Machine learning algorithms to forecast air quality: a survey

Manuel Méndez[1] · Mercedes G. Merayo[1] · Manuel Núñez[1]

Abstract

Air pollution is a risk factor for many diseases that can lead to death. Therefore, it is important to develop forecasting mechanisms that can be used by the authorities, so that they can anticipate measures when high concentrations of certain pollutants are expected in the near future. Machine Learning models, in particular, Deep Learning models, have been widely used to forecast air quality. In this paper we present a comprehensive review of the main contributions in the field during the period 2011–2021. We have searched the main scientific publications databases and, after a careful selection, we have considered a total of 155 papers. The papers are classified in terms of geographical distribution, predicted values, predictor variables, evaluation metrics and Machine Learning model.

Keywords Machine learning · Deep learning · Regression algorithms · Air quality

1 Introduction

The economical and urban development of cities has brought that the interest in environmental pollution has risen during the last years. Among other problems, air pollution has a big impact in human health, being a risk factor for many diseases that can lead to death. According to the World Health Organisation (WHO), air pollution is a "silent killer" that produces the premature death of almost seven million people each year,[1] including 600.000 children.

---

[1] https://www.scmp.com/news/world/europe/article/3149735/air-pollution-kills-7-million-year-says-who-it-tightens, last accessed on 2022-02-15.

Mercedes G. Merayo and Manuel Núñez have contributed equally to this work.

✉ Manuel Méndez
  manumend@ucm.es

  Mercedes G. Merayo
  mgmerayo@fdi.ucm.es

  Manuel Núñez
  mn@sip.ucm.es

1 Design and Testing of Reliable Systems Research Group, Universidad Complutense de Madrid, C/ Profesor José García Santesmases, 9, 28040 Madrid, Madrid, Spain

Air pollution forecasting is very useful for informing about the pollution level that will allow policy-makers to adopt measures for reducing its impact. For example, traffic restrictions could be imposed with the goal of avoiding high pollution episodes. Usually, the Air Quality Index (AQI) is used to indicate the air pollution level. AQI is a piece-wise linear function of the following pollutant concentrations: Ozone (O_3), particulate matters (PM2.5, PM10), carbon monoxide (CO), sulphur dioxide (SO_2) and nitrogen dioxide (NO_2). Nevertheless, there does not exist a global standard and different countries and regions have their own AQI indexes, based on their own air quality standards.

Machine Learning (ML) techniques are the most common methods to forecast air quality. Since the beginning of the 21st century, we can find hundreds of works in the literature that propose implementations of different models to get the best accuracy in the forecasting of air quality index or pollutant concentration prediction.

There has been much research on applying different ML algorithms to predict both the AQI and the level of concentration of specific pollutants related to air quality. In fact, there are some papers that collect applications of ML to air quality (e.g. Niharika Venkatadri and Rao 2014; Nemade 2019; Iskandaryan et al. 2020; Bharat Deshmukh et al. 2021; Sharma et al. 2021). Unfortunately, their scope is rather limited, the total number of analysed papers is very small and the considered classification categories are very restricted. Therefore, we are not aware of an extensive work that reviews this important field.

The thorough study of these papers provide us the followings insights:

1. The pollution and technological potential in countries are determinant factors to publish papers in this topic.
2. From 2011, with few exceptions, there is an increase on the number of papers per year in the topic. This trend is very clear between 2017 and 2020. However, this trend (surprisingly) changed in 2021. We concluded that this change was due to the emergence of COVID-19 data as main case study of research working on time series forecasting. In order to support this claim, we did a simple experiment: we queried *IEEE Xplore* using *'("Abstract":"air quality") AND ("Abstract":"machine learning" OR "Abstract":"neural network" OR "Abstract":"prediction" OR "Abstract":"regression")'* and obtained a total of 105 results in 2020. Then, we modified the query and the year: we replaced "air quality" by "COVID" and checked 2021. We obtained a total of 1108 results.
3. AQI is the most predicted pollutant measure and PM2.5 is the most predicted pollutant concentration.
4. The most employed predictor variables are pollutant features (50%) followed by weather features (35%) and the combination of both of them gives the best accuracy.
5. Evidence seems to point towards deep learning algorithms being more effective than regression algorithms.
6. Despite the rise in popularity of deep learning, the ratio of deep learning algorithms with respect to regression algorithms has remained almost constant over the years.

The rest of the paper is organised as follows. In Sect. 2, we describe the main ML algorithms that have been applied for prediction of air quality. In Sect. 3 we explain the review methodology that we have applied and how we select publications by taking into account different aspects. Specifically, we consider studied geographical area, predicted values, predictor variables and evaluation metrics. In Sect. 4, which constitutes the bulk of the paper, we describe the papers that we have included in this survey according to the underlying algoritm. Finally, in Sect. 5 we present our conclusions.

2 Machine learning and classical regression algorithms

Machine Learning is a branch of Artificial Intelligence that aims to provide computers with the ability to learn how to perform specific tasks without being explicitly programmed by a human. This technique is based on the design of models that learn from data and make decisions or predictions when new data are available. Deep Learning (DL) can be seen as an evolution of ML that uses a structure of multiple layers called Artificial Neural Network (ANN). DL algorithms require less involvement of humans because features are automatically extracted. However, an important difference with respect to other ML techniques is that DL requires massive data to work properly.

Although ML and DL are recent concepts, the first computer learning program was written by Arthur Samuel in 1952 and the first neural network was proposed by Frank Rossenblatt in 1957.[2] Since the 1990s, the development in both ML and DL has been significant, mainly due to the increment of computation power and the availability of large amounts of data.

There exist many ML approaches that can be applied to solve different problems. In this section, we will review only those algorithms that have been used for predicting pollutant measures. We can distinguish between the ones based on regression analysis and the ones using neural networks. Moreover, in the first category we will distinguish between the use of classical regression algorithms and ML algorithms.

2.1 Classical regression-based algorithms

Regression analysis is used to infer the relation between a dependent variable and a set of independent variables. On the basis of this relation, and using the values of the independent variables, the value of the dependent variable is estimated. Regression helps to predict a continuous value. Next we review classical algorithms to carry out regression.

- **Multiple Linear Regression (MLR).** Let y and $x_1, \ldots, x_p$ be the dependent variable and the independent variables, respectively. The goal of linear regression is to define a linear function $f(x_1, \ldots, x_p)$ that minimise the square mean error, that is,

$$min[(y - f(x_1, \ldots, x_p))^2]$$

  If $p = 1$ the algorithm is called *simple linear regression*.
- **Auto-Regressive Integrated Moving Average (ARIMA).** This model uses time series for forecasting, that is, it makes predictions based on past and present data. The ARIMA model includes three components. The auto-regressive component refers to the number of delays used in the model. For example, an auto-regressive value of 3 establishes that only the three previous values will be used to explain the current value. The integrated component represents the differentiation degree required to convert the time series into a stationary series. The last component, the moving average, refers to the number of past errors required to explain the current error. Again, if the value of the moving average is 3, then only the three previous error values can be used to explain

---

(a) Decision tree structure   (b) Random Forest structure   (c) KNN structure
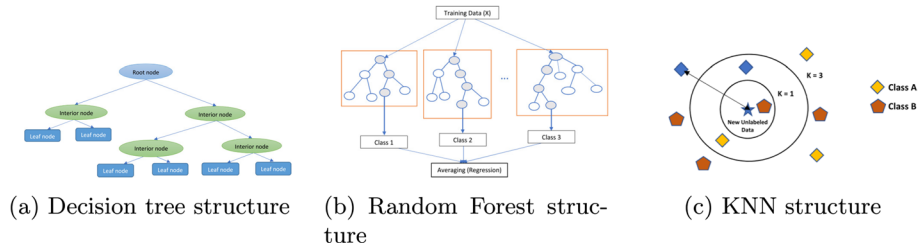
Fig. 1 Graphical representation of ML regression-based algorithms

the current error value. When the time series is seasonally, it uses an ARIMA variant called SARIMA.

2.2 Machine learning regression-based algorithms

In this section we review ML algorithms used for air quality prediction based on regression analysis.

- **Support vector regression (SVR).** Support vector machines are mainly used in classification problems. However, they can be also applied to regression. In this case, the approach is called support vector regression. Let y and $x_1, \dots, x_p$ be the dependent variable and the independent variables, respectively. Basically, SVR works as follows. First, a linear regression function, that is, a hyper-plane $h(x) = w_1 x_1 + \dots + w_p x_p + b$, must be defined. Then, a margin of tolerance ε is considered, expecting that all data will be at most at distance ε from the hyper-plane. In case the deviation of some points is larger than this value, *slack variables* $\xi, \xi' \geq 0$ can be introduced to cope with them. The final goal is to find the minimum of the function:

$$\min\left(\frac{1}{2}||w||^2 + C \sum_{i=1}^{N}(\xi + \xi') \right)$$

considering the restrictions

$$y - w \cdot x_i - b \leq \varepsilon + \xi$$
$$w \cdot x_i + b - y_i \leq \varepsilon + \xi'$$

In case of non-linear functions, SVR uses kernel functions to transform data into a higher dimensional space in order to develop a linear regression transformation.
- **Decision trees (DT).** The aim of this algorithm is to design a model for predicting a quantitative variable from a set of independent variables. The algorithm is based on a recursive partitioning. Trees are composed of decision nodes and leaves. DT regression usually is built by considering the standard deviation reduction to determine how to split a node in two or more branches. The root node is the first decision node that is divided on the basis of the most relevant independent variable. Nodes are split again by considering the variable with the less sum of squared estimate of errors (SSE) as the decision node. The dataset is divided based on the values of the selected variable. The process finishes when a previously established termination criterion is satisfied. The last nodes are known as leave nodes and provide the dependent variable predic-

tion. This value corresponds to the mean of the values associated to leaves. Figure 1a Balogun and Tella (2022) shows a graphical representation of the general structure of a standard Decision Tree.

- **Random Forest (RF).** Random Forest is based on the generation of several decision trees. The prediction will be the average of the predictions provided by the different trees. For the construction of each decision tree, a data sample is selected from the training dataset. The rest of the data will be used to estimate the decision tree error. The subset of independent variables that can be used for splitting each node are randomly selected. *Extremely randomised trees* (ERT) is a slightly modified random forest algorithm. Figure 1b Balogun and Tella (2022) shows a graphical representation of the general structure of a Random Forest Regressor.

- **K-nearest neighbours regression (KNN).** The *k*-nearest neighbours algorithm is often applied to classification problems although it can be also applied to regression problems. The idea of this algorithm is simple. Given a distance (Euclidean distance, Mahalanobis distance, etc) and a *k* value, the algorithm calculates the distance between a data point and the training dataset points for selecting the *k* nearest ones and establishes the average of them as prediction. An improvement of this algorithm is the algorithm known as *weighted k-nearest neighbours* (WKNN). In this case, the prediction considers a weighted arithmetic mean for calculating the prediction. Figure 1c Tella et al. (2021) shows a graphical representation of the general structure of a KNN model.

2.3 Deep learning algorithms

DL algorithms use ANNs. In this section, we will briefly describe different types of ANNs used in the literature for air quality prediction. In order to understand the internal behaviour of an ANN, we will introduce its structure.

An ANN is an algorithm based on the biological neuronal connections that are comprised of neurons or nodes. These connections are organised in three layer types. The input layer receives as input the original predictor variables. The output layer produces the predicted value for the given inputs. These two layers are connected by the the hidden layers. The hidden layers (more than one in the case of deep learning) contain non-observable neurons which are in charge of the computation. Each node in a layer is connected with nodes in the next layer. Each connection has associated a weight that is used to combine the inputs. Each node or neuron in the next layer receives the weighted value and transforms it by means of an activation function. Sigmoid and rectified linear unit functions are the most popular. The obtained result is the value that is passed as input to the nodes in the next layer. This process continues until the output layer is reached. At this point, the output prediction is produced.

The final aim of an ANN is to fit the weights to minimise an error function, commonly a quadratic function. To do this, the ANN uses the known as back-propagation algorithm. This algorithm employees the gradient descendent method using the layers partial derivatives to find the optimal weight of each node.

Now, we describe the different types of ANNs that have been used for predicting pollutant measures.

- **Multi-layer perceptron neural networks (MLP).** MLPs are classical neural networks with one or multiple hidden layers.

- **Convolutional neural networks (CNN).** CNNs are usually applied to images. They are an extension of MLP in which two types of layers are alternated: convolutional and pooling layers. The goal of convolutional layers is to extract features from the input image. A convolutional operation, which outputs convoluted features, is carried out by a matrix called kernel or filter. These features are the input of pooling layers. The aim of a pooling layer is to reduce the size of the convoluted features with the goal of decreasing the computational power required to process the image.
- **Recurrent neural networks (RNN).** RNNs deal with either time series or sequential data, that is, information ordered and related. RNNs have an internal memory in the sense that a neuron can feedback itself by receiving as input the output it has previously produced. This allows the model to acquire short-term memory which is essential to time series forecasting.
- **Long-short term memory neural networks (LSTM).** LSTMs are an extension of RNNs. They have an extended memory that allows to deal with long-term dependencies. LSTMs can remember information over arbitrary time intervals. The core component is the cell state that carries information throughout the processing of the data. The information is updated on the basis of three gates. Each of them controls the information that should be in the cell state using a sigmoid activation function. The forget gate determines which part of the previous state information should be forgotten. The input gate decides the new information that will be used to update the memory. Using an hyperbolic-tangent function, it creates a vector candidate to be added to the cell state. The last gate, known as output gate, uses a *tanh* activation function for determining which part of the updated cell state will be used as output.
- **Gated recurrent unit (GRU).** GRUs are a simplified version of LSTMs that combine the forget and the input gates. They provide similar results to the ones obtained by LSTMs.
- **Encoder-Decoder neural networks (EDNN).** The encoder-decoder model is a recurrent neural network used for sequence-to-sequence prediction problems. Its architecture consists of three components: encoder, intermediate vector and decoder. The encoder and the decoder are comprised of a set of recurrent units (usually, LSTM or GRU). In the case of the encoder, each unit processes an element of the input sequence and tries to encapsulate the associated information in the intermediate vector, for increasing the accuracy of the decoder prediction. The decoder uses this vector for producing the prediction.

