Tamil Nadu, India	Bagging ensemble	O ₃
2019	[155]	Bangi, Malaysia	ANN	O ₃	2019	[156]	Northern China	Stacking ensemble	PM _{2.5}
2020	[157]	Kanpur, India	MLP-NN	O ₃	2019	[158]	Southern China	BP-NN ensembles	PM _{2.5}
2020	[159]	USA, UK, and Italy	ANN	O ₃	2020	[160]	Tehran, Iran	Nonlinear ensemble model	AQI
2020	[161]	Five regions, Taiwan	BP-NN	PM _{2.5}	2020	[162]	Beijing, Tianjin, and Hebei of China	Stacking-driven ensemble model	PM _{2.5}
2020	[163]	Nan, Thailand	MLP-NN	PM ₁₀	Hybrid and Other Models				
Deep Learning Models					2002	[67]	Downtown, Hong Kong	Hybrid (PCA-RBF-NN)	PM ₁₀
2015	[88]	Central Italy	R-NN	O ₃	2007	[42]	Oporto, Portugal	Hybrid (PCA-FFNN)	O ₃
2016	[132]	Beijing City, China	Deep learning	PM _{2.5}	2008	[43]	Szeged, Hungary	PCA-SVM and PCA-ANN	NO _x
2017	[38]	Aarhus city, Denmark	Deep learning	O ₃	2008	[164]	Kuwait	Hybrid (PCR, ANN, and PCA assisted ANN)	O ₃
2017	[23]	Pescara, Italy	R-NN	PM ₁₀ and PM _{2.5}	2011	[41]	Greece and Finland	Hybrid (PCA-ANN)	PM ₁₀ and PM _{2.5}
2018	[165]	China	Deep learning	PM ₁₀	2011	[82]	Nantou, Taiwan	Hybrid	PM ₁₀ and NO _x

Table 2. Cont.

Year	Ref.	Study Region	Model Name	Outputs	Year	Ref.	Study Region	Model Name	Outputs
2018	[166]	Kuwait	Deep learning	O ₃	2011	[45]	Dilovasi, Turkey	Hybrid (PCA-MLP-NN)	O ₃
2019	[39]	Taipei City, Taiwan.	LSTM	PM _{2.5} , PM ₁₀ , and NO _x	2012	[47]	California and Texas, USA	HMM technique with Gamma distribution	O ₃
2019	[167]	Beijing, China.	Deep learning	PM _{2.5}	2012	[46]	Rub' Al Khali, Saudi Arabia	GMDH technique	O ₃
2019	[168]	Guangdong, China	Deep learning	PM _{2.5}	2012	[169]	Tehran, Iran	Type-2 fuzzy	CO
2019	[170]	Tehran, Iran	Deep learning	PM _{2.5}	2013	[171]	Delhi, India	PCA-ANN	AQI
2019	[73]	Beijing, China	Deep neural network	PM _{2.5} , NO ₂ , and SO ₂	2013	[172]	Gansu, China	Hybrid	O ₃
2020	[161]	Five regions, Taiwan	CNN and LSTM	PM _{2.5}	2015	[173]	Düzce Province, Turkey	Hybrid (PCA-ANN)	PM ₁₀
SVM Models					2015	[85]	Guangzhou, China	Hybrid	PM _{2.5}
2002	[174]	Hong Kong	SVM	RSP	2015	[81]	Salamanca, Mexico	Hybrid	PM ₁₀
2003	[175]	Mong Kok, Hong Kong	SVM and PCA-RBF-NN	RSP	2015	[176]	China	Hybrid (wavelet-based MLP-NN)	PM _{2.5}
2005	[177]	Hong Kong	SVM and RBF-NN	NO _x and RSP	2015	[178]	Delhi, India	PSO-ANFIS	SO ₂ and O ₃
2014	[29]	Rio de Janeiro, Brazil	SVM and ANN	O ₃	2015	[179]	Taiyuan, China	Hybrid ANN and hybrid SVM	PM ₁₀ and SO ₂
2014	[180]	Tehran, Iran	RBF-SVM	O ₃	2015	[181]	Agra, India	PCA-ANN	NO ₂
2014	[182]	India	SVM	O ₃	2016	[183]	Nan'an, China	Hybrid (wavelet-based BP-NN)	PM ₁₀ , SO ₂ , and NO ₂
2016	[132]	Beijing City, China	SVM	PM _{2.5}	2016	[184]	London, UK	Hybrid (PCA-ANN)	PM ₁₀ and PM _{2.5}
2017	[185]	Beijing, Tianjin, and Shijiazhuang of China	Multi-dimensional collaborative SVR	AQI	2017	[83]	Tunisia	Hybrid (RF and SVM)	PM ₁₀
2018	[28]	Spain	SVM	PM ₁₀	2017	[44]	Central London, UK	PCA-MLP-NN	NO ₂
2018	[30]	Singapore	SVM and BP-NN	CO ₂ and TVOC	2017	[68]	Taichang, China	Hybrid (PCA-GA-ANN)	AQI
2018	[186]	Beijing, China	Space-time SVM and global SVM.	PM _{2.5}	2017	[187]	Kunming and Yuxi, China	Hybrid (CI-FPA-SVM)	PM _{2.5} and PM ₁₀
2019	[188]	Beijing, China	SVR	AQI and NO _x	2017	[189]	Hebei Province, China	Hybrid method (PCA-CS-SVM)	PM _{2.5}

Table 2. Cont.

Year	Ref.	Study Region	Model Name	Outputs	Year	Ref.	Study Region	Model Name	Outputs
2020	[163]	Nan, Thailand	SVR	PM ₁₀	2017	[190]	Shangdianzi, China	Hybrid	PM _{2.5}
Evolutionary NN and SVM Models					2018	[191]	Kolkata, India	Type-2 fuzzy	AQI
2003	[192]	Downtown, Hong Kong	PSO-ANN	CO, NO _x , RSP	2018	[193]	Datong, Taiwan	ANFIS-WELM, ANFIS, WELM, GA-BPNN	CO, NO, PM _{2.5} , and PM ₁₀
2004	[194]	Northern Greece	GA-FF-NN	PM ₁₀	2018	[195]	Egypt	Hybrid	PM ₁₀
2004	[196]	Helsinki, Finland	GA-MLP-NN	NO ₂	2018	[197]	CZT, China	Hybrid Grey–Markov	PM ₁₀
2006	[31]	Athens, Greece	GA-ANN	PM ₁₀	2019	[198]	Taipei, Taiwan	Hybrid (multi-output and multi-tasking SVM)	PM _{2.5}
2013	[199]	European countries	GA-GR-NN	PM ₁₀	2019	[200]	Varanasi, India	Hybrid (PCA and ANN)	PM ₁₀
2014	[201]	New Zealand	GA-ANN	NO ₂	2019	[71]	Yancheng, China	Hybrid	PM _{2.5}
2015	[202]	Helsinki, Finland	GA-NN	PM ₁₀ and PM _{2.5}	2019	[72]	Hangzhou, China	Hybrid (RF and R-NN)	SO ₂ , NO ₂ , CO, PM _{2.5} , PM ₁₀ , and O ₃
2016	[203]	Aghdasiyeh, Iran	GA-BP-NN	PM ₁₀	2019	[204]	Jing-Jin-Ji, China	Hybrid model	PM _{2.5}
2017	[205]	India	GWO-SVM	AQI	2019	[76]	China	Hybrid	Indoor AQ
2018	[206]	China	GA-BP-NN	PM _{2.5}	2019	[48]	Wrocław, Poland	RFP technique	NO ₂
2019	[84]	Beijing, China	PSO-SVM	PM _{2.5}	2019	[80]	Beijing and Yibin, China	Hybrid	PM _{2.5}
2019	[207]	Beijing, China	PSO-SVM and GA-SVM	AQI	2019	[208]	China	Hybrid (wavelet-based ARIMA, ANN, and SVM)	PM _{2.5}
2019	[209]	Coc Sau, Vietnam	PSO-SVM	PM ₁₀	2019	[210]	China	CEEMD-CSA-GWO-SVR	NO ₂ and SO ₂
2019	[211]	Beijing, China	QPSO, PSO, GA, and GS-based SVM	NO ₂ and PM _{2.5}	2020	[212]	China	SCA-ELM	AQI
2020	[163]	Nan province, Thailand	GA-MLP-NN and GA-SVM	PM ₁₀	2020	[124]	Delhi, India	Wavelet-ANFIS-PSO and wavelet-ANFIS-GA	PM _{2.5}

Mlakar et al. [90] have illustrated how MLP-NN forecasted SO_2 concentrations for the next interval of 30 min in Šoštanj, Slovenia using wind data, air temperature, actual and historical SO_2 concentrations, emissions from the thermal power plant, and the time of day. Gómez-Sanchis et al. [108] have developed MLP-NN models to estimate ambient O_3 concentrations for an agriculture training center; the concentrations of the next three days were predicted independently using the same training center in Carcaixent, Spain using surface meteorological variables and vehicle emission variables as predictors. Spellman [93] and Moustris et al. [125] developed models using relevant data from multiple locations. Spellman [93] proposed an MLP-NN for estimating summer surface O_3 concentrations at five locations (London, Harwell, Birmingham, Leeds, and Strath Vaich) in the United Kingdom using surface meteorological variables as predictors. The authors found better generalization performance of the MLP-NN over conventional regression model. Moustris et al. The authors of [125] predicted maximum daily surface O_3 concentration in several locations in the greater Athens area, Greece using FF-NN. The authors used hourly meteorological data (wind speed and direction, air temperature, solar irradiance, barometric pressure, and relative humidity) and air pollution data (8-h average values of surface O_3 concentration and hourly NO_2 concentrations) as inputs for their developed model. De Gennaro et al. [120] predicted 24-h average PM_{10} concentrations at two sites (Montseny and Barcelona) in Spain employing MLP-NN. The authors used hourly particulate matters concentrations of the previous day, meteorological data, and air masses origin as inputs.

Kurt et al. [114] developed an air pollution prediction system for one region of the Greater Istanbul Area, Turkey. The authors attempted three different ways to predict the concentrations of three air pollutants (SO_2 , PM_{10} , and CO) for the next three days using FF-NN, where the meteorological data (general condition, day and night temperatures, wind speed and direction, pressure, and humidity) were used as inputs in the first experiment data. In the second experiment, the concentrations of the second and third days were predicted cumulatively using the previous days' model outputs. The performance of the proposed models improved in the third experiment due to the use of the effect of the day of the week as the input parameter. Kurt and Oktay [117] updated the prediction model used in their previous research [114] by considering spatial features along with the air pollutant data from ten different air quality monitoring stations in Istanbul to forecast the same air pollutants levels three days in advance. The authors concluded that the distance-based geographic model ensured better performance when compared to the non-geographic model. Therefore, the employed input variables were the meteorological data (general condition, day and night temperatures, humidity, pressure, wind speed, and direction), periodic data (day of the week and date), and air pollutant concentrations (SO_2 , PM_{10} , and CO). Ozdemir and Taner [123] modeled PM_{10} concentrations in Kocaeli, Turkey using ANN and multiple linear regression (MLR) techniques, and achieved better prediction accuracy with the ANN technique than the MLR. They used hourly meteorological data (air temperatures, wind speed and direction, relative humidity, and air pressure) and pollutant levels (monthly average, minimum, maximum, and standard deviation of monthly PM_{10} concentrations).

Perez and Reyes [110] forecasted hourly maximum PM_{10} concentrations for the next day using MLP-NN in Santiago, Chile as a function of measured concentrations of PM_{10} until 7 p.m., and measured and forecasted the meteorological variables. The meteorological variables were the difference between the maximum and minimum temperature on the present day, the difference between the maximum and minimum forecasted temperature of the next day, and the forecasted meteorological potential of atmospheric pollution. Schornobay-Lui et al. [149] employed MLP-NN and non-linear autoregressive exogenous (NARX) neural networks (NARX-NN) to model short-term (daily) and medium-term (monthly) PM_{10} concentrations in São Carlos, Brazil. They used only monthly average meteorological data (temperature, relative humidity, wind speed, and accumulated rainfall) for monthly PM_{10} concentration prediction, whereas for predicting daily PM_{10} concentration (for the next day) they used PM_{10} concentrations along with meteorological data as inputs. Ha et al. [128] employed RBF-NN to predict the 8-h maximum average O_3 concentrations in the Sydney basin in New South Wales, Australia using NO_x and VOC emissions, ambient temperature, location coordinates, and topography

as inputs. Wahid et al. [118] estimated the 8-h maximum average of O_3 concentration in the Sydney basin employing RBF-NN. They modeled O_3 concentration as a function of NO_x emission, ambient temperature, location coordinates, and topography.

ANN models are widely explored in AQM due to their satisfactory performances. However, the models can be unstable in a few cases as they are highly dependent on data and can easily fall into local optima instead of finding the global optima [213]. In addition, network training, the amount and quality of training data, and the network parameters (number of hidden layers, transfer function type, number of epochs, number of neurons, and initial weights and biases) significantly influence their performance. A conventional ANN works efficiently when the approximated function is relatively monotonic with only a few dimensions of the input features, but may not be efficient in other cases [214]. A monotonic function is either entirely non-increasing or non-decreasing and its first derivative (which need not be continuous) does not change significantly. They experience the greatest difficulty in approximating functions when the input features are not linearly separable [215]. In typical ANN models, all input variables are connected to neurons of the hidden layer that may affect the generalization ability [216].

3.2. Support Vector Machine Models

Boser et al. [217] introduced the SVM technique, which was primarily limited to solving classification problems, to efficient data analysis. With the passage of time, the researchers extended the technique to regression problems. In regression problems, the SVM technique constructs an optimal geometric hyperplane to distinguish the available data and to map them into the higher dimensional feature space by forming a separation surface through the employment of various functions, including sigmoidal, polynomial, and radial basis functions [218,219]. While they are dealing with regression problems, the SVM is known as support vector regression. However, in this paper, the term SVM is used for both the support vector machine and regression. The structural risk minimization principle adopted in SVM seeks to minimize an upper bound of the generalization error that results in better generalization performance when compared to the conventional techniques.