3 Review methodology

In order to collect the papers included in this survey, we searched the main online research databases of technical publishers. Specifically, we considered ACM Library, Elsevier Online Library, IEEEXplore, SpringerLink and Wiley Online Library. We based the search on the occurrence of some keywords on the abstracts of the publications, with the exception of SpringerLink. In this case, it was not possible to restrict the exploration to the abstract section, so we search in the title of the papers. The usage keywords were "machine learning" "neural network", "regression" and "prediction", each of them combined with "air quality". We only collected papers published between 2011 and 2021. The total number of papers obtained from all considered databases was 781. Next, we examined the papers more deeply to select only those papers where air quality prediction was the main topic.

Fig. 2 Flow chart for paper identification and selection
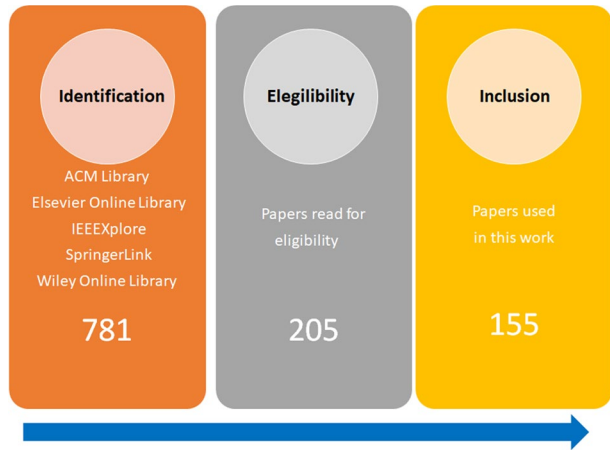
Fig. 3 Selected papers by database
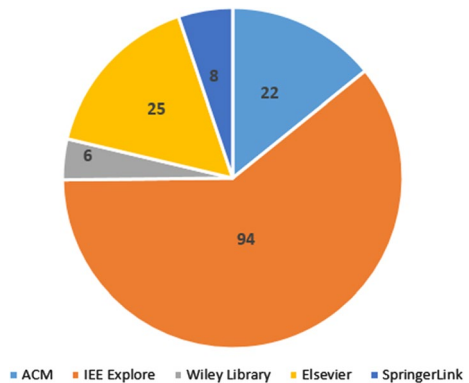
First, we removed duplicated papers and papers not written in English. In this phase, we also examined the publications in order to identified those ones that did not fall in the scope of the survey, that is, the application/development of ML algorithms to predict future values of AQI or pollutant concentrations. After filtering these papers, we obtained a set of 205 publications. Next, we went in depth in the reading of the papers for which it was not very clear whether they dealt with air quality prediction. In particular, we checked papers that did not make predictions about the future, but only analysed data to determine current AQI values. Finally, the total number of papers was reduced to 155. In Fig. 2, following the structure presented on a recent survey on the inter-correlation of climate change, air pollution and urban sustainability using novel machine learning algorithms and spatial information science Balogun et al. (2021), we present our flow chart for paper selection, screening, eligibility and inclusion.

In Figs. 3 and 4 we present the distribution of the final papers based on the search engines from which they were obtained and on the year of publication.

Next we analyse how we will classify publications in terms of geographical distribution, predicted values, predictor variables used for forecasting and evaluation metrics used to determine the models quality.
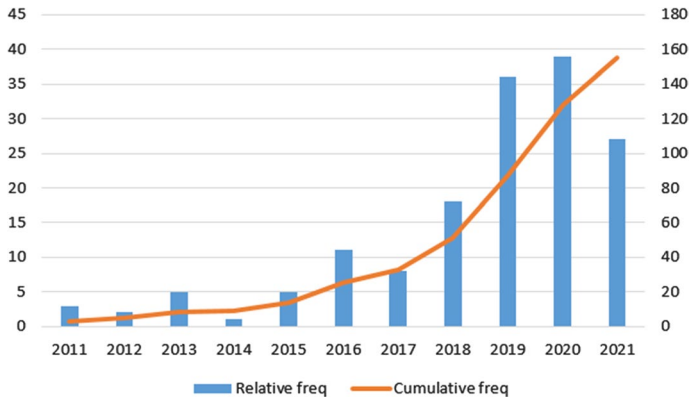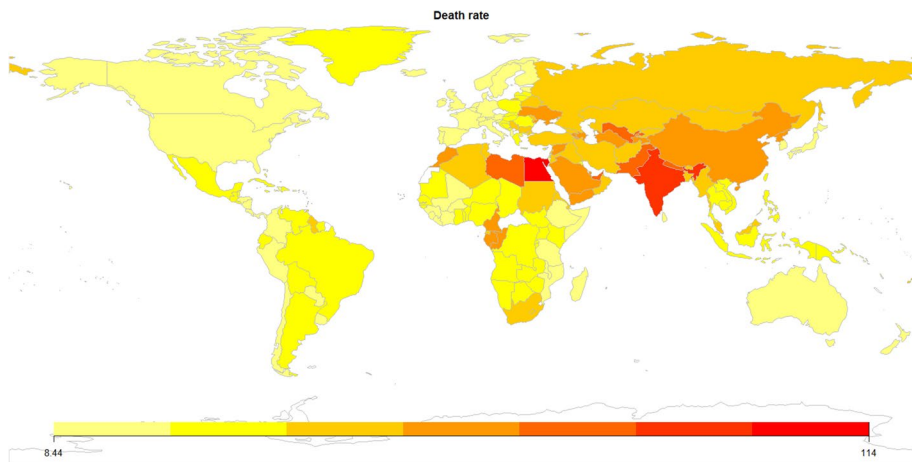
Fig. 4 Papers by publication year

Fig. 5 Death per 100.000 population caused by air pollution in 2017

3.1 Geographical areas

Air pollutant emissions strongly vary among different countries. Usually, developing countries are more polluted than developed countries. Specifically, East and South Asia are the most affected regions. According to *Our world in data*[3], in 2017, East Asia (9.75%), South Asia (7.63%) and Central Asia (6.19%) are the world regions that present the highest share of deaths by air pollution. In particular, China and India are the countries with more deaths due to this cause, 1.030.000 and 819.000, respectively, in absolute terms. In relative terms, Egypt (114.2 deaths per 100.000) and India (88 deaths per 100.000) are the most affected countries. As we can see in Fig. 5, Middle East and North Africa are also regions that suffer the air pollution impact in terms of human lives. In regions such as North Europe,

---

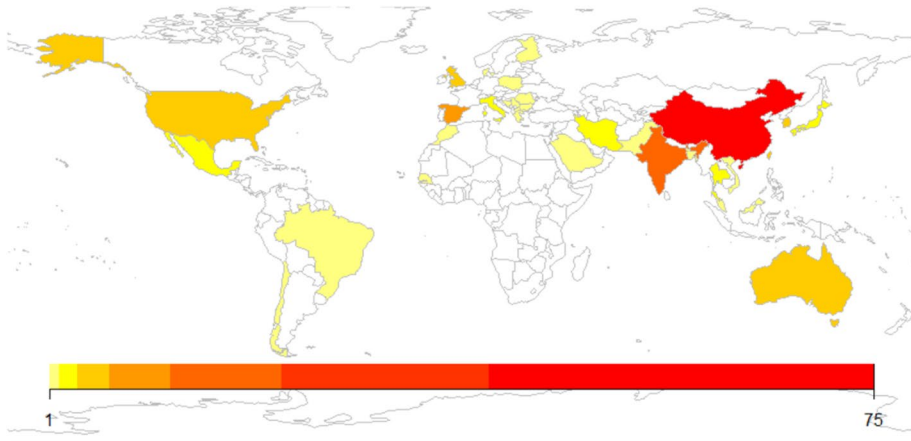[3] https://ourworldindata.org/, last accessed on 2022-02-15.

**Fig. 6** Number of papers by country in the survey

North America, Australia and New Zealand, the deaths caused by pollution are much lower than in the aforementioned regions.

Figure 6 illustrates how the specific country influences the number of publications related to the application of ML techniques for prediction of air pollution. The most influential advances in the field of ML applied to air quality prediction have been made by researchers in China. This fact is due to the scientific and technological potential of Chinese universities and the big air pollution problem they suffer. In this survey, more than half of the considered papers apply some ML technique to the air quality forecast of a Chinese city. The air quality of cities such as Beijing, Shanghai, Hangzhou and Guangdong has been analysed and predicted in many recent approaches using different ML algorithms (e.g. Ma et al. (2019); Yang et al. (2020); Lan et al. (2020); Han et al. (2021) among many others).

As we said before, India is also a country really harmed by air pollution, even more than China. According to the *IQ Air*,[4] nine of the top ten polluted cities in the world are Indian. Unfortunately, scientific and technological resources in India are lower than the Chinese ones and this is probably the main reason why papers that apply ML techniques to predict either air quality or pollutant indexes in Indian cities is very low. Nevertheless, many of the publications that compare different models to forecast air quality use data from Delhi. We can also see that in Europe and America, ML is not so widely used as in Asia for predicting air pollution. In Europe, Spain is the country in which most papers have been produced (7) followed by United Kingdom (6). Further on, one can see that in this survey there are papers studying air quality of many countries in East Europe. It may be because this is the European region most affected by pollution. In the rest of the world, Australia (6) and USA (4) lead the way.

---

[4] https://www.iqair.com/, last accessed on 2022-02-15.

Fig. 7 Publications by predicted pollutants
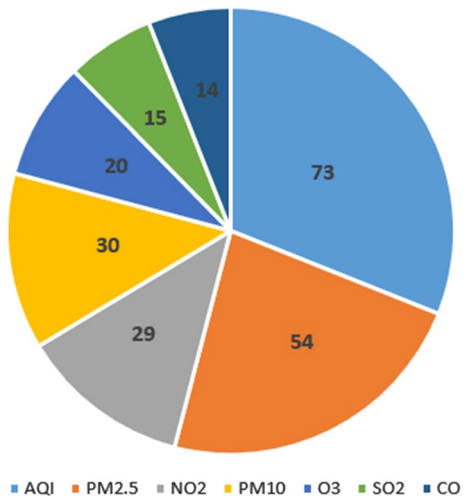
3.2 Predicted values

AQI is used by institutions to provide the population with information about the pollution degree of cities or neighbourhoods taking into account the main pollutant gasses. Therefore, it is natural that AQI is the most studied magnitude for being predicted in order to determine the air quality. It is worth pointing out that different countries use different AQI scales. We can mention the following four indexes.

- The Common Air Quality Index,[5] used in Europe since 2006, considers 5 risk categories and ranges from 0 (low risk) to 100 (air quality is unhealthy).
- The China Air Quality Index[6] has 6 risk categories and ranges from 0 (excellent) to 500 (severed polluted).
- The US Air Quality Index[7] runs from 0 to 500. These values are divided into 6 categories that correspond to different levels of air quality.
- In the case of United Kingdom,[8] it runs from 0 to 10, classifying these values into 4 levels.

There exist several pollutants that are also considered for their prediction: carbon monoxide(CO), ozone (O_3), sulphur dioxide (SO_2), nitrogen dioxide (NO_2), nitrogen oxides (NO_X) and particulates matter ($PM2.5$, $PM10$). However, we are not aware of a categorisation that clearly shows the harmfulness of each pollutant. The only data that could be used for analysing the level of danger of these pollutants is the one related to the premature deaths caused by each of them. Unfortunately, due to the difficulty in determining this fact,

---

[5] https://www.airqualitynow.eu/about_indices_definition.php, last accessed on 2022-02-15.

[6] https://datadrivenlab.org/air-quality-2/chinas-new-air-quality-index-how-does-it-measure-up/, last accessed on 2022-02-15.

[7] https://www.airnow.gov/aqi/aqi-basics/, last accessed on 2022-02-15.

[8] https://web.archive.org/web/20111115230340/http://www.comeap.org.uk/images/stories/Documents/Reports/comeap%20review%20of%20the%20uk%20air%20quality%20index.pdf, last accessed on 2022-02-15.
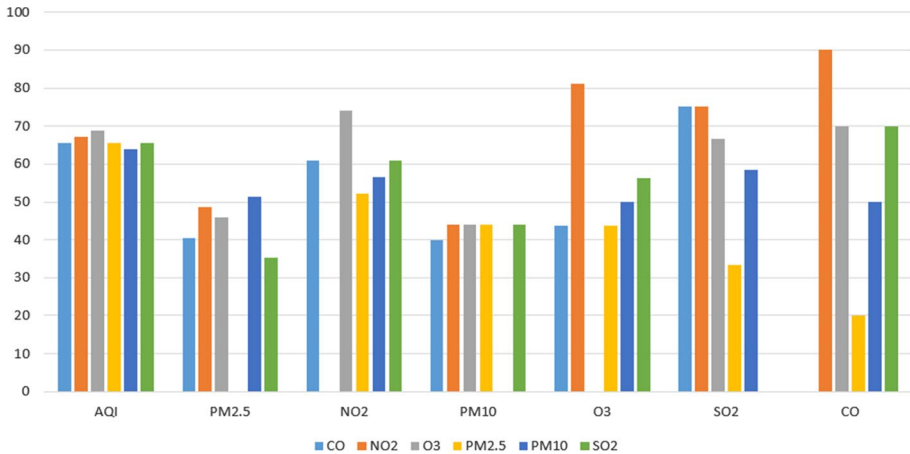
Fig. 8 Pollutant variables by dependent variable (percentage)

it is not easy to find this information. We just found data of premature deaths attributable to $PM2.5$, NO_2 and O_3 from the European Environment Agency.[9] According to this source, in 2018, $PM2.5$, NO_2 and O_3 were the cause of 379.000, 54.000 and 19.000, respectively, premature deaths in the European Union. There are no data for other pollutants. According to the World Health Organisation,[10] $PM10$ are unhealthy for human, but no so much as $PM2.5$ because of its size. In the case of CO, according to the UK Committee on the Medical Effects of Air Pollutants,[11] during the last years there has been an appreciable reduction of outdoor concentrations which makes that its danger has also decreased. These considerations help us to justify the correlation between the danger of the pollutants and the volume of papers devoted to forecasting them. Figure 7 shows the number of analysed publications in this survey for the prediction of each of the pollutants. Note that several papers predict more than one pollutant.

3.3 Predictor variables

In this Section we will analyse the independent and predictor variables that are used by models for pollutant forecasting. We have only considered those variables that are used by at least 5% of the models developed in the analysed publications. We distinguish three types of variables.

- *Pollutant variables* correspond to pollutant concentrations.
- *Weather variables* are associated with different weather elements: wind direction, wind speed, atmospheric pressure, rain, relative humidity, temperature and dew point.
- *Other variables* include different data such as year, season, month, week, weekday, hour, geolocation and visibility.

---

[9] https://www.eea.europa.eu, last accessed on 2022-02-15.
[10] https://www.who.int/, last accessed on 2022-02-15.
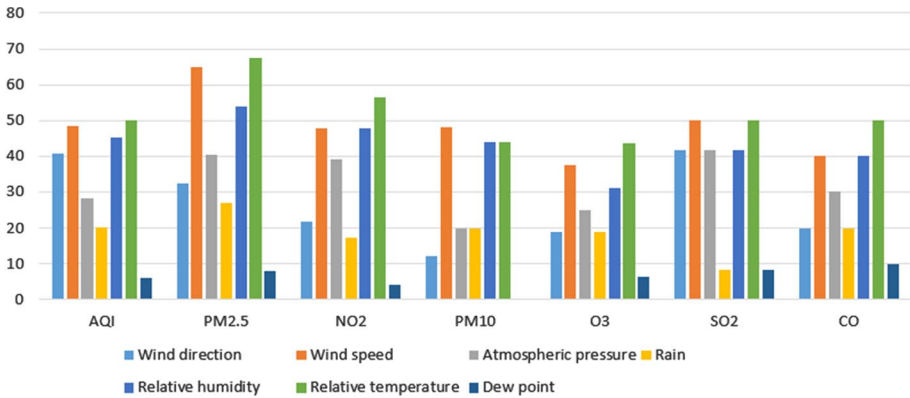[11] https://www.comeap.org.uk, last accessed on 2022-02-15.

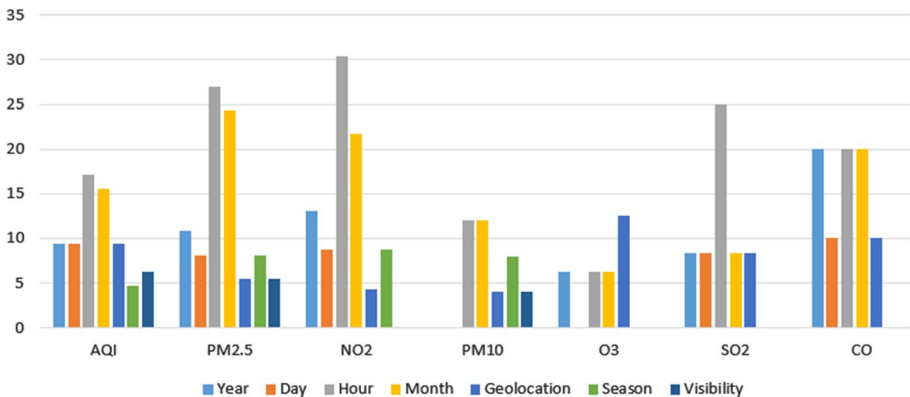**Fig. 9** Weather variables by dependent variable (percentage)

Fig. 10 Other variables by dependent variable (percentage)

Note that several papers included in this survey use two or more dependent variables.

Next, for each of these categories, we enumerate the percentage of analysed publications that use them as independent variables for predicting AQI, CO, O_3, SO_2, NO_2, $PM2.5$ and $PM10$.

Figure 8 shows the relevance of pollutant variables for predicting the different dependent variables. As can be noted, all the predictor variables are used in a similar percentage in the design of the analysed models. It is not surprising that all the pollutant variables are used in equivalent percentages for AQI prediction. This is due to the fact that all the pollutants have the same relevance in the piece-wise linear function to estimate the AQI value. However, there exist differences in the case of the prediction of other pollutants. It is remarkable the influence of NO_2 at the time of predicting CO. NO_2 is also the most important pollutant feature when other pollutants are predicted, like CO, O_3 or $PM10$. On the contrary, $PM2.5$ is the pollutant variable less used in the design of models for predicting other pollutant concentrations. The rest of pollutant variables are used in a similar percentage.
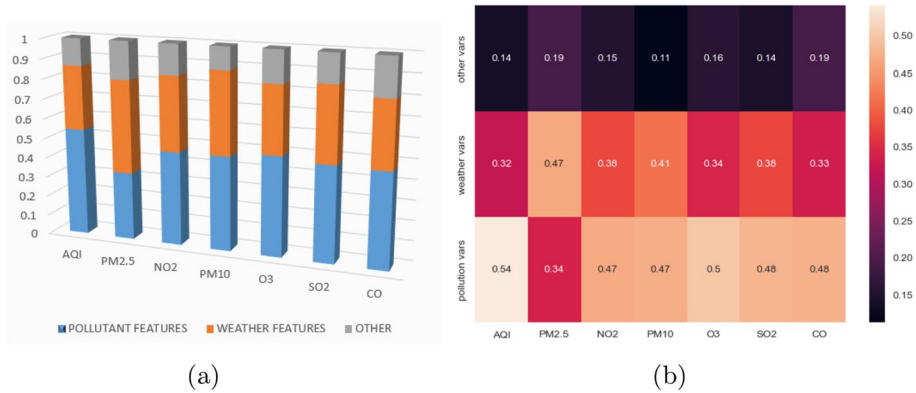
Fig. 11 Use of predictor variables by dependent variable

Figure 9 presents the use of weather variables as predictors in AQI and pollutant forecasting. The first thing that can be observed is that although all of them are considered to some extent in the design of the models, their presence in the different proposals is not as important as that of the pollutant variables. The relative temperature, the relative humidity and the wind speed are the weather predictors most included in the models (used in around a 50% of them). In a lower percentage we find the atmospheric pressure, the wind direction and the rainfall (used in around a 35% of the models). We can also detect a significant difference between the use of wind speed and wind direction in the prediction of all the pollutants: the latter is more widely used than the former. The less used variable of this type is the dew point, (used by less than 10% of the models). We can also observe that the dew point is not taken into account for predicting $PM10$.

Figure 10 shows the influence of other variables in the proposed models. In this category, the different approaches included in this survey mainly consider variables related to the timeline such as hour, day, month, weekend day and year. In addition, we also analysed the impact of the geographical location, that is, the geolocation and the visibility. Although other features appear in some models, like illumination, snow, land use and solar radiation, among others, as we previously said, we have only considered the variables used by at least 5% of the proposals. The predictor variables included in this category are not uniformly distributed in the models for the prediction of the different pollutants. In general, the most frequent variables are month and hour (used in around 25% of the models). These predictor variables are really significant in the predictions as external factors. In the case of the hour, the nighttime, for example, usually is associated with a reduction of the traffic and the industrial activity. Consequently, it should imply a decrease in the pollutant concentrations. In the case of the month, something similar happens. For instance, in many cities the work activity during summer is notably lower than during the rest of year. Therefore, a decrease in pollutant concentrations can be expected. On a second level we find geolocation, day and year (used in around 15% of the models). Finally, season and visibility are the less used variables of this category. In particular, season is not used in any proposed model for the prediction of CO, SO_2 and O_3. O_3 predictor models do not use day and $PM10$ predictor models do not use year. On the contrary, hour is the most used variable of this type in all the predictions except for predicting O_3.

Table 1 Range-dependent metrics

MSE	Mean square error	$\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$
RMSE	Root mean square error	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}$
MAE	Mean absolute error	$\frac{1}{n}\sum_{i=1}^{n}|Y_i - \hat{Y}_i|$
MRE	Mean relative error	$\frac{1}{n}\sum_{i=1}^{n}|\frac{Y_i - \hat{Y}_i}{Y_i}|$

Table 2 Percentage metrics

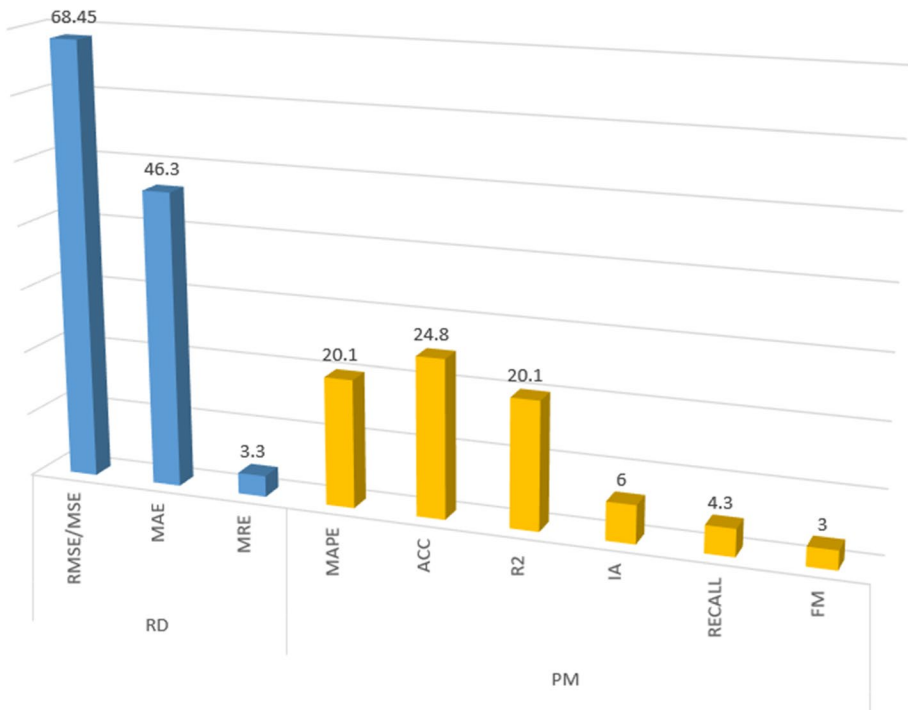| MAPE	Mean absolute percentage error	$\frac{100}{n}\sum_{i=1}^{n}|\frac{Y_i - \hat{Y}_i}{Y_i}|$
ACC	Accuracy	$1 - \frac{\sum_{i=1}^{n}|\hat{Y}_i - Y_i|}{\sum_{i=1}^{n}Y_i}$
R^2	R squared	$\sqrt{\frac{\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}}$
IA	Index of agreement	$1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n}(|Y_i - \hat{Y}_i| + |\hat{Y}_i - Y_i|)^2}$
RECALL	Sensitivity	$\frac{TP}{TP+FN}$
PREC	Positive predictive value	$\frac{TP}{TP+FP}$
FM	F measure	$\frac{PREC \cdot RECALL}{PREC + RECALL}$

Fig. 12 Use of evaluation metrics
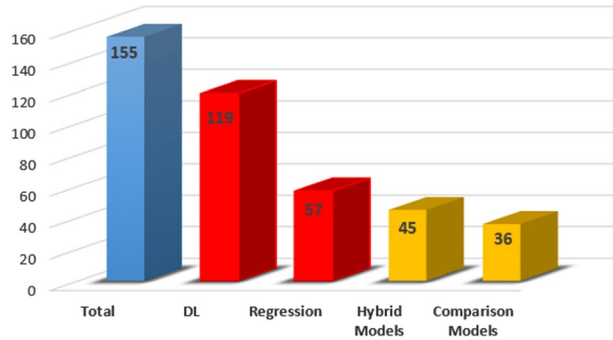
Fig. 13 Papers by algorithm type

Figure 11 illustrates the influence of the different types of variables in the prediction of the different pollutants. In all the cases, except *PM*2.5, the pollutant variables are the most used (around 45%), followed by the weather variables (around 35%). The rest of predictor variables are only used by around 15% of the models. However, in the case of *PM*2.5 the weather variables (50%) have an essential role being more used than the pollutants variables (33%).

3.4 Evaluation metrics

Evaluation metrics help us to determine the quality of an ML model. We classify them in two groups. On the one hand, range-dependent metrics compare different models that must be applied to the same dataset. On the other hand, percentage metrics compare models independently of the dataset used.

The range-dependent metrics used in the analysed publications are described in Table 1, where Y_i and $\hat{Y}_i$ correspond to the current and the predicted values of the *ith* observation, respectively, and n represents the number of observations. In Table 2 we can see the percentage metrics, where $\bar{Y}$ indicates the mean of the observations, *TP* the number of true positives, *FN* the number of false negatives and *FP* the number of false positives.

As we can see in Fig. 12, the most applied metrics in the analysed proposals are ACC (24.8% of the publications), R^2 and MAPE (20.1% each of them). Regarding range-dependent metrics, RMSE/MSE and MAE are the most used (68.45% and 46.3% of the publications, respectively). Nevertheless, as we previously said, this type of metrics does not allow comparisons between models.