Like other engineering problems, the SVM has also been employed in the AQM field. Ortiz-García et al. [27] applied SVM and MLP-NN for predicting hourly O_3 concentrations at five stations in the urban area of Madrid, Spain. They explored the influence of four previous O_3 measurements (at a given station and at the neighboring stations) and the influence of meteorological variables on O_3 predictions. Luna et al. [29] employed ANN and SVM to predict O_3 concentrations in Rio de Janeiro, Brazil using available primary pollutants (CO , NO_x , and O_3) and meteorological data (wind speed, solar irradiation, temperature, and moisture content). The authors found solar irradiation and temperature as the most dominant inputs and observed better predictability with SVM than ANN. Yang et al. [186] illustrated space-time SVM to predict hourly $PM_{2.5}$ concentrations in Beijing, China and compared the results with space-time ANN, ARIMA, and global SVM. The presented results confirmed the superiority of the space-time SVM over other techniques. The study employed historical $PM_{2.5}$ concentration data and meteorological data as inputs for the tested models. Wang et al. [175] illustrated an adaptive RBF-NN and an improved SVM for predicting RSP concentration in Mong Kok Roadside Gaseous Monitory Station, Hong Kong. The authors used six pollutant (SO_2 , NO_x , NO , NO_2 , CO , and RSP) concentrations and five meteorological variables (wind speed and direction, indoor and outdoor temperatures, and solar radiation) as inputs. The authors found better generalization performance in the SVM technique than the adaptive RBF-NN technique. However, the SVM technique was very slow compared to the RBF-NN technique. Mehdipour and Memarianfard [180] predicted tropospheric O_3 concentration in Tehran, Iran using SVM as a function of daily maximum pollutant concentrations (SO_2 , NO_2 , CO , $PM_{2.5}$, and PM_{10}) and daily average meteorological data (ambient air temperature, wind speed, and relative humidity). The authors found the RBF-kernel-function-based SVM as the best performing technique.

It is generally perceived that the SVM has higher generalization capability compared to the classical NN model, but the design of the SVM model heavily depends on the proper selection of kernel functions and their parameters. It is also reported that they can be very slow in the training phase [220].

3.3. Evolutionary Neural Network and Support Vector Machine Models

Evolutionary computation can be used to tune the weights and biases, to choose the number of hidden layers and neurons, to select the squashing functions, and to generate appropriate architectures for artificial neural networks [221–224]. They can also be employed to tune the key parameters of the SVM while solving classification or regression problems [25,209,225]. Among many evolutionary computational techniques, the genetic algorithm (GA), bat algorithm (BA), differential evolution (DE), particle swarm optimization (PSO), gravitational search algorithm (GSA), cuckoo search algorithm (CSA), sine cosine algorithm (SCA), grey wolf optimizer (GWO), and backtracking search algorithm (BSA) are well known for the aforementioned purposes.

Kapageridis and Triantafyllou [194] used GA to optimize the topology and learning parameters of time-lagged FF-NN. The model predicted a maximum 24-h moving average of PM_{10} concentration for the next day at two monitoring stations in northern Greece. The authors used meteorological data (minimum humidity, maximum temperature, the difference between the maximum and minimum temperature, and average wind speed) and the maximum PM_{10} concentration of the previous day as inputs. Niska et al. [196] introduced GA for selecting the inputs and designing the topology of an MLP-NN model for forecasting hourly NO_2 concentrations at a busy urban traffic station in Helsinki, Finland. Among the many available input variables, the authors selected NO_2 and O_3 concentrations, sine and cosine of the hour, sine of the weekday, temperature, wind speed and direction, solar elevation, and friction velocity as inputs for the GA-tuned MLP-NN-based on experience. Grivas and Chaloulakou [31] modeled hourly PM_{10} concentrations employing GA-ANN in Athens, Greece using the combination of meteorological data, air pollutant data (PM_{10} concentrations), and temporal data (hour of the day, day of the week, and seasonal index). Lu et al. [192] developed a PSO-tuned ANN approach for modeling air pollutant parameters (CO , NO_x , NO_2 , and RSP) for the downtown area of Hong Kong using original pollutant data. Liu et al. [84] illustrated the PSO-SVM technique for daily $PM_{2.5}$ grade predictions in Beijing, China as a function of meteorological data (average atmospheric pressure, relative humidity, and air temperature, maximum wind speed and direction, and cumulative precipitation) and hourly average air pollutant data ($PM_{2.5}$ and PM_{10}). The PSO-SVM displayed better accuracy compared to the GA-SVM, AdaBoost, and ANN models for $PM_{2.5}$ grade prediction. Chen et al. [207] employed a PSO-driven SVM technique to predict short-term atmospheric pollutant concentration in Temple of Heaven, Beijing, China and found a faster response for PSO-SVM over GA-SVM. Li et al. [211] employed quantum-behaved particle-swarm-optimization-assisted (QPSO) SVM to predict NO_2 and $PM_{2.5}$ concentrations for the next four hours in the Haidian District of Beijing, China. The authors used hourly measurements for five air pollutants ($PM_{2.5}$, NO_2 , CO , O_3 , and SO_2), and six meteorological parameters (weather condition, wind speed and direction, temperature, pressure, and relative humidity) as inputs for their developed model. The authors also found that their QPSO-SVM outperformed PSO-SVM, GA-SVM, and GS-SVM. The evolutionary algorithms can tackle the practical problems and obtain better generalization performance, but the process is computationally expensive.

3.4. Fuzzy Logic and Neuro-Fuzzy Models

The fuzzy logic model deals with the imprecisions and uncertainties of real-world applications using a set of manually extracted “if-then” rules [226–228]. It creates rules controlling conceivable relationships between input and output features through a fuzzification process that transforms the inferable features into the membership values. Then, it follows a defuzzification process to infer the quantifiable output according to the rules and the input data [122]. It is simple, flexible,

customizable, and can handle problems with imprecise and incomplete data. However, it is tiresome to develop rules and membership functions. Besides, it makes the analysis difficult as the outputs can be interpreted in a number of ways [229]. In response, fuzzy logic model and artificial neural networks (neuro-fuzzy systems) can be combined together, where the fuzzy logic performs as an inference mechanism under cognitive uncertainty and the neural network possesses the learning, adaptation, fault-tolerance, parallelism, and generalization capabilities [97]. Similar to many engineering problems, the neuro-fuzzy systems have also been employed to model air pollutant concentrations.

Carnevale et al. [36] have suggested a neuro-fuzzy approach to identify source-receptor models for O_3 and PM_{10} concentrations prediction in northern Italy by processing the simulations of a deterministic multi-phase modeling system. The authors used daily NO_x and VOC emissions for O_3 modeling, whereas NH_3 , NO_x , primary PM_{10} , SO_x , and VOC for PM_{10} modeling as input variables. Morabito and Versaci [37] demonstrated a fuzzy neural identification technique to forecast short-term hydrocarbon concentrations in local air at Villa San Giovanni, Italy. The authors used measured air pollutants (CO , NO , NO_2 , O_3 , PM_{10} , and SO_2) data, meteorological data, and vehicle movements as input variables. Hájek and Ole [104] predicted hourly average O_3 concentrations for the next hour at the Pardubice Dukla station, Czech Republic using air pollutant data (NO_x , NO , NO_2 , and O_3), temporal data (month of the year), and meteorological data of the present hour. They employed several soft computing techniques, including ANN, SVM, and the fuzzy logic model. The presented results confirmed the superiority of the fuzzy logic model over other models. Yildirim and Bayramoglu [97] developed an adaptive neuro-fuzzy inference system (ANFIS) for the estimation of SO_2 and total suspended particulate matter (TSP) in Zonguldak, Turkey. The authors used meteorological data (wind speed, pressure, precipitation, temperature, solar radiation, and relative humidity) and air pollutant concentrations (SO_2 or TSP) of the previous day as inputs. They found temperature and air pollutant concentration as the most dominant input variables. Yeganeh et al. [119] estimated ground-level $PM_{2.5}$ concentration using ANFIS, SVM, and BP-NN techniques in Southeast Queensland, Australia. They used satellite-based, meteorological, and geographical data as inputs. The presented results confirmed the superiority of the ANFIS model over other models. Yeganeh et al. [121] predicted the monthly mean NO_2 concentration by employing ANFIS in Queensland, Australia. The authors improved the prediction accuracy of their developed model using satellite-based and traffic data in conjunction with comprehensive meteorological and geographical data as inputs.

Tanaka et al. [91] adopted a self-organizing fuzzy identification algorithm for modeling CO concentrations for the next hour in a large city in Japan. The authors used CO concentration data, meteorological data (wind speed, temperature, and sunshine), and traffic volume for the previous four hours as inputs. Chung et al. [122] employed a fuzzy inference system (FIS) to predict $PM_{2.5}$ and Pb concentrations in mid-southern Taiwan using geographic coordinates and time as inputs. Heo and Kim [95] employed fuzzy logic and neural network models (neuro-fuzzy system) consecutively to forecast the hourly maximum O_3 concentrations for the next day at four monitoring sites in Seoul, Korea. The authors used air pollutants data (CO , NO_2 , O_3 , and SO_2) and meteorological data (wind speed and direction, temperature, relative humidity, and solar radiation) as input variables. The prediction accuracy of their developed model was continually verified and augmented through corrective measures. Mishra et al. [113] analyzed the haze episodes and developed a neuro-fuzzy system to forecast $PM_{2.5}$ concentrations during haze episodes in the urban area of Delhi, India. The air pollutants (CO , O_3 , NO_2 , SO_2 , and $PM_{2.5}$) and meteorological parameters (pressure, temperature, wind speed and direction, relative humidity, visibility, and dew point temperature) were used as inputs. The presented results confirmed the superiority of the neuro-fuzzy model over ANN and MLR models. Jain and Khare [102] developed a neuro-fuzzy model for predicting hourly ambient CO concentrations at urban intersections and roadways in Delhi, India. The authors used measured hourly CO data, meteorological data (sunshine hours, wind speed and direction, humidity, temperature, pressure, cloud cover, visibility, stability class, and mixing height), and traffic data (two-wheelers, three-wheelers,

diesel-powered, and gasoline-powered vehicles) as inputs. The authors adopted the correlation matrix technique, PCA, and fuzzy Delphi method to identify the most dominant input variables.

3.5. Deep Learning Models

Deep learning is a subset of machine learning techniques that are built using artificial neural networks. It deals with a sufficiently large amount of data without pre-processing. In deep learning, the ANN takes inputs and processes those using weights and biases in many hidden layers to spit out a prediction. Weights and biases are adjusted during the training processes to make a better prediction. This process can easily overcome the bottlenecks and overfitting issues related to the shallow multi-output long short-term memory model. It is employed to model both regression and classification problems [230]. The application of deep learning models for air quality modeling has been following an increasing trend.

Zhou et al. [39] proposed a multi-step-ahead air quality forecasting technique, employing a deep multi-output long short-term memory neural network model incorporating three deep learning algorithms (i.e., mini-batch gradient descent, dropout neuron, and L2 regularization) using meteorological and air quality data as inputs. The model predicted $PM_{2.5}$, PM_{10} , and NO_x concentrations simultaneously in Taipei City, Taiwan. Besides, the model overcame the bottlenecks and overfitting issues related to the shallow multi-output long short-term memory model. Athira et al. [165] predicted PM_{10} concentrations using deep learning technique (R-NN, long short-term memory, and gated recurrent unit) in China and found gated recurrent unit as the best performer amongst the employed techniques. The authors used meteorological data as inputs to the deep learning technique. Peng et al. [132] illustrated a spatiotemporal deep learning method for hourly average $PM_{2.5}$ concentration prediction in Beijing City, China and compared the prediction results with spatiotemporal ANN, ARIMA, and SVM. According to the presented results, the deep learning technique outperformed the other techniques. The authors used the air quality data for previous intervals as inputs. Tao et al. [167] proposed a deep learning-based short-term forecasting model to predict $PM_{2.5}$ concentrations at the US Embassy in Beijing, China.

3.6. Ensemble Models

Ensemble models combine a set of individual methods by means of aggregation rules to solve a given problem and produce a better-aggregated solution. Every single method is trained individually and produces a different and/or closer solution [231–234]. In general, these models use generated solutions of the individual techniques as inputs and produce ultimate solutions using a wide range of processing methods, including averaging, voting, boosting, bagging, and stacking [235]. Averaging (for regression problems) and voting (for classification problems) are the most popular due to their simplicity and better interpretability. These models are also known as committees of learning machines.

Valput et al. [40] modeled hourly NO_2 concentration in Madrid, Spain using ensemble neural networks as a function of meteorological data, pollutant data (hourly average NO_2 , NO , CO , and SO_2), and traffic data (intensity and road occupation). Maciag et al. [148] proposed a clustering-based ensemble of evolving spiking neural networks to forecast air pollutants (O_3 and PM_{10}) concentrations up to six hours in the Greater London Area. The presented results demonstrated the superiority of the proposed ensemble technique over ARIMA, MLP-NN, and singleton spiking neural network models. Bing et al. [131] modeled daily maximum O_3 concentration in the metropolitan area of Mexico City, Mexico using four individual soft computing techniques (ANN, SVM, RF, and MLR) and two ensemble techniques (linear and greedy). Among the employed soft computing techniques, the linear ensemble model outperformed other techniques. Feng et al. [158] developed a BP-NN ensemble to forecast the daily $PM_{2.5}$ concentration in Southern China using daily assimilated surface meteorological data (including air temperature, relative humidity, pressure, and winds) and daily fire pixel observation as inputs. The authors observed significant improvement in the air pollutant prediction capability of the employed ensembles technique. Wang and Song [139] designed a deep spatial-temporal ensemble

model to improve air quality pollutant ($PM_{2.5}$) prediction in 35 stations in Beijing, China using historical air pollutant data (CO , NO_2 , SO_2 , O_3 , PM_{10} , and $PM_{2.5}$) and meteorological data as inputs. The authors found their proposed model outperformed other models including linear regression and a deep neural network.

3.7. Hybrid and Other Models

A combination of multiple techniques is known as a hybrid modeling technique and these are very popular in modeling air pollutant concentrations throughout the world. The authors of [41] employed MLP-NN to forecast the daily mean PM_{10} and $PM_{2.5}$ concentrations for the next day in Greece and Finland where the input features were selected through the PCA method.

Sousa et al. [42] compared the effectiveness of PCA-based FF-NN with MLR and original data-based FF-NN to predict hourly O_3 concentration for the next day in Oporto, Portugal. The authors considered hourly O_3 , NO , and NO_2 concentrations and hourly mean temperature, wind velocity, and relative humidity as the inputs to the predictive models. Cortina-Januch et al. [81] modeled hourly average PM_{10} concentration for the next day in Salamanca, Mexico using clustering-based ANN as a function of historical pollutant data (PM_{10} concentration) and meteorological data of the current day. They used FCM and FKM clustering approaches to process the input data. Lin et al. [82] predicted three air pollutants (PM_{10} , NO_x , and NO_2) using immune-algorithm-tuned SVM (IA-SVM) in Nantou station, Taiwan from their historical measured values. The authors preprocessed the raw data employing logarithmic, scaling, and normalization techniques. They found logarithmic IA-SVM was the most effective prediction technique in terms of overall MAPE values. The combination of ANFIS and a weighted extreme learning machine (WELM) model outperformed neuro-genetic, ANFIS and other models in predicting air pollutant (CO , NO , $PM_{2.5}$, and PM_{10}) concentration in Datong, Taiwan using measured time series air pollutant data [193]. Cheng et al. [208] employed wavelet decomposition to mine useful information from the measured weather data and used ARIMA, ANN, and SVM to predict $PM_{2.5}$ in five Chinese cities. The authors confirmed the superiority of the hybrid approaches over non-hybrid approaches through exhaustive simulations. Bai et al. [183] demonstrated wavelet-decomposition-based BP-NN technique to estimate daily air pollutant (PM_{10} , SO_2 , and NO_2) concentration in Nan'an District of Chongqing, China. The authors used the wavelet coefficients of the previously measured air pollutant concentrations and local meteorological data as inputs. Apart from the reviewed single and hybrid AQM techniques, other modeling approaches including GMDH, HMM, and RFP techniques, have also been explored to model air pollutant concentration. For instance, the authors of [46] employed a GMDH approach to model O_3 concentration in Rub' Al Khali, Saudi Arabia using meteorological data (wind speed, temperature, pressure, relative humidity) and air pollutant data (NO and NO_2). Conversely, Zhang et al. [47] illustrated an HMM with Gamma distribution to predict O_3 level in California and Texas, USA as a function of measured air pollutant data (O_3 , NO , and NO_2) and measured and predicted meteorological data (temperature, wind speed, relative humidity, and solar radiation). The authors used the wavelet decomposition technique to reduce the input data size by extracting useful features. The authors of [48] proposed an RFP technique to model NO_2 concentration in Wrocław, Poland by using hourly meteorological data (air temperature, wind speed and direction, relative humidity, and atmospheric pressure), hourly traffic counts, and temporal data (month of the year and day of the week).

3.8. Generalized Overview

The authors of [236] and [237] proved that the single-hidden-layer feed forward neural networks are the universal approximators of continuous target functions. Similarly, Hammer and Gersmann [238] illustrated the universal function approximation capabilities of the support vector machine. Researchers investigated different variations of the fuzzy logic systems and demonstrated their capabilities as universal function approximators. Ying [239], proved that Takagi–Sugeno (TS) fuzzy systems are the universal approximators for a multi-input and single-output problem. Therefore,

the mentioned soft computing models have become theoretically valid candidates for modeling air pollutants. The models attempt to build relations between pollutant levels with the information related to other relevant pollutants, pollutant characteristics, meteorological conditions and pattern, and terrain types and characteristics. If multiple models are used for same air quality modeling problems, it is expected that the model performance will mainly depend on the selection of appropriate inputs, the size of datasets, the run-time, the hyper-parameters of the models, and the inherent model suitability for the specific problem in multi-dimensional space.

Therefore, researchers have started to model air pollutants using ANN, SVM, and fuzzy logic approaches from the beginning. With the passage of time, it has been observed that evolutionary optimization techniques are playing vital roles in achieving better generalization performance by tuning the key parameters of the primarily employed soft computing approaches in AQM. Likewise, researchers moved from fuzzy logic models to neuro-fuzzy and ANFIS models that enhance overall performance. The ensemble models exploit the advantages of the individual models and cutoff the drawbacks (if there are any). Besides, the deep learning methods are better candidates than the shallow machine learning techniques for handling massive and verities of information received from the advanced measurement infrastructures and sensors. Hence, the ensemble models perform better than the individual models and the deep learning models have been exhibiting superiority, especially for large datasets.

In general, environmental data are complex to model as they are functions of several variables that share a complicated and non-linear relationship with each other [240]. The soft computing models that address the limitations of other models and ensure higher prediction accuracies and reliability for a specific pollutant in specific air quality episode should not be considered as effective, accurate, and reliable in all cases. The discussed notion motivates the research for the exploration and exploitation of more techniques in the field of AQM that can vary based on area, pollutant type, and data availability.

4. Potential Soft Computing Models and Approaches

Among many potential techniques, different variations of artificial neural networks, evolutionary fuzzy and neuro-fuzzy models, ensemble and hybrid models, and knowledge-based models should be further explored. Besides, there is a continuous need for the development of a universal model, as most of the explored models are either site-dependent or pollutant dependent. This section discusses future research directions and potential soft computing models that can be investigated in air quality modeling throughout the world.

4.1. Variations of ANN Models

As can be observed from Section 3, ANN approaches were widely explored in AQM and in most cases MLP-NN, BP-NN, RBF-NN, or R-NN were employed. Other available variations of ANN (GR-NN, GC-NN, P-NN, W-NN, and others) models that successfully demonstrated their capabilities in modeling complex and non-linear problems in other engineering fields have not been explored significantly [11]. Many of them (extreme learning machine, multitasking, probabilistic, time delay, modular, and other hybrid neural networks) are rarely explored. Besides, deep neural network models received great attention in modeling PM_{2.5} concentrations, but other air pollutants have not been modeled significantly. Therefore, such unexplored and rarely explored variations of the neural networks can be investigated in future works for modeling all types of air pollutant concentrations.

4.2. Evolutionary Fuzzy and Neuro-Fuzzy Models

Fuzzy systems are the proven tools for many applications for modeling complex and non-linear problems. However, the lack of learning capabilities in the fuzzy systems has encouraged researchers to augment their capabilities by hybridizing them with the EO techniques [241–244]. Among the many EO techniques, GA, GWO, CSA, SCA, and PSO are widely used and well-known global search optimization approaches with the ability to explore a large search space for suitable solutions [245–247].

Besides, the type-2 fuzzy set is capable of handling more uncertainties than the type-1 fuzzy set that has been successfully applied in a wide range of areas [248–250]. Therefore, considering the potentiality of the fuzzy logic approaches, these can be explored in the field of AQM.

4.3. Group Method Data Handling Models and Functional Network Models

Long-term research in the field of neural networks and advanced statistical methods has contributed to the evolution of an abductive induction mechanism that is known as GMDH [251]. It automatically synthesizes abductive networks from a database of inputs and outputs with complex and nonlinear relationships. Other extensions of the neural network models include the functional network models (FNM) [252]. This determines the structure of a network and data using domain knowledge and estimates unknown neuron functions. Both GMDH and FNM were explored in many relevant applications [253–255]. These rarely explored extensions of the neural networks can be further investigated in AQM.

4.4. Case-Based Reasoning and Knowledge-Based Models

Case-based reasoning solves new problems by recalling the experiences and solutions of similar past problems [256]. It deals with the given problems following four steps, namely retrieve, reuse, revise, and retain [257]. Another soft computing technique, the knowledge-based system, attempts to solve problems by giving advice in a domain and utilizing the knowledge provided by a human expert [258]. Researchers have employed both techniques to solve many complex problems [259–261]. These techniques can be investigated in AQM, as none of them have yet been explored.

4.5. Ensemble and Hybrid Models

As discussed earlier, ensemble models employ multiple learning techniques in parallel and combine their outputs to produce a better generalization performance. In a real-world situation, they aim to manage the strengths and weaknesses of each model and end up with the best possible solutions [262]. Recently, such models received huge momentum in modeling AQM, but this was limited to a few specific pollutants (mainly PM_{2.5}). Researchers should invest more time into these attractive tools as they will become some of the most prominent tools for AQM in the future.

4.6. Development of Universal Models

Most of the discussed models are either site dependent or pollutant dependent. There is no guarantee that a specific model developed for a specific site will be stable and reliable for another location with different meteorological conditions. Therefore, there is always a need for the development of a universal model for AQM. Besides, the comparison between the site-specific models could be an attractive option for future research as it aids in developing site characterizations. Such research may enable the creation of guidelines for site-specific model development.

4.7. Appropriate Input Selection Methods

As discussed in Section 2, several approaches have been reported to reduce the input space by selecting the most dominant input variables. In addition, most of the approaches selected air pollutant and meteorological data as inputs. A few of the considered other types of data, including temporal, traffic, geographical, and sustainable data. Therefore, the present authors believe that the comparison of such input selection methods considering all available input data types could be an attractive field of research in AQM. Besides, the selection of proper decomposition components for the reduction of data dimensionality could be considered as another potential research direction, as the inclusion of many components in input space may result in model complexity and the accumulation of errors. Moreover, other available data pre-processing and feature extraction techniques employed for relevant fields could also be explored.

5. Conclusions

Soft computing models have become very popular in air quality modeling as they can efficiently model the complexity and non-linearity associated with air quality data. This article critically reviewed and discussed existing soft computing modeling approaches. Among the many available soft computing techniques, the artificial neural networks with variations of structures and the hybrid modeling approaches combining several techniques were widely explored in predicting air pollutant concentrations throughout the world. Other approaches, including support vector machines, evolutionary artificial neural networks and support vector machines, fuzzy logic, and neuro-fuzzy systems, have also been used in air quality modeling for several years. Recently, deep learning and ensemble models have received huge momentum in modeling air pollutant concentrations due to their wide range of advantages over other available techniques. Additionally, this research reviewed and listed all possible input variables for air quality modeling. It also discussed several input selection processes, including cross-correlation analysis, principal component analysis, random forest, learning vector quantization, rough set theory, and wavelet decomposition techniques. Besides, this article sheds light on several data recovery approaches for missing data, including linear interpolation, multivariate imputation by chained equations, and expectation-maximization imputation methods.

Finally, it proposed many advanced, reliable, and self-organizing soft computing models that are rarely explored and/or not explored in the field of air quality modeling. For instance, functional neural network models, variations of neural network models, evolutionary fuzzy and neuro-fuzzy systems, type-2 fuzzy logic models, group method data handling, case-based reasoning, ensemble, and hybrid models, and knowledge-based systems have the immense potential for modeling air pollutant concentrations. Moreover, the modelers can compare the effectiveness of several input selection processes to find the most suitable one for air quality modeling. Furthermore, they can attempt to build universal models instead of developing site-specific and pollutant-specific models. The authors believe that the findings of this review article will help researchers and decision-makers in determining the suitability and appropriateness of a particular model for a specific modeling context.

Author Contributions: Conceptualization, S.M.R.; methodology, S.M.R., and M.S.; formal analysis, S.M.R., M.S., A.N.K., A.A., and M.H.Z.; data curation, S.M.R., M.S., A.N.K., A.A., M.H.Z., and M.M.R.; writing—original draft preparation, S.M.R., M.S., and A.N.K.; writing—review and editing, M.M.R., A.A., and M.H.Z.; funding acquisition, M.M.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Deanship of Scientific Research at King Faisal University under Nasher Track (Grant No. 186357).

Acknowledgments: The co-authors acknowledge the support received from King Fahd University of Petroleum & Minerals.

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

List of Acronyms

Acronym	Explication
AI	Artificial Intelligence
ANFIS	Adaptive Neuro-Fuzzy Inference System
ANN	Artificial Neural Network
ANN	Artificial Neural Networks
ANVMD	Adaptive Noise and Variational Mode Decomposition
AQM	Air Quality Modeling
ARIMA	Autoregressive Integrated Moving Average
BA	Bat Algorithm
BP-NN	Backpropagation Neural Networks
BR	Bias Ratio

BSA	Backtracking Search Algorithm
CART	Classification and Regression Trees
CBR	Case-Based Reasoning
CCA	Cross-Correlation Analysis
CEEMD	Complimentary Ensemble Empirical Mode Decomposition
CH ₄	Methane
CO	Carbon Monoxide
CO ₂	Carbon Dioxide
CO _x	Oxides of Carbon
CSA	Cuckoo Search Algorithm
CW-SVM	Chance Theory Assisted Weighted Support Vector Machine
DE	Differential Evolution
EELM	Ensemble Extreme Learning Machine
ELM	Extreme Learning Machine
EM	Expectation Maximization
EO	Evolutionary Optimization
FCM	Fuzzy c-Means
FF-NN	Feed Forward Neural Network
FIS	Fuzzy Inference System
FKM	Fuzzy k-Means
FLM	Fuzzy Logic Model
FNM	Functional Network Models
FPA	Flower Pollination Algorithm
GA	Genetic Algorithm
GC-NN	Graph Convolutional Neural Network
GMDH	Group Method Data Handling
GR-NN	General Regression Neural Networks
GSA	Gravitational Search Algorithm
GWO	Grey Wolf Optimizer
HC	Hydrocarbons
HCHO	Formaldehyde
HMM	Hidden Markov Model
IA	Index of Agreement
IA-SVM	Immune-Algorithm-Tuned Support Vector Machine
KBS	Knowledge-Based System
LM-NN	Levenberg–Marquardt Neural Network
LVQ	Learning Vector Quantization
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MBE	Mean Bias Error
MLP-NN	Multilayer Perceptron Neural Networks
MLR	Multiple Linear Regression
NARX-NN	Non-Linear Autoregressive Exogenous Neural Networks
NH ₃	Ammonia
NMHC	Non-Methane Hydrocarbons
NO	Nitric Oxide
NO ₂	Nitrogen Dioxide
NO _x	Oxides of Nitrogen
O ₃	Ozone
Pb	Lead
PCA	Principal Component Analysis
PCR	Principal Component Regression
PM	Particulate Matters
P-NN	Pruned Neural Networks
P-NN	Pruned Neural Network

PSE	Proportion of Systematic Error
PSO	Particle Swarm Optimization
QPSO	Quantum-Behaved Particle Swarm Optimization
R	Correlation Coefficient
R ²	Determination Coefficient
RBF-NN	Radial Basis Function Neural Networks
RF	Random Forest
RFP	Random Forest Partitioned
RMSE	Root Mean Squared Error
R-NN	Recursive Neural Networks
RSP	Respirable Suspended Particle
RST	Rough Set Theory
RST	Rough Set Theory
SCA	Sine Cosine Algorithm
SMMLP-NN	Square Multilayer Perceptron Neural Networks
SO ₂	Sulfur Dioxide
SO _x	Oxides of Sulphur
SPI	Statistical Performance Indices
SVM	Support Vector Machine
TSP	Total Suspended Particulate
TSP	Total Suspended Particulate Matter
TT	Time Taken
TVOC	Total Volatile Organic Compounds
VARMA	Vector Autoregressive Moving-Average
VR	Variance Ratio
WELM	Weighted Extreme Learning Machine
W-NN	Ward Neural Networks