The fact that pollutant variables are the most used ones is related to the strong influence on the predicted variable. Weather variables are less used than pollutant ones. Nevertheless, all but *dew point* are frequently considered. This is not surprising because meteorology is highly related to air quality. For example, in the case of ozone it is important to take into account that it is formed in the presence of sunlight. Finally, other variables related to date and geolocation have influence on the predicted variable but it is not significant. Therefore, we think that the presence of a dependent variable is directly related to its influence in the variation of the pollutant concentrations and air quality index.
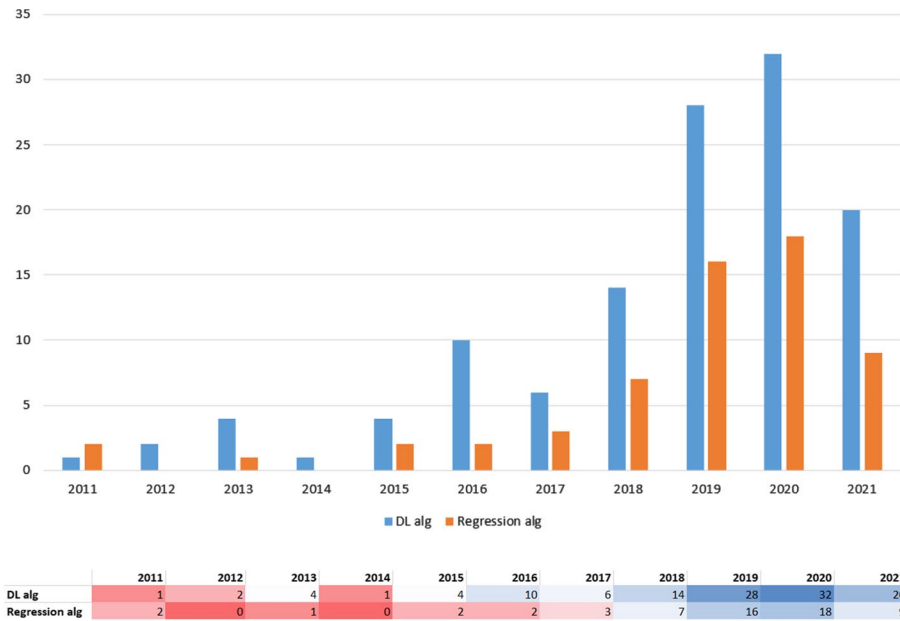
Fig. 14 Algorithm type used by year of publication

4 Classification, by used model, of the contributions on air quality forecasting

In this section we summarise the main contributions concerning the application of ML and classical regression algorithms to the design of air quality forecasting models.

It is worth pointing out that some of the publications that have been studied consider several algorithms in the proposed approaches. This happens in both publications that *compare* different models and those papers in which a *hybrid model* is proposed. A hybrid model uses several algorithms to improve the performance with respect to the use of a single algorithm. In some cases, these proposals combine both deep learning and regression algorithms. Therefore, they could be classified in both types. Taking this fact into account, Fig. 13 shows the amount of publications corresponding to the different categories. In our study, we have analysed 35 publications where comparisons between approaches are studied (22.5%) and 45 papers which present hybrid models (29%).

We can also notice that the use of DL algorithms in the design of the models is most frequent than the use of regression algorithms. A total of 119 papers deal with DL methods, which represents 76.8% of all the publications. Regression algorithms are considered only in 57 papers, that is, 36.8% of the total number of them. This difference may be due to the "machine independence" advance of DL methods with respect to the classical regression algorithms. It makes these algorithms fit better than regression algorithms in the case of air quality forecasting because it requires the use of many predictor variables whose distribution and correlation with the target variable are not regular. Figure 14 presents the evolution of the use of both types of algorithms during the last 10 years. The publications that consider algorithms corresponding to both deep learning and regression have been
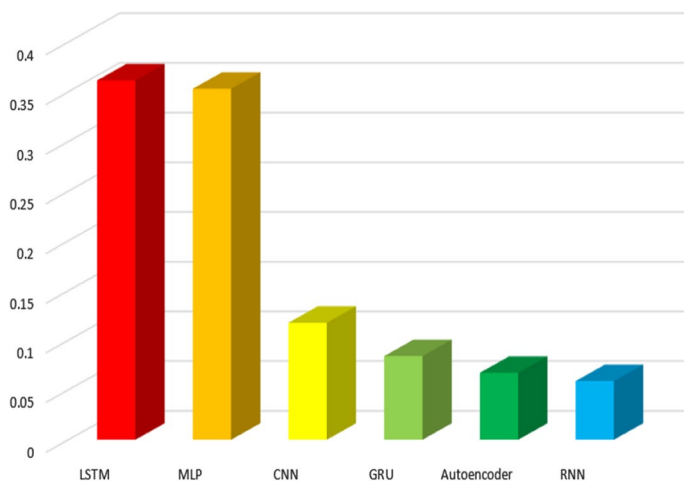
Fig. 15 DL algorithms type proportion respect to the total number of DL algorithms

included in both categories. Despite the number of publications corresponding to each of them, the increase of proposals during the last 3 years has been similar in both cases.

4.1 Deep learning algorithms

In this Section we consider those papers that apply a single DL algorithm. Approaches that combine different DL methods in the proposal will be described in Sect. 4.3. In the same way, those works that focus on comparing different algorithms will be discussed in Sect. 4.4.

Figure 15 shows the different DL algorithms used in the analysed publications. LSTM networks and MLP networks are the most frequently used (36.13% and 35.3% respect to all deep learning papers, respectively), followed by CNN. Next, we briefly describe the proposals which consider these algorithms.

Several papers apply MLP for predicting pollutant concentrations. In this group we can mention Paoli et al. (2011), focusing on O_3 in Corsica, Yao et al. (2012), predicting $PM2.5$ in Northern China, Huang et al. (2015), forecasting AQI in Shijiazhuang and Agarwal, Dedovic et al. (2016), predicting $PM10$ in Sarajevo, and Agarwal et al. (2020), considering $PM10$, $PM2.5$, NO_2 and O_3 in Delhi. In these proposals, the models are classical MLP neural networks without either significant improvements or variations. The highest number of hidden layers (15) corresponds to Dedovic et al. (2016). The highest number of predictor variables appears in Agarwal et al. (2020).

Other authors consider extensions and variations of the classical MLP model. Although the back-propagation algorithm is the most common training algorithm for neural networks, it can be improved as shown by Hoi et al. (2013). In order to make a prediction of the $PM10$ concentration in Macao, these authors use a Kalman filter in the learning algorithm to build a time-varying MLP neural network. Results show that the 3 layer time-varying MLP neural network is even more accurate than a 10 layer classical MLP neural network. Fu et al. (2015) developed an MLP network with rolling mechanism and accumulated generating operation of gray model to predict $PM2.5$ and $PM10$ concentrations in three Chinese cities. The original model was modified for improving its deficiency to

assess the possible correlation between different input variables. Jiang and Li (2016) propose an MLP model for forecasting AQI in three Chinese cities using geo-sensor data using participatory sensing. Moreover, inspired by the dropout mechanism, the authors introduce a new approach in order to handle the problem of unsensed variables. A classical MLP improved by a linear regression technique is proposed by Thanavanich et al. (2021) to forecast $PM2.5$ in Thailand. Huang et al. (2020) introduce an improved MLP model employing Particle Swarm optimisation, a meta-heuristic evolutionary algorithm which simulates the behaviour of particles in nature. The target data is AQI in China.

LSTMs are also used in air quality forecasting models due to its effective performance in time series. Several works propose an LSTM-based model for prediction of different pollutants. Kök et al. (2017); Schürholz et al. (2019); Alhirmizy and Qader (2019); Li et al. (2020); Schürholz et al. (2020) apply LSTM to forecast AQI in Aarhus, Brasov, Australia, Madrid, and Melbourne, respectively. Alsaedi and Liyakathunisa (2019) and Jiao et al. (2019) present approaches to predict the level of AQI in Beijing and London and in Shanghai, respectively. Zhou et al. (2019) focus on the forecasting of $PM2.5$, $PM10$ and NO_X in Taipei. Finally, Navares and Aznarte (2020) considered the levels of CO, NO_2, O_3, $PM10$, SO_2 and pollen in Madrid.

Many authors consider some variations of the classical LSTM model. Wang and Song (2018) introduce a three component algorithm to predict AQI in Beijing. The first component trains and combines dynamically several individual models that are based on weather patterns. The second component finds the spatial correlation between areas. The third component is an LSTM neural network. Song et al. (2019) propose an LSTM-Kalman model to predict three pollutants on experimental data-sets. Their results show that the proposed model has a higher accuracy than a classical LSTM model. Lan et al. (2020) propose an LSTM model to predict AQI in Hangzhou. The model presents a spatial optimisation component considering the correlations between variables and a temporal optimisation method considering the time window size. Their results also show the improvement of accuracy of the proposed model with respect to a classical LSTM. Seng et al. (2021) define a multi-index multi-output model based on LSTM to predict $PM2.5$, CO, NO_2, O_3 and SO_2 in Beijing. Han et al. (2021) train a Bayesian LSTM to predict AQI and $PM2.5$ in Beijing during a blue-sky period. Mao et al. (2021) proposed a temporal-sliding LSTM model to predict $PM2.5$ in the Jing-Jin-Ji region, in China. The model integrates the optimal time lag of spatio-temporal correlations to improve the model. Their results show that optimising the temporal-sliding LSTM reduces classical LSTM and MLP error. Jin et al. (2021) developed a nested LSTM network to forecast AQI in Beijing. Finally, Wang et al. (2021) propose an algorithm that uses a chi-square test to determine the most influential factors and an LSTM to forecast Shijiazhuang AQI. The results show that this model presents a smaller error than other classical methods.

Bidirectional LSTMs have been also considered in some works. The main difference between this algorithm and a classical LSTM is that the bidirectional LSTM preserves the past and future information, while the classical LSTM only preserves the past. This method is applied by Dua et al. (2019); Madaan et al. (2019) to predict $PM2.5$, $PM10$ and NO_2 in Delhi. Ma et al. (2019) consider a bidirectional LSTM and also apply transfer learning to transfer the knowledge learned from smaller temporal resolutions to larger temporal resolutions in order to forecast $PM2.5$, NO_2 and O_3 in Anhui, East of China. Ma et al. (2019) use a bidirectional LSTM combined with an inverse distance weighting technique for the spatio-temporal forecasting of $PM2.5$ in Guangdong. Ma et al. (2020) use a bidirectional LSTM, applying transfer learning, to predict $PM2.5$ also in Guangdong. Zhang et al. (2021) propose a bidirectional LSTM model, which includes an empirical mode decomposition,

to predict *PM*2.5 in Beijing. Zhang et al. (2021) apply a bidirectional LSTM after they decompose the original time series with a variation mode decomposition method to predict *PM*2.5 in several Chinese cities.

Since LSTMs are considered an improvement of RNNs, that is the reason why RNNs are rarely used in air quality prediction. Ong et al. (2016) propose an RNN algorithm to predict *PM*2.5 in Japan. They employ a dynamic method to pre-train the model based on multi-step-ahead time series prediction. Lim et al. (2019) apply RNN to predict *PM*10, O_3, SO_2, *CO* and NO_2. In this case, the model corresponds to an RNN without significant improvements or variations with respect to the classical model.

GRUs are used in a similar proportion as RNNs, that is, scarcely. Zhang et al. (2020); Sonawani et al. (2021); Liu et al. (2021) propose enhanced GRU models for predicting AQI in the Huaihai Economic Zone, NO_2 in Pune (India) and *PM*2.5 in Beijing, respectively. The first two contributions apply a bidirectional GRU.

Finally, it is worth noting that a very low percentage of publications consider other algorithms. Multitask learning is a subfield of ML that improves the learning of a task by taking into account information of a related task. The idea is that what is learned for a task can help to better learn other tasks. A deep belief network with multi-tasking learning is employed by Li et al. (2019) to obtain a high level of accuracy. Sun et al. (2020) propose a model that uses multitask methods and RNN to forecast AQI in China. Barrón-Adame et al. (2012) propose a self organising map neural network to forecast SO_2 in Salamanca, Mexico. Wahid et al. (2013) forecast the O_3 concentration in Sydney considering a radial basis neural network. Fuzzy inference neural networks are used by Pai et al. (2013) to forecast oxidant concentration in Tokyo and by Ganesh et al. (2017) to forecast AQI in Delhi. Metia et al. (2014) consider a Matérn function based on a fractional Kalman filtering to forecast O_3 concentration in Sydney. De Vito et al. (2015) deal with a dynamic multivariate regression to forecast air quality in Cambridge. Some papers Liu et al. (2016); Zhao et al. (2016); Zhu et al. (2020); Sharma et al. (2020) introduce extreme learning machine models to forecast AQI in Beijing, air quality in Helsinki metropolitan area, AQI in Liaoning (China) and particulates matter in four Australian cities, respectively. Mei et al. (2016) develop an Elman neural network model using a genetic algorithm (GA) to optimise the initial weights and the number of hidden layer nodes. This model is applied to forecast AQI from a public data-set chosen from the UCI ML repository. A wavelet neural network is proposed by Zhang et al. (2018) to make a short-term forecasting of AQI in Beijing. An air pollution forecasting system which uses deep distributed fusion network is proposed by Yi et al. (2018), providing fine-grained AQI forecasts for more than 300 Chinese cities. Li et al. (2018) consider data spatial-temporal features to propose a spatio-temporal aware sparse denoising auto-encoder to forecast AQI in China. Their results show that use of spatial-temporal features increases accuracy. Yan et al. (2018) use an EDNN to predict *PM*2.5 in Tianjin. Wu and Zhang (2019) propose a generative additive model to forecast *PM*2.5 in Beijing. A leveraging bagging ensemble learning algorithm is proposed by Liu et al. (2019) to predict the air quality grade in Beijing. Finally, Yi et al. (2020) apply a distributed fusion network for short-term predictions and a deep distributed cascade network for long-term predictions to forecast the value of AQI of the next 48 h and the daily average air quality of the next 7 days.