References

- Li, Y.; Guan, D.; Tao, S.; Wang, X.; He, K. A review of air pollution impact on subjective well-being: Survey versus visual psychophysics. *J. Clean. Prod.* **2018**, *184*, 959–968. [[CrossRef](#)]
- Ortíz-Alvarez, A.; Peralta, O.; Alvarez-Ospina, H.; Martínez-Arroyo, A.; Castro, T.; Páramo, V.H.; Ruiz-Suárez, L.G.; Garza, J.; Saavedra, I.; de la Luz Espinosa, M.; et al. Concentration profile of elemental and organic carbon and personal exposure to other pollutants from brick kilns in Durango, Mexico. *Air Qual. Atmos. Heal.* **2018**, *11*, 285–300. [[CrossRef](#)]
- Zhang, J.; Zhang, L.; Du, M.; Zhang, W.; Huang, X.; Zhang, Y.; Yang, Y.; Zhang, J.; Deng, S.; Shen, F.; et al. Identifying the major air pollutants base on factor and cluster analysis, a case study in 74 Chinese cities. *Atmos. Environ.* **2016**, *144*, 37–46. [[CrossRef](#)]
- Chen, Y.; Du, W.; Zhuo, S.; Liu, W.; Liu, Y.; Shen, G.; Wu, S.; Li, J.; Zhou, B.; Wang, G.; et al. Stack and fugitive emissions of major air pollutants from typical brick kilns in China. *Environ. Pollut.* **2017**, *224*, 421–429. [[CrossRef](#)] [[PubMed](#)]
- Hvidtfeldt, U.A.; Sørensen, M.; Geels, C.; Ketzel, M.; Khan, J.; Tjønneland, A.; Overvad, K.; Brandt, J.; Raaschou-Nielsen, O. Long-term residential exposure to PM_{2.5}, PM₁₀, black carbon, NO₂, and ozone and mortality in a Danish cohort. *Environ. Int.* **2019**, *123*, 265–272. [[CrossRef](#)] [[PubMed](#)]
- Ansari, M.; Ehrampoush, M.H. Meteorological correlates and AirQ+ health risk assessment of ambient fine particulate matter in Tehran, Iran. *Environ. Res.* **2019**, *170*, 141–150. [[CrossRef](#)]
- Liu, F.; Chen, G.; Huo, W.; Wang, C.; Liu, S.; Li, N.; Mao, S.; Hou, Y.; Lu, Y.; Xiang, H. Associations between long-term exposure to ambient air pollution and risk of type 2 diabetes mellitus: A systematic review and meta-analysis. *Environ. Pollut.* **2019**, *252*, 1235–1245. [[CrossRef](#)]
- Chen, X.; Wang, X.; Huang, J.; Zhang, L.; Song, F.; Mao, H.; Chen, K.; Chen, J.; Liu, Y.; Jiang, G.; et al. Nonmalignant respiratory mortality and long-term exposure to PM₁₀ and SO₂: A 12-year cohort study in northern China. *Environ. Pollut.* **2017**, *231*, 761–767. [[CrossRef](#)]

9. Liu, H.; Liu, S.; Xue, B.; Lv, Z.; Meng, Z.; Yang, X.; Xue, T.; Yu, Q.; He, K. Ground-level ozone pollution and its health impacts in China. *Atmos. Environ.* **2018**, *173*, 223–230. [\[CrossRef\]](#)
10. Alimisis, A.; Philippopoulos, K.; Tzanis, C.G.; Deligiorgi, D. Spatial estimation of urban air pollution with the use of artificial neural network models. *Atmos. Environ.* **2018**, *191*, 205–213. [\[CrossRef\]](#)
11. Cabaneros, S.M.; Calautit, J.K.; Hughes, B.R. A review of artificial neural network models for ambient air pollution prediction. *Environ. Model. Softw.* **2019**, *119*, 285–304. [\[CrossRef\]](#)
12. Taylan, O. Modelling and analysis of ozone concentration by artificial intelligent techniques for estimating air quality. *Atmos. Environ.* **2017**, *150*, 356–365. [\[CrossRef\]](#)
13. Reikard, G. Volcanic emissions and air pollution: Forecasts from time series models. *Atmos. Environ. X* **2019**, *1*, 100001. [\[CrossRef\]](#)
14. Silibello, C.; D'Allura, A.; Finardi, S.; Bolignano, A.; Sozzi, R. Application of bias adjustment techniques to improve air quality forecasts. *Atmos. Pollut. Res.* **2015**, *6*, 928–938. [\[CrossRef\]](#)
15. Sozzi, R.; Bolignano, A.; Ceradini, S.; Morelli, M.; Petenko, I.; Argentini, S. Quality control and gap-filling of PM₁₀ daily mean concentrations with the best linear unbiased estimator. *Environ. Monit. Assess.* **2017**, *189*, 562. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Kumar, A.; Goyal, P. Forecasting of air quality in Delhi using principal component regression technique. *Atmos. Pollut. Res.* **2011**, *2*, 436–444. [\[CrossRef\]](#)
17. Donnelly, A.; Misstear, B.; Broderick, B. Real time air quality forecasting using integrated parametric and non-parametric regression techniques. *Atmos. Environ.* **2015**, *103*, 53–65. [\[CrossRef\]](#)
18. Rybarczyk, Y.; Zalakeviciute, R.; Rybarczyk, Y.; Zalakeviciute, R. Machine Learning Approaches for Outdoor Air Quality Modelling: A Systematic Review. *Appl. Sci.* **2018**, *8*, 2570. [\[CrossRef\]](#)
19. Zhu, J.; Wu, P.; Chen, H.; Zhou, L.; Tao, Z.; Zhu, J.; Wu, P.; Chen, H.; Zhou, L.; Tao, Z. A Hybrid Forecasting Approach to Air Quality Time Series Based on Endpoint Condition and Combined Forecasting Model. *Int. J. Environ. Res. Public Health* **2018**, *15*, 1941. [\[CrossRef\]](#)
20. Hrust, L.; Klaić, Z.B.; Križan, J.; AntoniĆ, O.; Hercog, P. Neural network forecasting of air pollutants hourly concentrations using optimised temporal averages of meteorological variables and pollutant concentrations. *Atmos. Environ.* **2009**, *43*, 5588–5596. [\[CrossRef\]](#)
21. Radojević, D.; Antanasijević, D.; Perić-Grujić, A.; Ristić, M.; Pocajt, V. The significance of periodic parameters for ANN modeling of daily SO₂ and NO_x concentrations: A case study of Belgrade, Serbia. *Atmos. Pollut. Res.* **2019**, *10*, 621–628. [\[CrossRef\]](#)
22. Pawlak, I.; Jarosławski, J.; Pawlak, I.; Jarosławski, J. Forecasting of Surface Ozone Concentration by Using Artificial Neural Networks in Rural and Urban Areas in Central Poland. *Atmosphere* **2019**, *10*, 52. [\[CrossRef\]](#)
23. Biancofiore, F.; Busilacchio, M.; Verdecchia, M.; Tomassetti, B.; Aruffo, E.; Bianco, S.; Di Tommaso, S.; Colangeli, C.; Rosatelli, G.; Di Carlo, P. Recursive neural network model for analysis and forecast of PM₁₀ and PM_{2.5}. *Atmos. Pollut. Res.* **2017**, *8*, 652–659. [\[CrossRef\]](#)
24. Jothilakshmi, S.; Gudivada, V.N. Large scale data enabled evolution of spoken language research and applications. In *Handbook of Statistics*; Elsevier B.V.: Amsterdam, The Netherlands, 2016; Volume 35, pp. 301–340, ISBN 9780444637444.
25. Shafiullah, M.; Ijaz, M.; Abido, M.A.; Al-Hamouz, Z. Optimized Support Vector Machine & Wavelet Transform for Distribution Grid Fault Location. In Proceedings of the 2017 11th IEEE International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG), Cadiz, Spain, 4–6 April 2017; pp. 77–82.
26. Shahriar, M.S.; Shafiullah, M.; Rana, M.J. Stability enhancement of PSS-UPFC installed power system by support vector regression. *Electr. Eng.* **2017**, 1–12. [\[CrossRef\]](#)
27. Ortiz-García, E.G.; Salcedo-Sanz, S.; Pérez-Bellido, Á.M.; Portilla-Figueras, J.A.; Prieto, L. Prediction of hourly O₃ concentrations using support vector regression algorithms. *Atmos. Environ.* **2010**, *44*, 4481–4488. [\[CrossRef\]](#)
28. García Nieto, P.J.; Sánchez Lasheras, F.; García-Gonzalo, E.; de Cos Juez, F.J. PM₁₀ concentration forecasting in the metropolitan area of Oviedo (Northern Spain) using models based on SVM, MLP, VARMA and ARIMA: A case study. *Sci. Total Environ.* **2018**, *621*, 753–761. [\[CrossRef\]](#)
29. Luna, A.S.; Paredes, M.L.L.; de Oliveira, G.C.G.; Corrêa, S.M. Prediction of ozone concentration in tropospheric levels using artificial neural networks and support vector machine at Rio de Janeiro, Brazil. *Atmos. Environ.* **2014**, *98*, 98–104. [\[CrossRef\]](#)

30. Chen, S.; Mihara, K.; Wen, J. Time series prediction of CO₂, TVOC and HCHO based on machine learning at different sampling points. *Build. Environ.* **2018**, *146*, 238–246. [[CrossRef](#)]
31. Grivas, G.; Chaloulakou, A. Artificial neural network models for prediction of PM₁₀ hourly concentrations, in the Greater Area of Athens, Greece. *Atmos. Environ.* **2006**, *40*, 1216–1229. [[CrossRef](#)]
32. Tanaka, K.; Sugeno, M. Introduction to fuzzy modeling. In *Fuzzy Systems*; Springer: New York, NY, USA, 1998; pp. 63–89.
33. Hossain, M.I.; Khan, S.A.; Shafiullah, M.; Hossain, M.J. Design and implementation of MPPT controlled grid connected photovoltaic system. In Proceedings of the 2011 IEEE Symposium on Computers & Informatics, Kuala Lumpur, Malaysia, 20–23 March 2011; pp. 284–289.
34. Hoon Joo, Y.; Chen, G. Fuzzy systems modeling. In *Encyclopedia of Artificial Intelligence*; IGI Global: Hershey, PA, USA, 2011; pp. 734–743.
35. Jorquera, H.; Pérez, R.; Cipriano, A.; Espejo, A.; Victoria Letelier, M.; Acuña, G. Forecasting ozone daily maximum levels at Santiago, Chile. *Atmos. Environ.* **1998**, *32*, 3415–3424. [[CrossRef](#)]
36. Carnevale, C.; Finzi, G.; Pisoni, E.; Volta, M. Neuro-fuzzy and neural network systems for air quality control. *Atmos. Environ.* **2009**, *43*, 4811–4821. [[CrossRef](#)]
37. Morabito, F.C.; Versaci, M. Fuzzy neural identification and forecasting techniques to process experimental urban air pollution data. *Neural Networks* **2003**, *16*, 493–506. [[CrossRef](#)]
38. Ghoneim, O.A.; Doreswamy; Manjunatha, B.R. Forecasting of ozone concentration in smart city using deep learning. In *Proceedings of the 2017 International Conference on Advances in Computing, Communications and Informatics, ICACCI 2017*; Institute of Electrical and Electronics Engineers Inc.: Piscataway, NJ, USA, 2017; Volume 2017, pp. 1320–1326.
39. Zhou, Y.; Chang, F.-J.; Chang, L.-C.; Kao, I.-F.; Wang, Y.-S. Explore a deep learning multi-output neural network for regional multi-step-ahead air quality forecasts. *J. Clean. Prod.* **2019**, *209*, 134–145. [[CrossRef](#)]
40. Valput, D.; Navares, R.; Aznarte, J.L. Forecasting hourly NO₂ concentrations by ensembling neural networks and mesoscale models. *Neural. Comput. Appl.* **2019**. [[CrossRef](#)]
41. Voukantsis, D.; Karatzas, K.; Kukkonen, J.; Räsänen, T.; Karppinen, A.; Kolehmainen, M. Intercomparison of air quality data using principal component analysis, and forecasting of PM₁₀ and PM_{2.5} concentrations using artificial neural networks, in Thessaloniki and Helsinki. *Sci. Total Environ.* **2011**, *409*, 1266–1276. [[CrossRef](#)] [[PubMed](#)]
42. Sousa, S.I.V.; Martins, F.G.; Alvim-Ferraz, M.C.M.; Pereira, M.C. Multiple linear regression and artificial neural networks based on principal components to predict ozone concentrations. *Environ. Model. Softw.* **2007**, *22*, 97–103. [[CrossRef](#)]
43. Juhos, I.; Makra, L.; Tóth, B. Forecasting of traffic origin NO and NO₂ concentrations by Support Vector Machines and neural networks using Principal Component Analysis. *Simul. Model. Pract. Theory* **2008**, *16*, 1488–1502. [[CrossRef](#)]
44. Cabaneros, S.M.S.; Calautit, J.K.S.; Hughes, B.R. Hybrid Artificial Neural Network Models for Effective Prediction and Mitigation of Urban Roadside NO₂ Pollution. *Energy Procedia* **2017**, *142*, 3524–3530. [[CrossRef](#)]
45. Özbay, B.; Keskin, G.A.; Doğruparmak, Ş.Ç.; Ayberk, S. Predicting tropospheric ozone concentrations in different temporal scales by using multilayer perceptron models. *Ecol. Inform.* **2011**, *6*, 242–247. [[CrossRef](#)]
46. Rahman, S.M.; Khondaker, A.N.; Abdel-Aal, R. Self organizing ozone model for Empty Quarter of Saudi Arabia: Group method data handling based modeling approach. *Atmos. Environ.* **2012**, *59*, 398–407. [[CrossRef](#)]
47. Zhang, H.; Zhang, W.; Palazoglu, A.; Sun, W. Prediction of ozone levels using a Hidden Markov Model (HMM) with Gamma distribution. *Atmos. Environ.* **2012**, *62*, 64–73. [[CrossRef](#)]
48. Kamińska, J.A. A random forest partition model for predicting NO₂ concentrations from traffic flow and meteorological conditions. *Sci. Total Environ.* **2019**, *651*, 475–483. [[CrossRef](#)] [[PubMed](#)]
49. Ayturan, Y.A.; Ayturan, Z.C.; Altun, H.O. Air pollution modelling with deep learning: A review. *Inter. J. Environ. Pollut. Environ. Model.* **2018**, *1*, 58–62.
50. Zhou, K.; Xie, R. Review of neural network models for air quality prediction. In *Proceedings of the Advances in Intelligent Systems and Computing AISC*; Springer: Berlin/Heidelberg, Germany, 2020; Volume 1117, pp. 83–90.
51. Iskandaryan, D.; Ramos, F.; Trilles, S. Air Quality Prediction in Smart Cities Using Machine Learning Technologies based on Sensor Data: A Review. *Appl. Sci.* **2020**, *10*, 2401. [[CrossRef](#)]

52. Roy, K.; Mukherjee, A.; Jana, D.K. Prediction of maximum oil-yield from almond seed in a chemical industry: A novel type-2 fuzzy logic approach. *South African J. Chem. Eng.* **2019**, *29*, 1–9. [\[CrossRef\]](#)
53. Razzak, S.A.; Shafiullah, M.; Rahman, S.M.; Hossain, M.M.; Zhu, J. A Multigene Genetic Programming approach for modeling effect of particle size in a liquid–solid circulating fluidized bed reactor. *Chem. Eng. Res. Des.* **2018**, *134*, 370–381. [\[CrossRef\]](#)
54. Walton, R.; Binns, A.; Bonakdari, H.; Ebtehaj, I.; Gharabaghi, B. Estimating 2-year flood flows using the generalized structure of the Group Method of Data Handling. *J. Hydrol.* **2019**, *575*, 671–689. [\[CrossRef\]](#)
55. Shafiullah, M.; Juel Rana, M.; Shafiul Alam, M.; Abido, M.A. Online Tuning of Power System Stabilizer Employing Genetic Programming for Stability Enhancement. *J. Electr. Syst. Inf. Technol.* **2018**. [\[CrossRef\]](#)
56. Balaguer Ballester, E.; Camps, I.; Valls, G.; Carrasco-Rodriguez, J.; Soria Olivas, E.; del Valle-Tascon, S. Effective 1-day ahead prediction of hourly surface ozone concentrations in eastern Spain using linear models and neural networks. *Ecol. Modell.* **2002**, *156*, 27–41. [\[CrossRef\]](#)
57. Corani, G. Air quality prediction in Milan: Feed-forward neural networks, pruned neural networks and lazy learning. *Ecol. Modell.* **2005**, *185*, 513–529. [\[CrossRef\]](#)
58. Shlens, J. A Tutorial on Principal Component Analysis. *arXiv* **2014**, arXiv:1404.1100.
59. Wise, B.M.; Ricker, N.L.; Veltkamp, D.J. *Upset and Sensor Failure Detection in Multivariate Processes*; University of Washington: Seattle, WA, USA, 1989.
60. Joback, K.G. *A Unified Approach to Physical Property Estimation Using Multivariate Statistical Techniques*; Massachusetts Institute of Technology: Cambridge, MA, USA, 1984.
61. Fukunaga, K.; Koontz, W.L.G. Application of the Karhunen-Loève Expansion to Feature Selection and Ordering. *IEEE Trans. Comput.* **1970**, *C-19*, 311–318. [\[CrossRef\]](#)
62. Bakshi, B.R. Multiscale PCA with application to multivariate statistical process monitoring. *AIChE J.* **1998**, *44*, 1596–1610. [\[CrossRef\]](#)
63. MacGregor, J.F.; Kourti, T. Statistical process control of multivariate processes. *Control Eng. Pract.* **1995**, *3*, 403–414. [\[CrossRef\]](#)
64. Kresta, J.V.; Macgregor, J.F.; Marlin, T.E. Multivariate statistical monitoring of process operating performance. *Can. J. Chem. Eng.* **1991**, *69*, 35–47. [\[CrossRef\]](#)
65. Kramer, M.A. Nonlinear principal component analysis using autoassociative neural networks. *AIChE J.* **1991**, *37*, 233–243. [\[CrossRef\]](#)
66. Wold, S. Exponentially weighted moving principal components analysis and projections to latent structures. *Chemom. Intell. Lab. Syst.* **1994**, *23*, 149–161. [\[CrossRef\]](#)
67. Lu, W.Z.; Wang, W.J.; Fan, H.Y.; Leung, A.Y.T.; Xu, Z.B.; Lo, S.M.; Wong, J.C.K. Prediction of Pollutant Levels in Causeway Bay Area of Hong Kong Using an Improved Neural Network Model. *J. Environ. Eng.* **2002**, *128*, 1146–1157. [\[CrossRef\]](#)
68. Mu, B.; Li, S.; Yuan, S. An improved effective approach for urban air quality forecast. In *Proceedings of the 2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)*; IEEE: Piscataway, NJ, USA, 2017; pp. 935–942.
69. Dubey, A. Feature Selection Using Random Forest. Available online: <https://towardsdatascience.com/feature-selection-using-random-forest-26d7b747597f> (accessed on 28 August 2019).
70. Breiman, L.; Friedman, J.H.; Olshen, R.A.; Stone, C.J. *Classification and Regression Trees*; Routledge: Abingdon, UK, 2017; ISBN 9781315139470.
71. Shang, Z.; Deng, T.; He, J.; Duan, X. A novel model for hourly PM_{2.5} concentration prediction based on CART and EELM. *Sci. Total Environ.* **2019**, *651*, 3043–3052. [\[CrossRef\]](#)
72. Feng, R.; Zheng, H.; Gao, H.; Zhang, A.; Huang, C.; Zhang, J.; Luo, K.; Fan, J. Recurrent Neural Network and random forest for analysis and accurate forecast of atmospheric pollutants: A case study in Hangzhou, China. *J. Clean. Prod.* **2019**, *231*, 1005–1015. [\[CrossRef\]](#)
73. Li, J.; Shao, X.; Sun, R. A DBN-Based Deep Neural Network Model with Multitask Learning for Online Air Quality Prediction. *J. Control Sci. Eng.* **2019**, *2019*, 1–9. [\[CrossRef\]](#) [\[PubMed\]](#)
74. Sailaja, N.V.; Sree, L.P.; Mangathayaru, N. Rough set based feature selection approach for text mining. In *Proceedings of the 2016 2nd International Conference on Contemporary Computing and Informatics, IC3I 2016*; Institute of Electrical and Electronics Engineers Inc.: Piscataway, NJ, USA, 2016; pp. 40–45.
75. Zhang, Q.; Xie, Q.; Wang, G. A survey on rough set theory and its applications. *CAAI Trans. Intell. Technol.* **2016**, *1*, 323–333. [\[CrossRef\]](#)

76. Lei, L.; Chen, W.; Xue, Y.; Liu, W. A comprehensive evaluation method for indoor air quality of buildings based on rough sets and a wavelet neural network. *Build. Environ.* **2019**, 106296. [\[CrossRef\]](#)
77. Liu, J.; Zuo, B.; Zeng, X.; Vroman, P.; Rabenasolo, B. Nonwoven uniformity identification using wavelet texture analysis and LVQ neural network. *Expert Syst. Appl.* **2010**, 37, 2241–2246. [\[CrossRef\]](#)
78. Jiang, P.; Dong, Q.; Li, P. A novel hybrid strategy for PM_{2.5} concentration analysis and prediction. *J. Environ. Manag.* **2017**, 196, 443–457. [\[CrossRef\]](#)
79. Gan, K.; Sun, S.; Wang, S.; Wei, Y. A secondary-decomposition-ensemble learning paradigm for forecasting PM_{2.5} concentration. *Atmos. Pollut. Res.* **2018**, 9, 989–999. [\[CrossRef\]](#)
80. Zhao, F.; Li, W. A Combined Model Based on Feature Selection and WOA for PM_{2.5} Concentration Forecasting. *Atmosphere* **2019**, 10, 223. [\[CrossRef\]](#)
81. Cortina-Januchs, M.G.; Quintanilla-Dominguez, J.; Vega-Corona, A.; Andina, D. Development of a model for forecasting of PM₁₀ concentrations in Salamanca, Mexico. *Atmos. Pollut. Res.* **2015**, 6, 626–634. [\[CrossRef\]](#)
82. Lin, K.P.; Pai, P.F.; Yang, S.L. Forecasting concentrations of air pollutants by logarithm support vector regression with immune algorithms. *Appl. Math. Comput.* **2011**, 217, 5318–5327. [\[CrossRef\]](#)
83. Ben Ishak, A. Variable selection based on statistical learning approaches to improve PM₁₀ concentration forecasting. *J. Environ. Informatics* **2017**, 30, 79–94. [\[CrossRef\]](#)
84. Liu, W.; Guo, G.; Chen, F.; Chen, Y. Meteorological pattern analysis assisted daily PM_{2.5} grades prediction using SVM optimized by PSO algorithm. *Atmos. Pollut. Res.* **2019**. [\[CrossRef\]](#)
85. Liu, D.; Li, L.; Liu, D.; Li, L. Application Study of Comprehensive Forecasting Model Based on Entropy Weighting Method on Trend of PM_{2.5} Concentration in Guangzhou, China. *Int. J. Environ. Res. Public Health* **2015**, 12, 7085–7099. [\[CrossRef\]](#) [\[PubMed\]](#)
86. Karatzas, K.; Kaltsatos, S. Air pollution modelling with the aid of computational intelligence methods in Thessaloniki. *Simul. Model. Pract. Theory* **2007**, 15, 1310–1319. [\[CrossRef\]](#)
87. Andretta, M.; Eleuteri, A.; Fortezza, F.; Manco, D.; Mingozzi, L.; Serra, R.; Tagliaferri, R. Neural networks for sulphur dioxide ground level concentrations forecasting. *Neural Comput. Appl.* **2000**, 9, 93–100. [\[CrossRef\]](#)
88. Biancofiore, F.; Verdecchia, M.; Di Carlo, P.; Tomassetti, B.; Aruffo, E.; Busilacchio, M.; Bianco, S.; Di Tommaso, S.; Colangeli, C. Analysis of surface ozone using a recurrent neural network. *Sci. Total Environ.* **2015**, 514, 379–387. [\[CrossRef\]](#)
89. Gualtieri, G.; Carotenuto, F.; Finardi, S.; Tartaglia, M.; Toscano, P.; Gioli, B. Forecasting PM₁₀ hourly concentrations in northern Italy: Insights on models performance and PM₁₀ drivers through self-organizing maps. *Atmos. Pollut. Res.* **2018**, 9, 1204–1213. [\[CrossRef\]](#)
90. Boznar, M.; Lesjak, M.; Mlakar, P. A neural network-based method for short-term predictions of ambient SO₂ concentrations in highly polluted industrial areas of complex terrain. *Atmos. Environ. Part B. Urban Atmos.* **1993**, 27, 221–230. [\[CrossRef\]](#)
91. Tanaka, K.; Sano, M.; Watanabe, H. Modeling and control of carbon monoxide concentration using a neuro-fuzzy technique. *IEEE Trans. Fuzzy Syst.* **1995**, 3, 271–279. [\[CrossRef\]](#)
92. Ruiz-Suarez, J.C.; Mayora, O.A.; Smith-Perez, R.; Ruiz-Suarez, L.G. A Neural Network-based Prediction Model of Ozone for Mexico City. *WIT Trans. Ecol. Environ.* **1994**, 3. [\[CrossRef\]](#)
93. Spellman, G. An application of artificial neural networks to the prediction of surface ozone concentrations in the United Kingdom. *Appl. Geogr.* **1999**, 19, 123–136. [\[CrossRef\]](#)
94. Gardner, M.W.; Dorling, S.R. Neural network modelling and prediction of hourly NO_x and NO₂ concentrations in urban air in London. *Atmos. Environ.* **1999**, 33, 709–719. [\[CrossRef\]](#)
95. Heo, J.-S.; Kim, D.-S. A new method of ozone forecasting using fuzzy expert and neural network systems. *Sci. Total Environ.* **2004**, 325, 221–237. [\[CrossRef\]](#) [\[PubMed\]](#)
96. Pérez, P.; Trier, A.; Reyes, J. Prediction of PM_{2.5} concentrations several hours in advance using neural networks in Santiago, Chile. *Atmos. Environ.* **2000**, 34, 1189–1196. [\[CrossRef\]](#)
97. Yildirim, Y.; Bayramoglu, M. Adaptive neuro-fuzzy based modelling for prediction of air pollution daily levels in city of Zonguldak. *Chemosphere* **2006**, 63, 1575–1582. [\[CrossRef\]](#) [\[PubMed\]](#)
98. Perez, P.; Reyes, J. Prediction of Particulate Air Pollution using Neural Techniques. *Neural Comput. Appl.* **2001**, 10, 165–171. [\[CrossRef\]](#)
99. Nebot, À.; Acosta, J.; Mugica, V. Environmental modeling by means of genetic fuzzy systems. In Proceedings of the IEEE International Conference on Fuzzy Systems, London, UK, 23–26 July 2007.