We think that the popular use of LSTM is due to its high adaptability to solve sequential problems in general and time series problems in particular. In the case of MLP, we think that they are very often used due to is popularity, although its usage in time series problems can present some difficulties. CNN, GRU and Auto-encoders are algorithms whose main application field is very different to time series. We think that this is the cause of
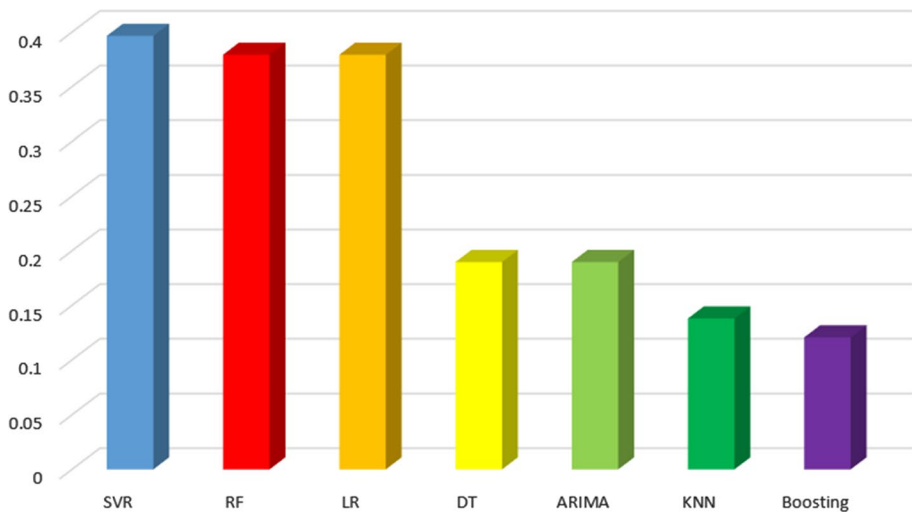
Fig. 16 Regression algorithms type proportion respect to the total number of regression algorithms

their low use. Finally, the popularity of RNNs has decreased in recent years because LSTM improves their performance.

4.2 Regression algorithms

In this section we consider contributions that apply a unique regression algorithm. Approaches that combine different regression methods in the proposal will be described in Sect. 4.3. Likewise, those papers that compare different algorithms will be discussed in Sect. 4.4.

Figure 16 shows the percentage of analysed publications that deal with the different regression algorithms. We can see that SVR, RF and linear regression are the most used methods in the analysed works. Note that the percentages add up to more than 1. The reason for this is that many papers introduce more than a single regression algorithm.

An SVR model with Gaussian and polynomial kernel is proposed by Sotomayor-Olmedo et al. (2011) to forecast PM2.5 and PM10 in London. A SVR model is developed by García Nieto et al. (2013) to built a non-dynamic forecasting model of AQI in Oviedo, Spain. SVR with quasi-linear kernel is designed by Zhu and Hu (2019) to predict CO, NO_X and NO_2 in an open data-set. In this case, their results show that quasi-linear SVR outperforms other learning methods.

Zhang et al. (2015) develop a model, in Apache Spark, using a parallelised RF algorithm to forecast PM2.5 in Beijing. The model takes into account 12 predictor variables and the results show that their improvements imply high accuracy in the model. Li et al. (2018) propose an algorithm to predict PM2.5, NO_2 and SO_2 in Beijing. They apply a method based on an RF for the selection of the most relevant factors. Initially, they obtain 20 predictor variables for forecasting the concentration of the pollutants but some of them are removed by using a selection method based also on an RF.

Linear regression (simple or multiple) is the most used *classical* regression algorithm in the literature. Barthwal and Acharya (2018) propose a linear regression model to predict AQI levels in New Delhi by taking into account 8 predictor variables.

Jato-Espino et al. (2018) use clustering to partition the air quality monitoring stations according to similarity characteristic. Then, they apply multi-linear regression to predict AQI in Catalonia. Durić and Vujović (2020) apply multi-linear regression, using 3 predictor variables, to forecast AQI in four stations of Belgrade.

A similar model is applied by Ma et al. (2020) for predicting AQI in North China using SPSS program. They use 7 predictor variables and obtain a highly accurate model. LASSO, a multi-linear regression regularise method for variable selection is used by Sethi et al. (2021) to predict AQI in Delhi.

Some works deal with ARIMA algorithms. Liu et al. (2018) use an ARIMA algorithm combined with numerical modelling to forecast $PM2.5$, O_3 and NO_2 in Hong Kong. The effectiveness of this combination is shown by empirical results. An ARIMA model is developed by Ngom et al. (2021) to forecast $PM10$ with a multi-agent real-time simulation in Dakar, Senegal. They also discuss the relevance of studying other parameters of the model. A seasonal ARIMA method is developed by Bhatti et al. (2021) to predict $PM2.5$ in Lahore, Pakistan. Hajmohammadi and Heydecker (2021) propose a vector auto-regressive moving-average model to predict $PM10$, NO, NO_X and NO_2 in London.

Another works consider the use of other regression algorithms. A method which uses rolling window regression is developed by Lang et al. (2019) to estimate trends in NO_2, NO_X and NO_2/NO_X in London. Wang and Kong (2019) develop a DT to predict AQI for different locations in China. The model is improved in the feature attribute value and in the weighting of the information gain. Results show that this model provides best accuracy than other improved DT models. Zhang et al. (2019) propose an improved light gradient boosting machine method to predict $PM2.5$ in Beijing and Matović and Vlahović (2021) propose a boosting algorithm to forecast 6 pollutants in Beijing.

We think that the popular use of SVR and RF is due to its great adaptability to different contexts and to different type of variables (this is specially remarkable in the case of RF). In the case of LR, its significant percentage of use can be due to being the basis of regression algorithms. DT is scarcely used because RF provides a clear improvement and, with the same knowledge, one can apply a DT or an RF indistinctly. In the case of ARIMA, we think that it is not used very often because its performance decreases when many predictor variables are included, as in the cases that concern us. Finally, KNN and Boosting are more popular to solve classification tasks than regression tasks.

4.3 Hybrid models

In this section, we will review papers that consider different algorithms in their proposals.

First, we review hybrid models that only deal with deep learning algorithms. Zheng et al. (2013) developed a co-training model to forecast the $PM2.5$ concentration in Beijing and Shanghai. The proposed model is a linear-chain conditional random field model used to model the temporal predictor variables and an MLP network with one hidden layer which takes spatially-related predictor variables as input. Lin et al. (2018) present a geo-context based diffusion convolutional recurrent neural network, a hybrid CNN-RNN model, for $PM2.5$ short-term forecasting in Los Angeles and Beijing. The combination of an EDNN and a GRU results in a model based on n-step recurrent prediction. This model, named *seq-2-seq*, is applied by Liu et al. (2019) to forecast AQI in Beijing. Zhang et al. (2020) propose a hybrid CNN-GRU multi-task learning model to predict $PM2.5$ in three districts of Lanzhou, China. Hu et al. (2021) combine an MLP network with a linear interpolation method to forecast $PM2.5$ and $PM10$ in two Chinese cities. A hybrid EDNN-LSTM model

was developed by Abirami and Chitra (2021) to predict AQI in Delhi. In this case, the EDNN encodes the spatial relations in data. A hybrid model based on a general regression neural network, a extreme learning machine and a variational mode decomposition, was proposed by Rahimpour et al. (2021) to predict AQI in Urmia, Iran.

The combination of LSTM and CNN increases the model efficiency due to both the LSTM long-term forecasting capability and the CNN capability to learn data features without the need of extracting them in advance. Hybrid models based on this combination are developed by Zhao et al. (2019); Li et al. (2020); Kiftiyani et al. (2021); Jovova and Trivodaliev (2021); Le et al. (2020); Putra et al. (2021) to predict AQI in the Beijing-Tianjin-Hebei region, PM2.5 in Beijing, 6 pollutant concentrations in Seoul, PM10 in Skopje, AQI in Seoul and PM2.5 in Taiwan, respectively. Previously, Soh et al. (2018) apply a similar hybridisation in which an MLP is also combined with a CNN and a spatio-temporal LSTM. A hybrid model which combines LSTM, GRU and CNN is developed by Chiang and Horng (2021) to forecast daily PM2.5 in Taiwan.

As we previously said, some proposals consider the use of GAs to determine the optimal initial weights. Zhenghua and Zhihui (2017) apply a hybrid GA-MLP model to predict AQI in Xuchang. Kang and Qu (2017) and Su and Xie (2020) apply this hybridisation to forecast AQI in Lanzhou and Dailan, respectively. The results show that GA optimisation reduces the model error with respect to a non-optimise MLP. Mu et al. (2017) and Yang et al. (2020) propose a similar algorithm to predict AQI in Taicang and Shanghai, respectively. In these works, principal component analysis (PCA) is used to reduce dimension before the application of the algorithm and shows its effectiveness for increasing the accuracy. Lin et al. (2020) develop a hybrid approach which includes a self-organising-maps neural network, a GA for optimisation and PCA for reducing the dimension of a neuro-fuzzy modelling, a combination of a neural network and fuzzy logic, to predict AQI in Taiwan.

Some publications propose other combinations of deep learning algorithms. Ao et al. (2019) use a bidirectional LSTM combined with a k-means clustering to forecast AQI in Qingdao. A Bayesian encoder-decoder hybrid RNN is developed by Han et al. (2020) to forecast PM2.5 and PM10 in London and Beijing. Zhao et al. (2020) propose a hybrid model which combines the advantages of both forward neural networks and recurrent neural networks to forecast AQI in Lanzhou and Xi'an. A hybridisation of GA and the Encoder-decoder method is developed by Nguyen et al. (2021) to predict PM2.5 concentrations in Hanoi and Taiwan. (Ouyang et al. 2021) combine a dynamic graph CNN with a GRU to forecast PM2.5 concentration in two datasets belonging to Beijing and London. A hybridisation between a CNN with residual connections, attention layers and a bidirectional LSTM is developed by Choi et al. (2021) to forecast PM2.5 and PM10 in Seoul. Colchado et al. (2021) propose a model composed by a KNN-attention pooling layer stacked to a classical MLP network to make a spatial prediction of PM2.5 in São Paulo.

Finally, we will review those papers which only consider the hybridisation of regression algorithms. Yi et al. (2019) propose a hybrid boosting-random forest model to classify AQI in Fangchenggang and Beijing. The model uses samples grouped bootstrap instead of a single bootstrap method in the random sampling phase. Results show that this improvement provides best accuracy than an RF method. Gocheva-Ilieva et al. (2020) propose a hybrid model that combines RF and ARIMA: RF is used to analyse meteorological feature relations and ARIMA is used to correct the model. The model is applied to forecast the concentration of 7 pollutants in Dimitrovgrad, Bulgaria. Hybrid models based on the combination of ARIMA, optimised extreme learning machine and fuzzy time series models are developed by Li et al. (2020) to forecast AQI in Shijiazhuang, Zhengzhou and Guangzhou.
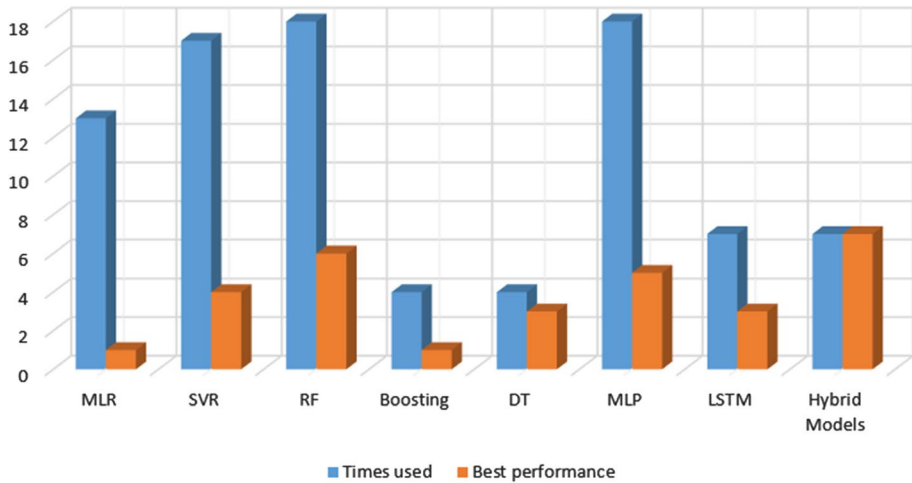
Fig. 17 Algorithm type used in papers comparing models and number of times showing the best performance

Some works also propose hybridisation of deep learning and regression algorithms. Bougoudis et al. (2016) propose a hybrid machine learning system which combines RF, self-organising maps and fuzzy inference systems that has been applied to forecast AQI in Athens. Chen et al. (2017) propose a hybrid model which combines an SVR and an Elman neural network to predict PM2.5 in Wuhan. Wang et al. (2017) present a hybrid model where they optimise the parameters using four probabilistic parameter models. Then, an improved MLP neural network with signal processing techniques is employed to predict AQI in several Chinese cities. Qi et al. (2018) introduce a hybrid model which combines an ECNN and a spatio-temporal semi-supervised regression model to forecast fine-grained AQI in Beijing. An ERT-MLP hybrid model is developed by Eslami et al. (2019) to forecast O_3 concentration in Seoul. Zhu et al. (2020) propose a model which combines a complementary ensemble empirical mode decomposition to pre-process data and statistical machine learning models to forecast AQI in 17 port cities in China. An unsupervised model is proposed by Guo et al. (2020) to predict PM2.5 in Hubei, China. The model corresponds to a hybrid of a GRU and an inverse distance weighted KNN. GRU is also combined with multi-linear regression by Lin et al. (2021) and applied to forecast PM2.5 concentration in Taiwan. Janarthanan et al. (2021) propose a hybridisation of SVR and LSTM to predict AQI in Chennai.