100. Chelani, A.B.; Chalapati Rao, C.; Phadke, K.; Hasan, M. Prediction of sulphur dioxide concentration using artificial neural networks. *Environ. Model. Softw.* **2002**, *17*, 159–166. [\[CrossRef\]](#)
101. Abdul-Wahab, S.; Al-Alawi, S. Assessment and prediction of tropospheric ozone concentration levels using artificial neural networks. *Environ. Model. Softw.* **2002**, *17*, 219–228. [\[CrossRef\]](#)
102. Jain, S.; Khare, M. Adaptive neuro-fuzzy modeling for prediction of ambient CO concentration at urban intersections and roadways. *Air Qual. Atmos. Heal.* **2010**, *3*, 203–212. [\[CrossRef\]](#)
103. 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. [\[CrossRef\]](#)
104. Hájek, P.; Olej, V. Ozone prediction on the basis of neural networks, support vector regression and methods with uncertainty. *Ecol. Inform.* **2012**, *12*, 31–42. [\[CrossRef\]](#)
105. Varela, C.A.R.; Rey, M.A.M.; Varela, A.R.; Nieto, L.D.A. Genetic Fuzzy System for the prediction of air pollution level by Particulate Matter—Case study: Bogotá. *Ingeniería* **2012**, *17*, 1. [\[CrossRef\]](#)
106. Ordieres, J.B.; Vergara, E.P.; Capuz, R.S.; Salazar, R.E. Neural network prediction model for fine particulate matter (PM_{2.5}) on the US–Mexico border in El Paso (Texas) and Ciudad Juárez (Chihuahua). *Environ. Model. Softw.* **2005**, *20*, 547–559. [\[CrossRef\]](#)
107. Abd Rahman, N.H.; Lee, M.H.; Latif, M.T.; Suhartono, S. Forecasting of Air Pollution Index with Artificial Neural Network. *J. Teknol.* **2013**, *63*. [\[CrossRef\]](#)
108. Gómez-Sanchis, J.; Martín-Guerrero, J.D.; Soria-Olivas, E.; Vila-Francés, J.; Carrasco, J.L.; del Valle-Tascón, S. Neural networks for analysing the relevance of input variables in the prediction of tropospheric ozone concentration. *Atmos. Environ.* **2006**, *40*, 6173–6180. [\[CrossRef\]](#)
109. Azadeh, A.; Sheikhalishahi, M.; Saberi, M.; Mostaghimi, M.H. An intelligent multivariate approach for optimum forecasting of daily ozone concentration in large metropolitans with incomplete inputs. *Int. J. Product. Qual. Manag.* **2013**, *12*, 209. [\[CrossRef\]](#)
110. Perez, P.; Reyes, J. An integrated neural network model for PM₁₀ forecasting. *Atmos. Environ.* **2006**, *40*, 2845–2851. [\[CrossRef\]](#)
111. Savic, M.; Mihajlovic, I.N.; Zivkovic, Z. An ANFIS-Based Air Quality Model for Prediction of SO₂ Concentration in Urban Area. *SSRN Electron. J.* **2013**. [\[CrossRef\]](#)
112. Solaiman, T.A.; Coulibaly, P.; Kanaroglou, P. Ground-level ozone forecasting using data-driven methods. *Air Qual. Atmos. Heal.* **2008**, *1*, 179–193. [\[CrossRef\]](#)
113. Mishra, D.; Goyal, P.; Upadhyay, A. Artificial intelligence based approach to forecast PM_{2.5} during haze episodes: A case study of Delhi, India. *Atmos. Environ.* **2015**, *102*, 239–248. [\[CrossRef\]](#)
114. Kurt, A.; Gulbagci, B.; Karaca, F.; Alagha, O. An online air pollution forecasting system using neural networks. *Environ. Int.* **2008**, *34*, 592–598. [\[CrossRef\]](#)
115. Mihalache, S.F.; Popescu, M.; Oprea, M. Particulate matter prediction using ANFIS modelling techniques. In *Proceedings of the 2015 19th International Conference on System Theory, Control and Computing, ICSTCC 2015—Joint Conference SINTES 19, SACCs 15, SIMSIS 19*; Institute of Electrical and Electronics Engineers Inc.: Piscataway, NJ, USA, 2015; pp. 895–900.
116. Prasad, K.; Gorai, A.K.; Goyal, P. Development of ANFIS models for air quality forecasting and input optimization for reducing the computational cost and time. *Atmos. Environ.* **2016**, *128*, 246–262. [\[CrossRef\]](#)
117. Kurt, A.; Oktay, A.B. Forecasting air pollutant indicator levels with geographic models 3 days in advance using neural networks. *Expert Syst. Appl.* **2010**, *37*, 7986–7992. [\[CrossRef\]](#)
118. Wahid, H.; Ha, Q.P.; Duc, H.; Azzi, M. Neural network-based meta-modelling approach for estimating spatial distribution of air pollutant levels. *Appl. Soft Comput.* **2013**, *13*, 4087–4096. [\[CrossRef\]](#)
119. Yeganeh, B.; Hewson, M.G.; Clifford, S.; Knibbs, L.D.; Morawska, L. A satellite-based model for estimating PM_{2.5} concentration in a sparsely populated environment using soft computing techniques. *Environ. Model. Soft.* **2017**, *88*, 84–92. [\[CrossRef\]](#)
120. De Gennaro, G.; Trizio, L.; Di Gilio, A.; Pey, J.; Pérez, N.; Cusack, M.; Alastuey, A.; Querol, X. Neural network model for the prediction of PM₁₀ daily concentrations in two sites in the Western Mediterranean. *Sci. Total Environ.* **2013**, *463–464*, 875–883. [\[CrossRef\]](#)
121. Yeganeh, B.; Hewson, M.G.; Clifford, S.; Tavassoli, A.; Knibbs, L.D.; Morawska, L. Estimating the spatiotemporal variation of NO₂ concentration using an adaptive neuro-fuzzy inference system. *Environ. Model. Soft.* **2018**, *100*, 222–235. [\[CrossRef\]](#)

122. Chung, C.-J.; Hsieh, Y.-Y.; Lin, H.-C. Fuzzy inference system for modeling the environmental risk map of air pollutants in Taiwan. *J. Environ. Manag.* **2019**, *246*, 808–820. [[CrossRef](#)]
123. Özdemir, U.; Taner, S. Impacts of Meteorological Factors on PM₁₀: Artificial Neural Networks (ANN) and Multiple Linear Regression (MLR) Approaches. *Environ. Forensics* **2014**, *15*, 329–336. [[CrossRef](#)]
124. Bhardwaj, R.; Pruthi, D. Evolutionary Techniques for Optimizing Air Quality Model. *Proc. Comput. Sci.* **2020**, *167*, 1872–1879. [[CrossRef](#)]
125. Moustris, K.P.; Proias, G.T.; Larissi, I.K.; Nastos, P.T.; Koukouletsos, K.V.; Paliatsos, A.G. Prognosis of maximum daily surface ozone concentration within the greater Athens urban area, Greece. *Glob. NEST J.* **2014**, *16*, 873–882. [[CrossRef](#)]
126. Zeinalnezhad, M.; Chofreh, A.G.; Goni, F.A.; Klemeš, J.J. Air pollution prediction using semi-experimental regression model and Adaptive Neuro-Fuzzy Inference System. *J. Clean. Prod.* **2020**, *261*, 121218. [[CrossRef](#)]
127. Faris, H.; Alkasassbeh, M.; Rodan, A. Artificial Neural Networks for Surface Ozone Prediction: Models and Analysis. *Polish J. Environ. Stud.* **2014**, *23*, 341–348.
128. Ha, Q.P.; Wahid, H.; Duc, H.; Azzi, M. Enhanced radial basis function neural networks for ozone level estimation. *Neurocomputing* **2015**, *155*, 62–70. [[CrossRef](#)]
129. Solazzo, E.; Bianconi, R.; Vautard, R.; Appel, K.W.; Moran, M.D.; Hogrefe, C.; Bessagnet, B.; Brandt, J.; Christensen, J.H.; Chemel, C.; et al. Model evaluation and ensemble modelling of surface-level ozone in Europe and North America in the context of AQMEII. *Atmos. Environ.* **2012**, *53*, 60–74. [[CrossRef](#)]
130. Russo, A.; Lind, P.G.; Raischel, F.; Trigo, R.; Mendes, M. Neural network forecast of daily pollution concentration using optimal meteorological data at synoptic and local scales. *Atmos. Pollut. Res.* **2015**, *6*, 540–549. [[CrossRef](#)]
131. Bing, G.; Meré, J.O.; Cabrera, C.B. Prediction models for ozone in metropolitan area of Mexico City based on artificial intelligence techniques. *Int. J. Inf. Dec. Sci.* **2015**, *7*, 115. [[CrossRef](#)]
132. Li, X.; Peng, L.; Hu, Y.; Shao, J.; Chi, T. Deep learning architecture for air quality predictions. *Environ. Sci. Pollut. Res.* **2016**, *23*, 22408–22417. [[CrossRef](#)]
133. Niu, M.; Wang, Y.; Sun, S.; Li, Y. A novel hybrid decomposition-and-ensemble model based on CEEMD and GWO for short-term PM_{2.5} concentration forecasting. *Atmos. Environ.* **2016**, *134*, 168–180. [[CrossRef](#)]
134. 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](#)]
135. Niu, M.; Gan, K.; Sun, S.; Li, F. Application of decomposition-ensemble learning paradigm with phase space reconstruction for day-ahead PM_{2.5} concentration forecasting. *J. Environ. Manag.* **2017**, *196*, 110–118. [[CrossRef](#)]
136. Perez, P.; Gramsch, E. Forecasting hourly PM_{2.5} in Santiago de Chile with emphasis on night episodes. *Atmos. Environ.* **2016**, *124*, 22–27. [[CrossRef](#)]
137. Li, L.; Zhang, J.; Qiu, W.; Wang, J.; Fang, Y. An ensemble spatiotemporal model for predicting PM_{2.5} concentrations. *Int. J. Environ. Res. Public Health* **2017**, *14*, 549. [[CrossRef](#)]
138. Hur, S.K.; Oh, H.R.; Ho, C.H.; Kim, J.; Song, C.K.; Chang, L.S.; Lee, J.B. Evaluating the predictability of PM₁₀ grades in Seoul, Korea using a neural network model based on synoptic patterns. *Environ. Pollut.* **2016**, *218*, 1324–1333. [[CrossRef](#)] [[PubMed](#)]
139. Wang, J.; Song, G. A Deep Spatial-Temporal Ensemble Model for Air Quality Prediction. *Neurocomputing* **2018**, *314*, 198–206. [[CrossRef](#)]
140. Peng, H.; Lima, A.R.; Teakles, A.; Jin, J.; Cannon, A.J.; Hsieh, W.W. Evaluating hourly air quality forecasting in Canada with nonlinear updatable machine learning methods. *Air Qual. Atmos. Heal.* **2017**, *10*, 195–211. [[CrossRef](#)]
141. Rahimi, A. Short-term prediction of NO₂ and NO_x concentrations using multilayer perceptron neural network: A case study of Tabriz, Iran. *Ecol. Process.* **2017**, *6*. [[CrossRef](#)]
142. Zhai, B.; Chen, J. Development of a stacked ensemble model for forecasting and analyzing daily average PM_{2.5} concentrations in Beijing, China. *Sci. Total Environ.* **2018**, *635*, 644–658. [[CrossRef](#)]
143. Shtein, A.; Kloog, I.; Schwartz, J.; Silibello, C.; Michelozzi, P.; Gariazzo, C.; Viegi, G.; Forastiere, F.; Karnieli, A.; Just, A.C.; et al. Estimating Daily PM_{2.5} and PM₁₀ over Italy Using an Ensemble Model. *Environ. Sci. Technol.* **2019**. [[CrossRef](#)]

144. Gao, M.; Yin, L.; Ning, J. Artificial neural network model for ozone concentration estimation and Monte Carlo analysis. *Atmos. Environ.* **2018**, *184*, 129–139. [\[CrossRef\]](#)
145. Brasseur, G.P.; Xie, Y.; Petersen, A.K.; Bouarar, I.; Flemming, J.; Gauss, M.; Jiang, F.; Kouznetsov, R.; Kranenburg, R.; Mijling, B.; et al. Ensemble forecasts of air quality in eastern China—Part 1: Model description and implementation of the MarcoPolo–Panda prediction system, version 1. *Geosci. Model Dev.* **2019**, *12*, 33–67. [\[CrossRef\]](#)
146. Di, Q.; Amini, H.; Shi, L.; Kloog, I.; Silvern, R.; Kelly, J.; Sabath, M.B.; Choirat, C.; Koutrakis, P.; Lyapustin, A.; et al. An ensemble-based model of PM_{2.5} concentration across the contiguous United States with high spatiotemporal resolution. *Environ. Int.* **2019**, *130*, 104909. [\[CrossRef\]](#)
147. Palomares-Salas, J.C.; González-de-la-Rosa, J.J.; Agüera-Pérez, A.; Sierra-Fernández, J.M.; Florencias-Oliveros, O. Forecasting PM₁₀ in the Bay of Algeciras Based on Regression Models. *Sustainability* **2019**, *11*, 968. [\[CrossRef\]](#)
148. Maciag, P.S.; Kasabov, N.; Kryszkiewicz, M.; Bembenik, R. Air pollution prediction with clustering-based ensemble of evolving spiking neural networks and a case study for London area. *Environ. Model. Softw.* **2019**, *118*, 262–280. [\[CrossRef\]](#)
149. Schornobay-Lui, E.; Alexandrina, E.C.; Aguiar, M.L.; Hanisch, W.S.; Corrêa, E.M.; Corrêa, N.A. Prediction of short and medium term PM₁₀ concentration using artificial neural networks. *Manag. Environ. Qual. An Int. J.* **2019**, *30*, 414–436. [\[CrossRef\]](#)
150. Masih, A. Application of ensemble learning techniques to model the atmospheric concentration of SO₂. *Glob. J. Environ. Sci. Manag.* **2019**, *5*, 309–318. [\[CrossRef\]](#)
151. Ventura, L.M.B.; de Oliveira Pinto, F.; Soares, L.M.; Luna, A.S.; Gioda, A. Forecast of daily PM_{2.5} concentrations applying artificial neural networks and Holt–Winters models. *Air Qual. Atmos. Heal.* **2019**, *12*, 317–325. [\[CrossRef\]](#)
152. Masih, A. Application of Random Forest Algorithm to Predict the Atmospheric Concentration of NO₂. In *Proceedings of the 2019 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology, USBREIT 2019*; Institute of Electrical and Electronics Engineers Inc.: Piscataway, NJ, USA, 2019; pp. 252–255.
153. Mohebbi, M.R.; Karimi Jashni, A.; Dehghani, M.; Hadad, K. Short-Term Prediction of Carbon Monoxide Concentration Using Artificial Neural Network (NARX) Without Traffic Data: Case Study: Shiraz City. *Iran. J. Sci. Technol. Trans. Civ. Eng.* **2019**, *43*, 533–540. [\[CrossRef\]](#)
154. Mohan, S.; Saranya, P. A novel bagging ensemble approach for predicting summertime ground-level ozone concentration. *J. Air Waste Manag. Assoc.* **2019**, *69*, 220–233. [\[CrossRef\]](#)
155. Abdul Aziz, F.A.B.; Rahman, N.; Mohd Ali, J. Tropospheric ozone formation estimation in Urban City, Bangi, Using Artificial Neural Network (ANN). *Comput. Intell. Neurosci.* **2019**, *2019*. [\[CrossRef\]](#) [\[PubMed\]](#)
156. Liu, H.; Xu, Y.; Chen, C. Improved pollution forecasting hybrid algorithms based on the ensemble method. *Appl. Math. Model.* **2019**, *73*, 473–486. [\[CrossRef\]](#)
157. Resmi, C.T.; Nishanth, T.; Satheesh Kumar, M.K.; Balachandramohan, M.; Valsaraj, K.T. Long-term variations of air quality influenced by surface ozone in a coastal site in India: Association with synoptic meteorological conditions with model simulations. *Atmosphere* **2020**, *11*, 193.
158. Feng, X.; Fu, T.-M.; Cao, H.; Tian, H.; Fan, Q.; Chen, X. Neural network predictions of pollutant emissions from open burning of crop residues: Application to air quality forecasts in southern China. *Atmos. Environ.* **2019**, *204*, 22–31. [\[CrossRef\]](#)
159. Savi, F.; Nemitz, E.; Coyle, M.; Aitkenhead, M.; Frumau, K.; Gerosa, G.; Finco, A.; Gruening, C.; Goded, I.; Loubet, B.; et al. Neural Network Analysis to Evaluate Ozone Damage to Vegetation Under Different Climatic Conditions. *Front. For. Glob. Chang.* **2020**, *3*, 42. [\[CrossRef\]](#)
160. Shishegaran, A.; Saeedi, M.; Kumar, A.; Ghiasinejad, H. Prediction of air quality in Tehran by developing the nonlinear ensemble model. *J. Clean. Prod.* **2020**, *259*, 120825. [\[CrossRef\]](#)
161. Kow, P.Y.; Wang, Y.S.; Zhou, Y.; Kao, I.F.; Issermann, M.; Chang, L.C.; Chang, F.J. Seamless integration of convolutional and back-propagation neural networks for regional multi-step-ahead PM_{2.5} forecasting. *J. Clean. Prod.* **2020**, *261*, 121285. [\[CrossRef\]](#)
162. Sun, W.; Li, Z. Hourly PM_{2.5} concentration forecasting based on feature extraction and stacking-driven ensemble model for the winter of the Beijing-Tianjin-Hebei area. *Atmos. Pollut. Res.* **2020**. [\[CrossRef\]](#)
163. Photphanloet, C.; Lipikorn, R. PM₁₀ concentration forecast using modified depth-first search and supervised learning neural network. *Sci. Total Environ.* **2020**, *727*, 138507. [\[CrossRef\]](#)