4.4 Papers comparing different models

In this section we review papers that compare different approaches with the goal of showing which one of them outperforms the others under different scenarios. These papers represent 23.5% of all the papers analysed in our work. It is worth pointing out that most of these works focus on DL algorithms because they outperform regression algorithms more than 60% of the cases when both types of algorithms are compared. Figure 17 summarises the number of times that each model is used and when these models show better results than the considered alternatives.

Some works show that LSTM is the most accurate algorithm when it is compared with some deep learning and regression algorithms. Thaweephol and Wiwatwattana (2019) compare a usual LSTM model and a seasonal ARIMA with exogenous regressor for PM2.5 forecasting in the Bangkok urban area. The experiments indicate that LSTM has a low level of error for each of the time steps. Chowdhury et al. (2020) compare several models (DT, RF, SVM, Bagging, MLP or LSTM) to forecast AQI in the city of Dhaka, Bangladesh. Results show that LSTM models present the optimal performance when predicting hourly and daily values. Zhoul et al. (2020) compare a hybrid prophet-LSTM model with a hybrid prophet-SVM model to forecast AQI in Nanjing from 2014 to 2019. The results of the performed experiments show that hybrid prophet-LSTM presents the best performance and the highest accuracy. Cheng and Peng (2021) develop a hybrid model which combines an LSTM network with a fuzzy algorithm to forecast AQI in Taiwan.

Other publications conclude that MLP is the most accurate algorithm. Five predictive models (SVR, Naive-Bayes, RF, MLP and KNN) were compared by Amado and Dela Cruz (2018) to forecast AQI in China. All the models obtained an accuracy higher than 94%, corresponding the highest one to the MLP model. Abdullah et al. (2019) compare an MLP model with an MLR to forecast PM10 in Kuala Lumpur, Malaysia. Results show that the highest R^2 is obtained when using MLP. Several regressor models (MLR, KNN, RF, DT or MLP) were compared by Maheshwari and Lamba (2019) to predict PM2.5 in Beijing. Although the performance of all the models was comparable, MLP achieved the highest true positive rate with an accuracy value of 95.4. Bouakline et al. (2020) compare three models (SVR, MLP and RF) to forecast PM10 in Casablanca, Morocco. Again, their results show that MLP has the best accuracy.

In the same way, there are works that compare several models and conclude that RF is the most accurate algorithm. Ochando et al. (2015) compare several algorithms (MLR, KNN, M5P, RF, SVM or MLP) to predict different pollutant concentrations in Valencia, Spain. RF obtained the highest accuracy. Contreras and Ferri (2016) compare five algorithms (MLR, quantile regression, KNN, M5P -a decision tree learner for regression task-and RF) to forecast the level of several pollutants also in Valencia. In this case, RF also reached the highest accuracy. Shawabkeh et al. (2018) compare an MLP algorithm with an SVR model to forecast the concentration of benzene in a city of Italy. MLP obtains the best results both in short-term and long-term cases. Ameer et al. (2019) make a comparison between four models (RF, DT, MLP, Boosting) to predict PM2.5 of several Chinese cities. RF obtains the best accuracy. Li et al. (2019) make a comparison between logistic regression and RF to forecast AQI in California. The best performance was exhibited by the RF model. Srikamdee and Onpans (2019) compare three different methods (MLR, MLP and genetic programming) to forecast AQI in several areas of Thailand. Somehow surprisingly, the MLR and the genetic programming approach obtain the best performance in clean and unhealthy air quality situations. Kuo et al. (2019) develop a comparison to predict AQI in Taipei during one year. This paper compare an MLP, an RNN and an RNN with a Gaussian process. The last one obtains the highest accuracy. Mahanta et al. (2019) compare the accuracy of some prediction models (original MLR, two MLR variants, MLP, RF, extra trees, boosting and KNN) to forecast AQI in New Delhi. The results show that the extra trees regression model obtains the highest accuracy. Pasupuleti et al. (2020) consider three algorithms (RF, DT, MLR) to forecast the concentration of several pollutants in Spain. Various regressor methods (MLP, RF, DT and SVR) are compared by Kaur Bamrah et al. (2020) for predicting AQI in India by adopting terrain features. In these last two papers, RF obtains the highest accuracy. Zheng et al. (2020) compare several regression and DL models (ARIMA, LSTM, SVR, RF, MLR or several boosting models) to forecast AQI in

Beijing and Hong Kong. Results conclude that ensemble models outperform the accuracy of individual models. Eight regression models are compared by Hota et al. (2021) to forecast AQI in Indian cities during the COVID-19 crisis. Results show that boosting models obtain the best results. Yarragunta et al. (2021) compare six regression algorithms (DT, KNN, SVR, MLR, RF and Naive-Bayes) to predict AQI in Delhi. In this case, the highest accuracy is obtained by the DT algorithm. Chakradhar Reddy et al. (2021) develop a comparison between six supervised ML models (LR, RF, DT, SVR, KNN and Naive-Bayes) to forecast AQI in New Delhi. Results show that DT obtains the higher accuracy (very close to 100%).

There are several papers where SVM obtains the best results when it is compared with other algorithms. In the comparison developed by Ganesh et al. (2017) between four regression models to forecast AQI in India and USA, SVM with Gaussian kernel gets a maximum R^2 value. The algorithms are evaluated in two different cities, which provides a high consistency. Srivastava et al. (2018) compare eight ML models and the best R^2 value is obtained by the SVR and MLP models. However, the results are substantially worse than the ones of the previous paper. Mahalingam et al. (2019) compare an SVR with an MLP, getting the former the best accuracy (97.3%). Gu et al. (2020) compare an SVR algorithm that has been optimised by using PSO with an SVR algorithm that has been combined with an improved PSO algorithm to forecast AQI in Shenzhen. Results show that the improved algorithm requires less execution time and has higher accuracy. Sunori et al. (2021) compare a regression model (DT) and a DL model (general regression neural network) to forecast PM10 in Uttarakhand, India. The experiments show that the regression model has better results. Verma et al. (2021) also compare a ML model (ARIMA) and a DL model (LSTM) to forecast AQI in Delhi. In this case, LSTM outperforms ARIMA in three different metrics.

Several studies show that hybrid models perform usually better than single models. Metia et al. (2016) compare the performance of an extended fractional Kalman filter and an extended Kalman filter to predict two pollutant emissions in Sidney. The results show that the first model obtains the lowest error. An Elman RNN, a Jordan RNN and a hybrid Elman-Jordan RNN are compared by Septiawan and Endah (2018) to forecast several pollutant concentrations in London. The results show that Jordan RNN obtains the lowest RMSE. Ali Shah et al. (2019) compare a phase-space reconstruction algorithm combined with an RF, with an SVM and with an MLP neural network to forecast air quality in Masfalah, Saudi Arabia. The combination with the MLP network obtains the best results. Four prediction models [ARIMA, principal component regression (PCR), ARIMA-PCR hybridisation and combination of ARIMA and Gene Expression Programming (GCP)] are compared by Shishegaran et al. (2020) to forecast daily air quality in Tehran. The ARIMA-GCP model obtains the lowest error by far. Ozturk (2021) compares a hybrid model which combines RF, SVR and radial basis function regressor with an LSTM and a GRU to forecast CO and NO_2 in a provided dataset. Results show that the hybrid model obtains the lowest error. To forecast AQI in Beijing, Yan et al. (2021) compare four models (CNN, LSTM, hybrid CNN-LSTM and an spatio-temporal model). The overall forecasting based on the LSTM is considered as the optimal model for all stations.

We conclude this section with those algorithms that have been included in a lower percentage of papers where different approaches are compared. Campalani et al. (2011) compare three co-kriging models (2-variate ordinary co-kriging, 2-variate universal co-kriging and 4-variate ordinary co-kriging) to estimate PM10 concentration in Emilia Romagna, Italy. Their results show that the 2-variate ordinary co-kriging obtained the lowest error. This is the only analysed work which employs co-kriging methods to forecast.

Asadollahfardi et al. (2016) make a comparison between two different neural networks (MLP and RBNN) to forecast *PM*2.5 in Karaj, Iran. In this case, RBNN achieves the maximum R^2 value. MLR has showed the best results in Karatzas et al. (2018), where the authors compare an MLP, an RF and an MLR to *PM*10 in Gdansk, Poland, and Athens. Multi-linear regression obtained the lowest level of error. Radial-basis neural networks show the best performance in two works. Yadav and Nath (2018) compare a RBNN and a generalised regression neural network to forecast *PM*10 in India. The obtained results show that RBNN minimises the error.

5 Conclusions

The aim of this paper is to give a survey including the approaches presented on air quality forecasting since 2011. In order to select papers, we searched a previously defined query in the most widely used scientific databases: *ACM Library*, *Elsevier Online Library*, *IEEE Xplore*, *Springer-Link* and *Wiley Online Library*. After removing papers not related to the main topic of the survey, 155 publications were selected. Then, several characteristics of the papers were extracted and analysed.

In geographic terms, the analysed papers show a direct correlation between the most polluted and the most studied countries.

In temporal terms, there is a clearly increasing trend in the number of papers published per year. This is mainly due to the popularisation of ML in recent times and to the increase of awareness about problem of pollution. However, this trend changes in 2021. We have evidence to think that this change is due to the fact that many researchers are using COVID-19 data to validate their models.

Concerning the studied pollutant measures, our analysis shows that AQI is used in approximately half of the studied papers. Concerning pollutant concentrations, *PM*2.5 is the most predicted, in a total of 54 papers. The hazard of this pollutant could be the main reason of that.

We have observed that pollutants features are the type most widely used (around 50% of the papers). Weather variables are also used very often (around 40% of the papers).

Focusing on ML techniques, the results show that DL algorithms are most popular than regression algorithms. Furthermore, some hybrid algorithms include both types of algorithms in the same model. The DL algorithms most used are LSTM and MLP (27.6% and 26.9% of the total respectively). Other DL algorithms less frequently used are CNN, RNN, GRU and auto-encoders. Regression algorithms most used are SVR (14.7% of the total) and RF (14.1% of the papers). Other regression algorithms less frequently used are DT, ARIMA, KNN and Boosting.

Finally, we would like to look beyond the current work. We can mention recent work that give us an idea of the new trends in the application of Machine Learning to forecast air quality. A comprehensive survey on the correlation between air quality and climate change to develop techniques and models that enhance early warning mechanisms and support an effective response to climate-change-induced air pollution, thereby fostering sustainable cities and societies, has been recently presented Balogun et al. (2021). Concerning trends on new algorithms, we must mention the use of Deep Transformer Networks. Although they were initially developed to solve natural language processing tasks, their use has been extended in recent years to other fields such as time series and, by extension, to air quality forecasting. Transformers have been recently used to forecast *PM*2.5 concentration in two

Chinese cities (Zhang et al. 2022) and ozone concentration in Madrid, Spain Méndez et al. (2022). Another interesting family of models which are recently increasing their popularity in air quality forecasting are Graph Neural Networks. The main characteristic of this type of networks is the information obtained from the dynamic interaction between the neighbours (e.g. different cities, different neighbourhoods, different streets) that are weighted depending on the distance. In this line, a recent work Li et al. (2021) has applied a Graph Neural Network in three real-world datasets to predict *PM*2.5 concentration, obtaining better performance than baseline models. We can also mention the recent application of Temporal Convolutional Networks (TCNs) to air quality prediction, in particular, to predict *PM*2.5 concentrations. In this line, a multi-output TCN was used to predict *PM*2.5 and *PM*10 concentrations in multiple sites Samal et al. (2021), a multi-directional TCN was proposed to forecast *PM*2.5 in a dataset with missing values Samal et al. (2021), and a Gaussian-TCN model was recently used to predict *CO* concentration Ni et al. (2023). Finally, it is worth to mention the recent application of Complex Event Processing (CEP) Corral-Plaza et al. (2021) to analyse and predict air quality Semlali et al. (2021).