164. Al-Alawi, S.M.; Abdul-Wahab, S.A.; Bakheit, C.S. Combining principal component regression and artificial neural networks for more accurate predictions of ground-level ozone. *Environ. Model. Soft.* **2008**, *23*, 396–403. [\[CrossRef\]](#)
165. Athira, V.; Geetha, P.; Vinayakumar, R.; Soman, K.P. DeepAirNet: Applying Recurrent Networks for Air Quality Prediction. *Proc. Comput. Sci.* **2018**, *132*, 1394–1403. [\[CrossRef\]](#)
166. Freeman, B.S.; Taylor, G.; Gharabaghi, B.; Thé, J. Forecasting air quality time series using deep learning. *J. Air Waste Manag. Assoc.* **2018**, *68*, 866–886. [\[CrossRef\]](#)
167. Tao, Q.; Liu, F.; Li, Y.; Sidorov, D. Air Pollution Forecasting Using a Deep Learning Model Based on 1D Convnets and Bidirectional GRU. *IEEE Access* **2019**, *7*, 76690–76698. [\[CrossRef\]](#)
168. Ma, J.; Cheng, J.C.P.; Lin, C.; Tan, Y.; Zhang, J. Improving air quality prediction accuracy at larger temporal resolutions using deep learning and transfer learning techniques. *Atmos. Environ.* **2019**, 116885. [\[CrossRef\]](#)
169. Zarandi, M.H.F.; Faraji, M.R.; Karbasian, M. Interval type-2 fuzzy expert system for prediction of carbon monoxide concentration in mega-cities. *Appl. Soft Comput.* **2012**, *12*, 291–301. [\[CrossRef\]](#)
170. Zamani Joharestani, M.; Cao, C.; Ni, X.; Bashir, B.; Talebiesfandarani, S.; Zamani Joharestani, M.; Cao, C.; Ni, X.; Bashir, B.; Talebiesfandarani, S. PM_{2.5} Prediction Based on Random Forest, XGBoost, and Deep Learning Using Multisource Remote Sensing Data. *Atmosphere* **2019**, *10*, 373. [\[CrossRef\]](#)
171. Kumar, A.; Goyal, P. Forecasting of Air Quality Index in Delhi Using Neural Network Based on Principal Component Analysis. *Pure Appl. Geophys.* **2013**, *170*, 711–722. [\[CrossRef\]](#)
172. Zhang, W.Y.; Wang, J.J.; Liu, X.; Wang, J.Z. Prediction of Ozone Concentration in Semi-Arid Areas of China Using a Novel Hybrid Model. *J. Environ. Inform.* **2013**, *22*, 68–77. [\[CrossRef\]](#)
173. Taşpınar, F. Improving artificial neural network model predictions of daily average PM₁₀ concentrations by applying principle component analysis and implementing seasonal models. *J. Air Waste Manag. Assoc.* **2015**, *65*, 800–809. [\[CrossRef\]](#)
174. Lu, W.; Wang, W.; Leung, A.Y.T.; Lo, S.M.; Yuen, R.K.K.; Xu, Z.; Fan, H. Air pollutant parameter forecasting using support vector machines. In Proceedings of the International Joint Conference on Neural Networks, Honolulu, HI, USA, 12–17 May 2002; Volume 1, pp. 630–635.
175. Wang, W.; Xu, Z.; Weizhen Lu, J. Three improved neural network models for air quality forecasting. *Eng. Comput.* **2003**, *20*, 192–210. [\[CrossRef\]](#)
176. Feng, X.; Li, Q.; Zhu, Y.; Hou, J.; Jin, L.; Wang, J. Artificial neural networks forecasting of PM_{2.5} pollution using air mass trajectory based geographic model and wavelet transformation. *Atmos. Environ.* **2015**, *107*, 118–128. [\[CrossRef\]](#)
177. Lu, W.-Z.; Wang, W.-J. Potential assessment of the “support vector machine” method in forecasting ambient air pollutant trends. *Chemosphere* **2005**, *59*, 693–701. [\[CrossRef\]](#)
178. Sharma, S.; Kalra, U.; Srivathsan, S.; Rana, K.P.S.; Kumar, V. Efficient air pollutants prediction using ANFIS trained by Modified PSO algorithm. In *Proceedings of the 4th International Conference on Reliability, Infocom Technologies and Optimization: Trends and Future Directions, ICRITO 2015*; Institute of Electrical and Electronics Engineers Inc.: Piscataway, NJ, USA, 2015.
179. Wang, P.; Liu, Y.; Qin, Z.; Zhang, G. A novel hybrid forecasting model for PM₁₀ and SO₂ daily concentrations. *Sci. Total Environ.* **2015**, *505*, 1202–1212. [\[CrossRef\]](#)
180. Mehdipour, V.; Memarianfard, M. Ground-level O₃ sensitivity analysis using support vector machine with radial basis function. *Int. J. Environ. Sci. Technol.* **2019**, *16*, 2745–2754. [\[CrossRef\]](#)
181. Mishra, D.; Goyal, P. Development of artificial intelligence based NO₂ forecasting models at Taj Mahal, Agra. *Atmos. Pollut. Res.* **2015**, *6*, 99–106. [\[CrossRef\]](#)
182. Faris, H.; Ghatasheh, N.; Rodan, A.; Abu-Faraj, M.M.; Cornuelle, B.D. Predicting Surface Ozone Concentrations using Support Vector Regression. *Life Sci. J.* **2014**, *11*, 126–131.
183. Bai, Y.; Li, Y.; Wang, X.; Xie, J.; Li, C. Air pollutants concentrations forecasting using back propagation neural network based on wavelet decomposition with meteorological conditions. *Atmos. Pollut. Res.* **2016**, *7*, 557–566. [\[CrossRef\]](#)
184. Suleiman, A.; Tight, M.R.; Quinn, A.D. Hybrid Neural Networks and Boosted Regression Tree Models for Predicting Roadside Particulate Matter. *Environ. Model. Assess.* **2016**, *21*, 731–750. [\[CrossRef\]](#)
185. Liu, B.-C.; Binaykia, A.; Chang, P.-C.; Tiwari, M.K.; Tsao, C.-C. Urban air quality forecasting based on multi-dimensional collaborative Support Vector Regression (SVR): A case study of Beijing-Tianjin-Shijiazhuang. *PLoS ONE* **2017**, *12*, e0179763. [\[CrossRef\]](#)

186. Yang, W.; Deng, M.; Xu, F.; Wang, H. Prediction of hourly PM_{2.5} using a space-time support vector regression model. *Atmos. Environ.* **2018**, *181*, 12–19. [\[CrossRef\]](#)
187. Li, W.; Kong, D.; Wu, J. A New Hybrid Model FPA-SVM Considering Cointegration for Particulate Matter Concentration Forecasting: A Case Study of Kunming and Yuxi, China. *Comput. Intell. Neurosci.* **2017**, 2017. [\[CrossRef\]](#)
188. Liu, H.; Li, Q.; Yu, D.; Gu, Y. Air Quality Index and Air Pollutant Concentration Prediction Based on Machine Learning Algorithms. *Appl. Sci.* **2019**, *9*, 4069. [\[CrossRef\]](#)
189. Sun, W.; Sun, J. Daily PM_{2.5} concentration prediction based on principal component analysis and LSSVM optimized by cuckoo search algorithm. *J. Environ. Manag.* **2017**, *188*, 144–152. [\[CrossRef\]](#) [\[PubMed\]](#)
190. Qiao, J.; Cai, J.; Han, H.; Cai, J. Predicting PM_{2.5} concentrations at a regional background station using second order self-organizing fuzzy neural network. *Atmosphere* **2017**, *8*, 10. [\[CrossRef\]](#)
191. Debnath, J.; Majumder, D.; Biswas, A. Air quality assessment using weighted interval type-2 fuzzy inference system. *Ecol. Inform.* **2018**, *46*, 133–146. [\[CrossRef\]](#)
192. Lu, W.Z.; Fan, H.Y.; Lo, S.M. Application of evolutionary neural network method in predicting pollutant levels in downtown area of Hong Kong. *Neurocomputing* **2003**, *51*, 387–400. [\[CrossRef\]](#)
193. Li, Y.; Jiang, P.; She, Q.; Lin, G. Research on air pollutant concentration prediction method based on self-adaptive neuro-fuzzy weighted extreme learning machine. *Environ. Pollut.* **2018**, *241*, 1115–1127. [\[CrossRef\]](#) [\[PubMed\]](#)
194. Kapageridis, I.K.; Triantafyllou, A.G. A Genetically Optimised Neural Network For Prediction Of Maximum Hourly PM₁₀ Concentration—Semantic Scholar. In *Proceedings of the 12th International Conference on Modeling, Monitoring and Management of Air Pollution*; Wessex Institute of Technology: Rhodes, Greece, 2004; pp. 161–170.
195. Eldakhly, N.M.; Aboul-Ela, M.; Abdalla, A. A Novel Approach of Weighted Support Vector Machine with Applied Chance Theory for Forecasting Air Pollution Phenomenon in Egypt. *Int. J. Comput. Intell. Appl.* **2018**, *17*. [\[CrossRef\]](#)
196. Niska, H.; Hiltunen, T.; Karppinen, A.; Ruuskanen, J.; Kolehmainen, M. Evolving the neural network model for forecasting air pollution time series. *Eng. Appl. Artif. Intell.* **2004**, *17*, 159–167. [\[CrossRef\]](#)
197. Xu, S.; Zou, B.; Shafi, S.; Sternberg, T. A hybrid Grey-Markov/ LUR model for PM₁₀ concentration prediction under future urban scenarios. *Atmos. Environ.* **2018**, *187*, 401–409. [\[CrossRef\]](#)
198. Zhou, Y.; Chang, F.-J.; Chang, L.-C.; Kao, I.-F.; Wang, Y.-S.; Kang, C.-C. Multi-output support vector machine for regional multi-step-ahead PM_{2.5} forecasting. *Sci. Total Environ.* **2019**, *651*, 230–240. [\[CrossRef\]](#)
199. Antanasijević, D.Z.; Pocajt, V.V.; Povrenović, D.S.; Ristić, M.D.; Perić-Grujić, A.A. PM₁₀ emission forecasting using artificial neural networks and genetic algorithm input variable optimization. *Sci. Total Environ.* **2013**, *443*, 511–519. [\[CrossRef\]](#)
200. Yadav, V.; Nath, S. Novel hybrid model for daily prediction of PM₁₀ using principal component analysis and artificial neural network. *Int. J. Environ. Sci. Technol.* **2019**, *16*, 2839–2848. [\[CrossRef\]](#)
201. Elangasinghe, M.A.; Singhal, N.; Dirks, K.N.; Salmond, J.A. Development of an ANN-based air pollution forecasting system with explicit knowledge through sensitivity analysis. *Atmos. Pollut. Res.* **2014**, *5*, 696–708. [\[CrossRef\]](#)
202. De Mattos Neto, P.S.G.; Cavalcanti, G.D.C.; Madeiro, F.; Ferreira, T.A.E. An Approach to Improve the Performance of PM Forecasters. *PLoS ONE* **2015**, *10*, e0138507. [\[CrossRef\]](#) [\[PubMed\]](#)
203. Asghari Esfandani, M.; Nematzadeh, H. Predicting air pollution in Tehran: Genetic algorithm and back propagation neural network. *J. AI Data Min.* **2016**, *4*, 49–54.
204. Qi, Y.; Li, Q.; Karimian, H.; Liu, D. A hybrid model for spatiotemporal forecasting of PM_{2.5} based on graph convolutional neural network and long short-term memory. *Sci. Total Environ.* **2019**, *664*, 1–10. [\[CrossRef\]](#)
205. Saxena, A.; Shekhawat, S. Ambient Air Quality Classification by Grey Wolf Optimizer Based Support Vector Machine. *J. Environ. Public Health* **2017**, 2017. [\[CrossRef\]](#)
206. Wang, Z.; Long, Z. PM_{2.5} Prediction Based on Neural Network. In *Proceedings of the 11th International Conference on Intelligent Computation Technology and Automation, ICICTA 2018*; Institute of Electrical and Electronics Engineers Inc.: Piscataway, NJ, USA, 2018; pp. 44–47.
207. Chen, S.; Wang, J.Q.; Zhang, H.Y. A hybrid PSO-SVM model based on clustering algorithm for short-term atmospheric pollutant concentration forecasting. *Technol. Forecast. Soc. Change* **2019**, *146*, 41–54. [\[CrossRef\]](#)
208. Liu, H.; Jin, K.; Duan, Z. Air PM_{2.5} concentration multi-step forecasting using a new hybrid modeling method: Comparing cases for four cities in China. *Atmos. Pollut. Res.* **2019**. [\[CrossRef\]](#)