References

Abdullah S, Ismail M, Ahmed AN, Mansor WNW (2019) Big data analytics and artificial intelligence in air pollution studies for the prediction of particulate matter concentration. In: Proceedings of the 3rd International Conference on Telecommunications and Communication Engineering. ICTCE '19, pp 90–94. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/33695 55.3369557

Abirami S, Chitra P (2021) Regional air quality forecasting using spatiotemporal deep learning. J Clean Prod 283:125341. https://doi.org/10.1016/j.jclepro.2020.125341

Agarwal S, Sharma S, Suresh R, Rahman MH, Vranckx S, Maiheu B, Blyth L, Janssen S, Gargava P, Shukla VK, Batra S (2020) Air quality forecasting using artificial neural networks with real time dynamic error correction in highly polluted regions. Sci Total Environ 735:139454. https://doi.org/10.1016/j.scitotenv.2020.139454

Alhirmizy S, Qader B (2019) Multivariate time series forecasting with LSTM for madrid, spain pollution. In: 2019 International Conference on Computing and Information Science and Technology and Their Applications (ICCISTA), pp 1–5 (2019). https://doi.org/10.1109/ICCISTA.2019.8830667

Ali Shah SA, Aziz W, Ahmed Nadeem MS, Almaraashi M, Shim S-O, Habeebullah TM, Mateos C (2019) A novel Phase Space Reconstruction- (PSR-) based predictive algorithm to forecast atmospheric particulate matter concentration. Sci Program. https://doi.org/10.1155/2019/6780379

Alsaedi AS, Liyakathunisa L (2019) Spatial and temporal data analysis with deep learning for air quality prediction. In: 2019 12th International Conference on Developments in eSystems Engineering (DeSE), pp 581–587 . https://doi.org/10.1109/DeSE.2019.00111

Amado TM, Dela Cruz JC (2018) Development of machine learning-based predictive models for air quality monitoring and characterization. In: TENCON 2018—2018 IEEE Region 10 Conference, pp 0668–0672. https://doi.org/10.1109/TENCON.2018.8650518

Ameer S, Shah MA, Khan A, Song H, Maple C, Islam SU, Asghar MN (2019) Comparative analysis of machine learning techniques for predicting air quality in smart cities. IEEE Access 7:128325–128338. https://doi.org/10.1109/ACCESS.2019.2925082

Ao D, Cui Z, Gu D (2019) Hybrid model of air quality prediction using k-means clustering and deep neural network. In: 2019 Chinese Control Conference (CCC), pp 8416–8421. https://doi.org/10.23919/ChiCC.2019.8865861

Asadollahfardi G, Madinejad M, Aria SH, Motamadi V (2016) Predicting particulate matter (PM2.5) concentrations in the air of Shahr-e Ray City, Iran, by using an Artificial Neural Network. Environmental Quality Management 25(4), 71–83. https://doi.org/10.1002/tqem.21464

Balogun A-L, Tella A (2022) Modelling and investigating the impacts of climatic variables on ozone concentration in Malaysia using correlation analysis with random forest, decision tree regression, linear regression, and support vector regression. Chemosphere 299:134250. https://doi.org/10.1016/j.chemosphere.2022.134250

Balogun A-L, Tella A, Baloo L, Adebisi N (2021) A review of the inter-correlation of climate change, air pollution and urban sustainability using novel machine learning algorithms and spatial information science. Urban Climate 40:100989. https://doi.org/10.1016/j.uclim.2021.100989

Barrón-Adame JM, Ibarra-Manzano OG, Vega-Corona A, Cortina-Januchs MG, Andina D (2012) Air pollution data classification by SOM Neural Network. In: World Automation Congress 2012, pp 1–5. https://ieeexplore.ieee.org/abstract/document/6320993

Barthwal A, Acharya D (2018) An internet of things system for sensing, analysis forecasting urban air quality. In: 2018 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), pp 1–6 . https://doi.org/10.1109/CONECCT.2018.8482397

Bharat Deshmukh S, Prakash Shirsat K, Prashant Dhotre S, Ravsaheb Jejurkar P (2021) A survey on machine learning-based prediction of air quality index. Int J Adv Res Innov Ideas Educ 7:1203–1207

Bhatti UA, Yan Y, Zhou M, Ali S, Hussain A, Qingsong H, Yu Z, Yuan L (2021) Time series analysis and forecasting of air pollution particulate matter ($PM_2.5$): an sarima and factor analysis approach. IEEE Access 9:41019–41031. https://doi.org/10.1109/ACCESS.2021.3060744

Bouakline O, ARJDAL K, KHOMSI K, SEMANE N, ELIDRISSI A, NAFIRI S, NAJMI H (2020) Prediction of daily PM10 concentration using machine learning. In: 2020 IEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS), pp 1–5. https://doi.org/10.1109/ICECOCS50124.2020.9314380

Bougoudis I, Demertzis K, Iliadis L (2016) HISYCOL a hybrid computational intelligence system for combined machine learning: The case of air pollution modeling in Athens. Neural Comput Appl 27(5):1191–1206. https://doi.org/10.1007/s00521-015-1927-7

Campalani P, Nguyen TNT, Mantovani S, Mazzini G (2011) On the automatic prediction of PM10 with in-situ measurements, satellite aot retrievals and ancillary data. In: 2011 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), pp 093–098. https://doi.org/10.1109/ISSPIT.2011.6151541

Chakradhar Reddy K, Nagarjuna Reddy K, Brahmaji Prasad K, Selvi Rajendran P (2021) The prediction of quality of the air using supervised learning. In: 2021 6th International Conference on Communication and Electronics Systems (ICCES), pp 1–5. https://doi.org/10.1109/ICCES51350.2021.9488983

Chen Y, Li F, Deng Z, Chen X, He J (2017) PM2.5 forecasting with hybrid LSE model-based approach. Softw Pract Exp 47(3):379–390. https://doi.org/10.1002/spe.2413

Cheng J-C, Peng H-C (2021) Air quality forecast and evaluation based on long short-term memory network and fuzzy algorithm. In: 2021 IEEE 4th International Conference on Knowledge Innovation and Invention (ICKII), pp 88–91. https://doi.org/10.1109/ICKII51822.2021.9574707

Chiang P-W, Horng S-J (2021) Hybrid time-series framework for daily-based PM2.5 forecasting. IEEE Access 9:104162–104176. https://doi.org/10.1109/ACCESS.2021.3099111

Choi J, Kim J, Jung K (2021) Air quality prediction with 1-dimensional convolution and attention on multimodal features. In: 2021 IEEE International Conference on Big Data and Smart Computing (BigComp), pp 196–202. https://doi.org/10.1109/BigComp51126.2021.00045

Chowdhury A-S, Uddin MS, Tanjim MR, Noor F, Rahman RM (2020) Application of data mining techniques on air pollution of Dhaka city. In: 2020 IEEE 10th International Conference on Intelligent Systems (IS), pp 562–567. https://doi.org/10.1109/IS48319.2020.9200125

Colchado LE, Villanueva E, Ochoa-Luna J (2021) A neural network architecture with an attention-based layer for spatial prediction of fine particulate matter. In: 2021 IEEE 8th International Conference on Data Science and Advanced Analytics (DSAA), pp 1–10. https://doi.org/10.1109/DSAA53316.2021.9564200

Contreras L, Ferri C (2016) Wind-sensitive interpolation of urban air pollution forecasts. Procedia Comput Sci 80(C):313–323. https://doi.org/10.1016/j.procs.2016.05.343

Corral-Plaza D, Ortiz G, Medina-Bulo I, Boubeta-Puig J (2021) MEdit4CEP-SP: a model-driven solution to improve decision-making through user-friendly management and real-time processing of heterogeneous data streams. Knowl Based Syst 213:106682. https://doi.org/10.1016/j.knosys.2020.106682

De Vito S, Delli Veneri P, Esposito E, Salvato M, Bright V, Jones RL, Popoola O (2015) Dynamic multivariate regression for on-field calibration of high speed air quality chemical multi-sensor systems. In: 2015 XVIII AISEM Annual Conference, pp 1–3. https://doi.org/10.1109/AISEM.2015.7066794

Dedovic MM, Avdakovic S, Turkovic I, Dautbasic N, Konjic T (2016) Forecasting PM10 concentrations using neural networks and system for improving air quality. In: 2016 XI International Symposium on Telecommunications (BIHTEL), pp 1–6. https://doi.org/10.1109/BIHTEL.2016.7775721

Dua R, Madaan D, Mukherjee P, Lall B (2019) Real time attention based bidirectional long short-term memory networks for air pollution forecasting. In: 2019 IEEE Fifth International Conference on Big Data Computing Service and Applications (BigDataService), pp 151–158. https://doi.org/10.1109/BigDataService.2019.00027

Durić M, Vujović D (2020) Short-term forecasting of air pollution index in Belgrade, Serbia. Meteorol Appl 27(5):1946. https://doi.org/10.1002/met.1946

Eslami E, Salman AK, Choi Y, Sayeed A, Lops Y (2019) A data ensemble approach for real-time air quality forecasting using extremely randomized trees and deep neural networks. Neural Comput Appl 32:7563–7579. https://doi.org/10.1007/s00521-019-04287-6

Fu M, Wang W, Le Z, Khorram MS (2015) Prediction of particular matter concentrations by developed feed-forward neural network with rolling mechanism and gray model. Neural Comput Appl 26(8):1789–1797. https://doi.org/10.1007/s00521-015-1853-8

Ganesh SS, Modali SH, Palreddy SR, Arulmozhivarman P (2017) Forecasting air quality index using regression models: A case study on delhi and houston. In: 2017 International Conference on Trends in Electronics and Informatics (ICEI), pp 248–254. https://doi.org/10.1109/ICOEI.2017.8300926

Ganesh SS, Reddy NB, Arulmozhivarman P (2017) Forecasting air quality index based on mamdani fuzzy inference system. In: 2017 International Conference on Trends in Electronics and Informatics (ICEI), pp 338–341. https://doi.org/10.1109/ICOEI.2017.8300944

García Nieto PJ, Combarro EF, Del Coz Díaz JJ, Montañés E (2013) A svm-based regression model to study the air quality at local scale in Oviedo urban area (Northern Spain): a case study. Appl Math Comput 219(17):8923–8937. https://doi.org/10.1016/j.amc.2013.03.018

Gocheva-Ilieva SG, Ivanov AV, Livieris IE (2020) High performance machine learning models of large scale air pollution data in urban area. Cybern Inf Technol 20(6):49–60. https://doi.org/10.2478/cait-2020-0060

Gu K, Zhou Y, Sun H, Zhao L, Liu S (2020) Prediction of air quality in Shenzhen based on neural network algorithm. Neural Comput Appl. https://doi.org/10.1007/s00521-019-04492-3

Guo C, Liu G, Lyu L, Chen C-H (2020) An unsupervised PM2.5 estimation method with different spatio-temporal resolutions based on kidw-tcgru. IEEE Access 8:190263–190276. https://doi.org/10.1109/ACCESS.2020.3032420

Hajmohammadi H, Heydecker B (2021) Multivariate time series modelling for urban air quality. Urban Clim 37:100834. https://doi.org/10.1016/j.uclim.2021.100834

Han Y, Lam JCK, Li VO, Zhang Q (2020) A domain-specific Bayesian deep-learning approach for air pollution forecast. IEEE Trans Big Data. https://doi.org/10.1109/TBDATA.2020.3005368

Han Y, Li VOK, Lam JCK, Pollitt M (2021) How BLUE is the sky? Estimating air qualities in Beijing during the Blue Sky Day period (2008–2012) by Bayesian Multi-task LSTM. Environ Sci Policy 116:69–77. https://doi.org/10.1016/j.envsci.2020.10.015

Hoi KI, Yuen KV, Mok KM (2013) Improvement of the multilayer perceptron for air quality modelling through an adaptive learning scheme. Comput Geosci 59:148–155. https://doi.org/10.1016/j.cageo.2013.06.002

Hota L, Dash PK, Sahoo KS, Gandomi AH (2021) Air quality index analysis of indian cities during covid-19 using machine learning models: a comparative study. In: 2021 8th International Conference on Soft Computing Machine Intelligence (ISCMI), pp 27–31. https://doi.org/10.1109/ISCMI53840.2021.9654925

Huang Y, Xiang Y, Zhao R, Cheng Z (2020) Air quality prediction using improved pso-bp neural network. IEEE Access 8:99346–99353. https://doi.org/10.1109/ACCESS.2020.2998145

Huang M, Zhang T, Wang J, Zhu L (2015) A new air quality forecasting model using data mining and artificial neural network. In: 2015 6th IEEE International Conference on Software Engineering and Service Science (ICSESS), pp 259–262. https://doi.org/10.1109/ICSESS.2015.7339050

Hu J, Zhang X, Liang D, Zhong F (2021) A novel linear interpolation algorithm based on spatial distribution for improvement temporal resolution of $pm_{2.5}$ and PM_{10}. In: 2021 7th International Conference on Big Data and Information Analytics (BigDIA), pp.60–66. https://doi.org/10.1109/BigDIA53151.2021.9619651

Iskandaryan D, Ramos F, Trilles S (2020) Air quality prediction in smart cities using machine learning technologies based on sensor data: a review. Appl Sci. https://doi.org/10.3390/app10072401

Janarthanan R, Partheeban P, Somasundaram K, Navin Elamparithi P (2021) A deep learning approach for prediction of air quality index in a metropolitan city. Sustain Cities Soc 67:102720. https://doi.org/10.1016/j.scs.2021.102720

Jato-Espino D, Castillo-Lopez E, Rodriguez-Hernandez J, Ballester-Muñoz F (2018) Air quality modelling in Catalonia from a combination of solar radiation, surface reflectance and elevation. Sci Total Environ 624:189–200. https://doi.org/10.1016/j.scitotenv.2017.12.139

Jiang J-Y, Li C-T (2016) Forecasting geo-sensor data with participatory sensing based on dropout neural network. In: Proceedings of the 25th ACM International on Conference on Information and Knowledge Management. CIKM '16, pp 2033–2036. Association for Computing Machinery, New York, NY, USA . https://doi.org/10.1145/2983323.2983902

Jiao Y, Wang Z, Zhang Y (2019) Prediction of air quality index based on LSTM. In: 2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC), pp 17–20. https://doi.org/10.1109/ITAIC.2019.8785602

Jin N, Zeng Y, Yan K, Ji Z (2021) Multivariate air quality forecasting with nested long short term memory neural network. IEEE Trans Ind Inf 17(12):8514–8522. https://doi.org/10.1109/TII.2021.3065425

Jovova L, Trivodaliev K (2021) Air pollution forecasting using CNN-LSTM deep learning model. In: 2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO), pp 1091–1096. https://doi.org/10.23919/MIPRO52101.2021.9596860

Kang Z, Qu Z (2017) Application of bp neural network optimized by genetic simulated annealing algorithm to prediction of air quality index in lanzhou. In: 2017 2nd IEEE International Conference on Computational Intelligence and Applications (ICCIA), pp 155–160. https://doi.org/10.1109/CIAPP.2017.8167199

Karatzas K, Katsifarakis N, Orlowski C, Sarzyński A (2018) Revisiting urban air quality forecasting: a regression approach. Vietnam J Comput Sci 5(2):177–184. https://doi.org/10.1007/s40595-018-0113-0