209. Bui, X.-N.; Lee, C.W.; Nguyen, H.; Bui, H.-B.; Long, N.Q.; Le, Q.-T.; Nguyen, V.-D.; Nguyen, N.-B.; Moayed, H.; Bui, X.-N.; et al. Estimating PM₁₀ Concentration from Drilling Operations in Open-Pit Mines Using an Assembly of SVR and PSO. *Appl. Sci.* **2019**, *9*, 2806. [\[CrossRef\]](#)
210. Zhu, S.; Qiu, X.; Yin, Y.; Fang, M.; Liu, X.; Zhao, X.; Shi, Y. Two-step-hybrid model based on data preprocessing and intelligent optimization algorithms (CS and GWO) for NO₂ and SO₂ forecasting. *Atmos. Pollut. Res.* **2019**, *10*, 1326–1335. [\[CrossRef\]](#)
211. Li, X.; Luo, A.; Li, J.; Li, Y. Air Pollutant Concentration Forecast Based on Support Vector Regression and Quantum-Behaved Particle Swarm Optimization. *Environ. Model. Assess.* **2019**, *24*, 205–222. [\[CrossRef\]](#)
212. Wang, J.; Du, P.; Hao, Y.; Ma, X.; Niu, T.; Yang, W. An innovative hybrid model based on outlier detection and correction algorithm and heuristic intelligent optimization algorithm for daily air quality index forecasting. *J. Environ. Manag.* **2020**, *255*, 109855. [\[CrossRef\]](#) [\[PubMed\]](#)
213. Bai, L.; Wang, J.; Ma, X.; Lu, H.; Bai, L.; Wang, J.; Ma, X.; Lu, H. Air Pollution Forecasts: An Overview. *Int. J. Environ. Res. Public Health* **2018**, *15*, 780. [\[CrossRef\]](#) [\[PubMed\]](#)
214. Haykin, S. *Neural Networks: A Comprehensive Foundation*; Prentice Hall: Upper Saddle River, NJ, USA, 1994; ISBN 0132733501.
215. Park, D.; Rilett, L.R.; Han, G. Spectral Basis Neural Networks for Real-Time Travel Time Forecasting. *J. Transp. Eng.* **1999**, *125*, 515–523. [\[CrossRef\]](#)
216. Afandizadeh, S.; Kianfar, J. A Hybrid Neuro-Genetic Approach to Short-Term Traffic Volume Prediction. *Int. J. Civ. Eng.* **2009**, *7*, 41–48.
217. Boser, B.E.; Guyon, I.M.; Vapnik, V.N. A training algorithm for optimal margin classifiers. In *5th Annual Workshop on Computational Learning Theory—COLT '92, Proc. Conf.*; Association for Computing Machinery: New York, NY, USA, 1992; pp. 144–152.
218. Vapnik, V. *The Nature of Statistical Learning Theory*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2013; ISBN 1475732643.
219. Vapnik, V.; Golowich, S.E.; Smola, A.J. Support Vector Method for Function Approximation, Regression Estimation, and Signal Processing. *Adv. Neural Inf. Process. Syst.* **1997**, *9*, 281–287.
220. Burges, C.J.C. A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining Know. Discov.* **1998**, *2*, 121–167. [\[CrossRef\]](#)
221. Shin, Y.; Kim, Z.; Yu, J.; Kim, G.; Hwang, S. Development of NO_x reduction system utilizing artificial neural network (ANN) and genetic algorithm (GA). *J. Clean. Prod.* **2019**, *232*, 1418–1429. [\[CrossRef\]](#)
222. Masiur Rahman, S.; Khondaker, A.N.; Imtiaz Hossain, M.; Shafiullah, M.; Hasan, M.A. Neurogenetic modeling of energy demand in the United Arab Emirates, Saudi Arabia, and Qatar. *Environ. Prog. Sustain. Energy* **2017**, *36*. [\[CrossRef\]](#)
223. Ijaz, M.; Shafiullah, M.; Abido, M.A. Classification of power quality disturbances using Wavelet Transform and Optimized ANN. In *Proceedings of the 18th International Conference on Intelligent System Application to Power Systems (ISAP)*, Porto, Portugal, 11–16 September 2015; pp. 1–6.
224. Vonk, E.; Jain, L.C.; Johnson, R.P.; Ray, P. *Automatic Generation of Neural Network Architecture Using Evolutionary Computation*; World Scientific Inc.: River Edge, NJ, USA, 1997; ISBN 9810231067.
225. Shafiullah, M.; Abido, M.; Abdel-Fattah, T. Distribution Grids Fault Location employing ST based Optimized Machine Learning Approach. *Energies* **2018**, *11*, 2328. [\[CrossRef\]](#)
226. Karray, F.O.; De Silva, C.W. *Soft Computing and Intelligent Systems Design: Theory, Tools, and Applications*; Pearson/Addison Wesley: Boston, MA, USA, 2004; ISBN 0321116178.
227. Von Altrock, C. *Fuzzy Logic and Neuro Fuzzy Applications Explained*; Prentice Hall PTR: Upper Saddle River, NJ, USA, 1995; ISBN 0133684652.
228. Tanyildizi, H. Fuzzy logic model for prediction of mechanical properties of lightweight concrete exposed to high temperature. *Mater. Des.* **2009**, *30*, 2205–2210. [\[CrossRef\]](#)
229. Godil, S.; Shamim, M.; Enam, S.; Qidwai, U. Fuzzy logic: A “simple” solution for complexities in neurosciences. *Surg. Neurol. Int.* **2011**, *2*. [\[CrossRef\]](#)
230. Allibhai, E. *Building A Deep Learning Model Using Keras*; Towards Data Science Inc.: Toronto, ON, Canada, 2018.
231. Idri, A.; Hosni, M.; Abran, A. Systematic Mapping Study of Ensemble Effort Estimation. In *Proceedings of the 11th International Conference on Evaluation of Novel Software Approaches to Software Engineering*, Rome, Italy, 27–28 April 2016; pp. 132–139.

232. Assi, K.J.; Shafiullah, M.; Md Nahiduzzaman, K.; Mansoor, U. Travel-To-School Mode Choice Modelling Employing Artificial Intelligence Techniques: A Comparative Study. *Sustainability* **2019**, *11*, 4484. [\[CrossRef\]](#)
233. Werbin-Ofir, H.; Dery, L.; Shmueli, E. Beyond majority: Label ranking ensembles based on voting rules. *Expert Syst. Appl.* **2019**, *136*, 50–61. [\[CrossRef\]](#)
234. Hosni, M.; Abnane, I.; Idri, A.; Carrillo de Gea, J.M.; Fernández Alemán, J.L. Reviewing ensemble classification methods in breast cancer. *Comput. Methods Programs Biomed.* **2019**, *177*, 89–112. [\[CrossRef\]](#)
235. Demir, N. Ensemble Methods: Elegant Techniques to Produce Improved Machine Learning Results. Available online: <https://www.toptal.com/machine-learning/ensemble-methods-machine-learning> (accessed on 2 August 2019).
236. Hornik, K. Approximation capabilities of multilayer feedforward networks. *Neural Networks* **1991**, *4*, 251–257. [\[CrossRef\]](#)
237. Leshno, M.; Lin, V.Y.; Pinkus, A.; Schocken, S. Multilayer feedforward networks with a nonpolynomial activation function can approximate any function. *Neural Networks* **1993**, *6*, 861–867. [\[CrossRef\]](#)
238. Hammer, B.; Gersmann, K. A note on the universal approximation capability of support vector machines. *Neural Process. Lett.* **2003**, *17*, 43–53. [\[CrossRef\]](#)
239. Ying, H. General Takagi-Sugeno fuzzy systems are universal approximators. In *Proceedings of the 1998 IEEE International Conference on Fuzzy Systems Proceedings—IEEE World Congress on Computational Intelligence*; Institute of Electrical and Electronics Engineers Inc.: Piscataway, NJ, USA, 1998; Volume 1, pp. 819–823.
240. Morabito, F.C.; Marino, D.; Ricca, B. Management of uncertainty in environmental problems: An assesment of technical aspects and policies. In *Handbook of Management under Uncertainty*; Springer: Berlin/Heidelberg, Germany, 2001; pp. 747–799.
241. Herrera, F.; Lozano, M.; Verdegay, J.L. Dynamic and heuristic fuzzy connectives-based crossover operators for controlling the diversity and convergence of real-coded genetic algorithms. *Int. J. Intell. Syst.* **1998**, *11*, 1013–1040. [\[CrossRef\]](#)
242. Cordon, O.; Gomide, F.; Herrera, F.; Hoffmann, F.; Magdalena, L. Ten years of genetic fuzzy systems: Current framework and new trends. *Fuzzy Sets Syst.* **2004**, *141*, 5–31. [\[CrossRef\]](#)
243. Hassan, M.R.; Arafat, S.M.; Begg, R.K. Fuzzy-Genetic Model for the Identification of Falls Risk Gait. In *Proceedings of the Procedia Computer Science*; Elsevier B.V.: Amsterdam, The Netherlands, 2016; Volume 82, pp. 4–11.
244. Chatterjee, A.; Chatterjee, R.; Matsuno, F.; Endo, T. Neuro-fuzzy state modeling of flexible robotic arm employing dynamically varying cognitive and social component based PSO. *Meas. J. Int. Meas. Confed.* **2007**, *40*, 628–643. [\[CrossRef\]](#)
245. Dolgoplov, P.; Konstantinov, D.; Rybalchenko, L.; Muhitovs, R. Optimization of train routes based on neuro-fuzzy modeling and genetic algorithms. In *Proceedings of the Procedia Computer Science*; Elsevier B.V.: Amsterdam, The Netherlands, 2019; Volume 149, pp. 11–18.
246. Ashish, K.; Dasari, A.; Chattopadhyay, S.; Hui, N.B. Genetic-neuro-fuzzy system for grading depression. *Appl. Comput. Informatics* **2018**, *14*, 98–105. [\[CrossRef\]](#)
247. Douiri, M.R. Particle swarm optimized neuro-fuzzy system for photovoltaic power forecasting model. *Sol. Energy* **2019**, 91–104. [\[CrossRef\]](#)
248. Karnik, N.N.; Mendel, J.M. Applications of type-2 fuzzy logic systems: Handling the uncertainty associated with surveys. In *Proceedings of the FUZZ-IEEE'99. 1999 IEEE International Fuzzy Systems. Conference Proceedings (Cat. No.99CH36315)*; IEEE: Piscataway, NJ, USA, 1999; Volume 3, pp. 1546–1551.
249. Shafaei Bajestani, N.; Vahidian Kamyad, A.; Nasli Esfahani, E.; Zare, A. Prediction of retinopathy in diabetic patients using type-2 fuzzy regression model. *Eur. J. Oper. Res.* **2018**, *264*, 859–869. [\[CrossRef\]](#)
250. Sharifian, A.; Ghadi, M.J.; Ghavidel, S.; Li, L.; Zhang, J. A new method based on Type-2 fuzzy neural network for accurate wind power forecasting under uncertain data. *Renew. Energy* **2018**, *120*, 220–230. [\[CrossRef\]](#)
251. Barron, A. Predicted squared error: A criterion for automatic model selection. In *Proceedings of the Self-Organizing Methods in Modeling*; Marcel Dekker: New York, NY, USA, 1984; pp. 87–103.
252. Castillo, E. Functional Networks. *Neural Process. Lett.* **1998**, *7*, 151–159. [\[CrossRef\]](#)
253. Zhou, G.; Zhou, Y.; Huang, H.; Tang, Z. Functional networks and applications: A survey. *Neurocomputing* **2019**, *335*, 384–399. [\[CrossRef\]](#)
254. Wu, J.; Wang, Y.; Zhang, X.; Chen, Z. A novel state of health estimation method of Li-ion battery using group method of data handling. *J. Power Sources* **2016**, *327*, 457–464. [\[CrossRef\]](#)

- 255. Liu, H.; Duan, Z.; Wu, H.; Li, Y.; Dong, S. Wind speed forecasting models based on data decomposition, feature selection and group method of data handling network. *Measurement* **2019**, *148*, 106971. [[CrossRef](#)]
- 256. Kolodner, J.L. An introduction to case-based reasoning. *Artif. Intell. Rev.* **1992**, *6*, 3–34. [[CrossRef](#)]
- 257. Aamodt, A.; Plaza, E. Case-based reasoning: Foundational issues, methodological variations, and system approaches. *J. AI Commun.* **1994**, *7*, 39–59. [[CrossRef](#)]
- 258. Meyer, M.D.; Watson, L.S.; Walton, M.; Skinner, R.E. *Artificial Intelligence in Transportation: Information for Application*; National Research Council: Washington, DC, USA, 2007.
- 259. Abutair, H.Y.A.; Belghith, A. Using Case-Based Reasoning for Phishing Detection. In *Proceedings of the Procedia Computer Science*; Elsevier B.V.: Amsterdam, The Netherlands, 2017; Volume 109, pp. 281–288.
- 260. Raza, B.; Kumar, Y.J.; Malik, A.K.; Anjum, A.; Faheem, M. Performance prediction and adaptation for database management system workload using Case-Based Reasoning approach. *Inf. Syst.* **2018**, *76*, 46–58. [[CrossRef](#)]
- 261. Blondet, G.; Le Duigou, J.; Boudaoud, N. A knowledge-based system for numerical design of experiments processes in mechanical engineering. *Expert Syst. Appl.* **2019**, *122*, 289–302. [[CrossRef](#)]
- 262. Sammut, C.; Webb, G.I. *Encyclopedia of Machine Learning and Data Mining*, 2nd ed.; Springer: New York, NY, USA, 2017; ISBN 9781489976857.

© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).