Kaur Bamrah S, R SK, S GK (2020) Application of random forests for air quality estimation in india by adopting terrain features. In: 2020 4th International Conference on Computer, Communication and Signal Processing (ICCCSP), pp 1–6. https://doi.org/10.1109/ICCCSP49186.2020.9315252

Kiftiyani U, Nazhifah SA (2021) Deep learning models for air pollution forecasting in Seoul South Korea. In: 2021 International Conference on Software Engineering Computer Systems and 4th International Conference on Computational Science and Information Management (ICSECS-ICOCSIM), pp 547–551. https://doi.org/10.1109/ICSECS52883.2021.00106

Kök İ, Şimşek MU, Özdemir S (2017) A deep learning model for air quality prediction in smart cities. In: 2017 IEEE International Conference on Big Data (Big Data), pp 1983–1990 . https://doi.org/10.1109/BigData.2017.8258144

Kuo RJ, Prasetyo B, Wibowo BS (2019) Deep learning-based approach for air quality forecasting by using recurrent neural network with gaussian process in taiwan. In: 2019 IEEE 6th International Conference on Industrial Engineering and Applications (ICIEA), pp 471–474. https://doi.org/10.1109/IEA.2019.8715113

Lan Y, Dai Y (2020) Urban air quality prediction based on space-time optimization LSTM model. In: 2020 3rd International Conference on Artificial Intelligence and Big Data (ICAIBD), pp 215–222 . https://doi.org/10.1109/ICAIBD49809.2020.9137441

Lang PE, Carslaw DC, Moller SJ (2019) A trend analysis approach for air quality network data. Atmos Environ X 2:100030. https://doi.org/10.1016/j.aeaoa.2019.100030

Le, V.-D., Bui, T.-C., Cha, S.-K.: Spatiotemporal deep learning model for citywide air pollution interpolation and prediction. In: 2020 IEEE International Conference on Big Data and Smart Computing (BigComp), pp 55–62 (2020). https://doi.org/10.1109/BigComp48618.2020.00-99

Li J, Shao X, Sun R, Visioli A (2019) A dbn-based deep neural network model with multitask learning for online air quality prediction. J Control Sci Eng. https://doi.org/10.1155/2019/5304535

Li H, Wang J, Yang H (2020) A novel dynamic ensemble air quality index forecasting system. Atmos Pollut Res 11(8):1258–1270. https://doi.org/10.1016/j.apr.2020.04.010

Li T, Hua M, Wu X (2020) A hybrid CNN-LSTM model for forecasting particulate matter (PM2.5). IEEE Access 8:26933–26940. https://doi.org/10.1109/ACCESS.2020.2971348

Li L, Li Z, Reichmann L, Woodbridge D (2019) A scalable and reliable model for real-time air quality prediction. In: 2019 IEEE SmartWorld, Ubiquitous Intelligence Computing, Advanced Trusted Computing, Scalable Computing Communications, Cloud Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), pp 51–57. https://doi.org/10.1109/SmartWorld-UIC-ATC-SCALCOM-IOP-SCI.2019.00053

Lim YB, Aliyu I, Lim CG (2019) Air pollution matter prediction using recurrent neural networks with sequential data. In: Proceedings of the 2019 3rd International Conference on Intelligent Systems, Metaheuristics & Swarm Intelligence. ISMSI 2019, pp 40–44. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3325773.3325788

Lin Y-C, Lee S-J, Ouyang C-S, Wu C-H (2020) Air quality prediction by neuro-fuzzy modeling approach. Appl Soft Comput 86:105898. https://doi.org/10.1016/j.asoc.2019.105898

Lin C-Y, Chang Y-S, Abimannan S (2021) Ensemble multifeatured deep learning models for air quality forecasting. Atmos Pollut Res 12(5):101045. https://doi.org/10.1016/j.apr.2021.03.008

Lin Y, Mago N, Gao Y, Li Y, Chiang Y-Y, Shahabi C, Ambite JL (2018) Exploiting spatiotemporal patterns for accurate air quality forecasting using deep learning. In: Proceedings of the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. SIGSPATIAL '18, pp 359–368. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3274895.3274907

Li J, Shao X, Zhao H (2018) An online method based on random forest for air pollutant concentration forecasting. In: 2018 37th Chinese Control Conference (CCC), pp 9641–9648 . https://doi.org/10.23919/ChiCC.2018.8483621

Li Y, Shen X, Han D, Sun J, Shen Y (2018) Spatio-temporal-aware sparse denoising autoencoder neural network for air quality prediction. In: 2018 5th IEEE International Conference on Cloud Computing and Intelligence Systems (CCIS), pp 96–100. https://doi.org/10.1109/CCIS.2018.8691167

Liu T, Lau AKH, Sandbrink K, Fung JCH (2018) Time series forecasting of air quality based on regional numerical modeling in Hong Kong. J Geophys Res 123(8):4175–4196. https://doi.org/10.1002/2017JD028052

Liu B, Yan S, Li J, Qu G, Li Y, Lang J, Gu R (2019) A sequence-to-sequence air quality predictor based on the n-step recurrent prediction. IEEE Access 7:43331–43345. https://doi.org/10.1109/ACCESS.2019.2908081

Liu B, Yan S, Li J, Li Y, Lang J, Qu G (2021) A spatiotempor.al recurrent neural network for prediction of atmospheric PM2.5: a case study of Beijing. IEEE Trans Soc Syst 8(3):578–588. https://doi.org/10.1109/TCSS.2021.3056410

Liu B, Yan S, Li J, Li Y (2016) Forecasting PM2.5 concentration using spatio-temporal extreme learning machine. In: 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA), pp 950–953. https://doi.org/10.1109/ICMLA.2016.0171

Liu W, Zhang H, Liu Q (2019) An air quality grade forecasting approach based on ensemble learning. In: 2019 International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM), pp 87–91. https://doi.org/10.1109/AIAM48774.2019.00024

Li W, Yi L, Yin X (2020) Real time air monitoring, analysis and prediction system based on internet of things and LSTM. In: 2020 International Conference on Wireless Communications and Signal Processing (WCSP), pp 188–194. https://doi.org/10.1109/WCSP49889.2020.9299738

Li D, Yu H, Geng Y-a, Li X, Li Q (2021) DDGNet: A dual-stage dynamic spatio-temporal graph network for PM2.5 forecasting. In: 2021 IEEE International Conference on Big Data (Big Data), pp 1679–1685. https://doi.org/10.1109/BigData52589.2021.9671941

Ma J, Cheng JCP, Lin C, Tan Y, Zhang J (2019) Improving air quality prediction accuracy at larger temporal resolutions using deep learning and transfer learning techniques. Atmos Environ 214:116885. https://doi.org/10.1016/j.atmosenv.2019.116885

Ma J, Ding Y, Gan VJL, Lin C, Wan Z (2019) Spatiotemporal prediction of PM2.5 concentrations at different time granularities using IDW-BLSTM. IEEE Access 7:107897–107907. https://doi.org/10.1109/ACCESS.2019.2932445

Ma J, Li Z, Cheng JCP, Ding Y, Lin C, Xu Z (2020) Air quality prediction at new stations using spatially transferred bi-directional long short-term memory network. Sci Total Environ 705:135771. https://doi.org/10.1016/j.scitotenv.2019.135771

Madaan D, Dua R, Mukherjee P, Lall B (2019) Vayuanukulani: Adaptive memory networks for air pollution forecasting. In: 2019 IEEE Global Conference on Signal and Information Processing (GlobalSIP), pp 1–5. https://doi.org/10.1109/GlobalSIP45357.2019.8969343

Ma L, Gao Y, Zhao C (2020) Research on machine learning prediction of air quality index based on spss. In: 2020 International Conference on Computer Network, Electronic and Automation (ICCNEA), pp 1–5. https://doi.org/10.1109/ICCNEA50255.2020.00011

Mahalingam U, Elangovan K, Dobhal H, Valliappa C, Shrestha S, Kedam G (2019) A machine learning model for air quality prediction for smart cities. In: 2019 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET), pp 452–457. https://doi.org/10.1109/WiSPNET45539.2019.9032734

Mahanta S, Ramakrishnudu T, Jha RR, Tailor N (2019) Urban air quality prediction using regression analysis. In: TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON), pp 1118–1123. https://doi.org/10.1109/TENCON.2019.8929517

Maheshwari K, Lamba S (2019) Air quality prediction using supervised regression model. In: 2019 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT), vol. 1, pp 1–7. https://doi.org/10.1109/ICICT46931.2019.8977694

Mao W, Wang W, Jiao L, Zhao S, Liu A (2021) Modeling air quality prediction using a deep learning approach: method optimization and evaluation. Sustain Cit Soc 65:102567. https://doi.org/10.1016/j.scs.2020.102567

Matović K, Vlahović N (2021) Air quality prediction in smart city. In: 2021 15th International Conference on Advanced Technologies, Systems and Services in Telecommunications (TELSIKS), pp 287–290 . https://doi.org/10.1109/TELSIKS52058.2021.9606405

Mei G, Zhang H, Zhang B (2016) Improving elman neural network model via fusion of new feedback mechanism and genetic algorithm. In: 2016 International Conference on Progress in Informatics and Computing (PIC), pp 69–73. https://doi.org/10.1109/PIC.2016.7949469

Méndez M, Montero C, Núñez M (2022) Using deep transformer based models to predict ozone levels. In: Intelligent Information and Database Systems—14th Asian Conference, ACIIDS 2022, Part I. Lecture Notes in Computer Science, vol 13757, pp 169–182. Springer. https://doi.org/10.1007/978-3-031-21743-2_14

Metia S, Oduro SD, Duc HN, Ha Q (2016) Inverse air-pollutant emission and prediction using extended fractional Kalman filtering. IEEE J Sel Top Appl Earth Observ Remote Sens 9(5):2051–2063. https://doi.org/10.1109/JSTARS.2016.2541958

Metia S, Oduro SD, Ha QP, Due H (2014) Air pollution prediction using matérn function based extended fractional Kalman filtering. In: 2014 13th International Conference on Control Automation Robotics Vision (ICARCV), pp 758–763. https://doi.org/10.1109/ICARCV.2014.7064399

Mu B, Li S, Yuan S (2017) An improved effective approach for urban air quality forecast. In: 2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), pp 935–942. https://doi.org/10.1109/FSKD.2017.8393403

Navares R, Aznarte JL (2020) Predicting air quality with deep learning LSTM: towards comprehensive models. Eco Inform 55:101019. https://doi.org/10.1016/j.ecoinf.2019.101019

Nemade S (2019) A survey on different machine learning techniques for air quality forecasting for urban air pollution. Int J Res Appl Sci Eng Technol 7:2185–2194. https://doi.org/10.22214/ijraset.2019.4395

Ngom B, Diallo M, Seyc MR, Drame MS, Cambier C, Marilleau N (2021) PM10 data assimilation on real-time agent-based simulation using machine learning models: case of dakar urban air pollution study. In: 2021 IEEE/ACM 25th International Symposium on Distributed Simulation and Real Time Applications (DS-RT), pp 1–4. https://doi.org/10.1109/DS-RT52167.2021.9576143

Nguyen MH, Le Nguyen P, Nguyen K, Le VA, Nguyen T-H, Ji Y (2021) PM2.5 prediction using genetic algorithm-based feature selection and encoder-decoder model. IEEE Access 9:57338–57350. https://doi.org/10.1109/ACCESS.2021.3072280

Ni S, Jia P, Xu Y, Zeng L, Li X, Xu M (2023) Prediction of CO concentration in different conditions based on gaussian-TCN. Sens Actuators B 376:133010. https://doi.org/10.1016/j.snb.2022.133010

Niharika Venkatadri M, Rao P (2014) A survey on air quality forecasting techniques. Int J Comput Sci Inf Technol 5(1):103–107. https://doi.org/10.15680/IJIRSET.2019.0805015

Ochando LC, Julián CIF, Ochando FC, Ferri C (2015) Airvlc: An application for real-time forecasting urban air pollution. In: Proceedings of the 2nd International Conference on Mining Urban Data—Volume 1392. MUD'15, pp 72–79. CEUR-WS.org, Aachen, DEU

Ong BT, Sugiura K, Zettsu K (2016) Dynamically pre-trained deep recurrent neural networks using environmental monitoring data for predicting PM2.5. Neural Comput Appl 27(6):1553–1566. https://doi.org/10.1007/s00521-015-1955-3

Ouyang X, Yang Y, Zhang Y, Zhou W (2021) Spatial-temporal dynamic graph convolution neural network for air quality prediction. In: 2021 International Joint Conference on Neural Networks (IJCNN), pp 1–8. https://doi.org/10.1109/IJCNN52387.2021.9534167

Ozturk A (2021) Accuracy improvement in air-quality forecasting using regressor combination with missing data imputation. Comput Intell 37(1):226–252. https://doi.org/10.1111/coin.12399

Pai T-Y, Hanaki K, Su H-C, Yu L-F (2013) A 24-h forecast of oxidant concentration in Tokyo using neural network and fuzzy learning approach. Clean Soil Air Water 41(8):729–736. https://doi.org/10.1002/clen.201000067

Paoli C, Notton G, Nivet M-L, Padovani M, Savelli J-L (2011) A neural network model forecasting for prediction of hourly ozone concentration in corsica. In: 2011 10th International Conference on Environment and Electrical Engineering, pp 1–4 . https://doi.org/10.1109/EEEIC.2011.5874661

Pasupuleti VR, Uhasri Kalyan P, Srikanth Reddy HK (2020) Air quality prediction of data log by machine learning. In: 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), pp 1395–1399. https://doi.org/10.1109/ICACCS48705.2020.9074431

Putra KT, Prayitno Cahyadi EF, Mamonto AS, Berutu SS, Muharom S (2021) Forecasting air quality using massive-scale wsn based on convolutional LSTM network. In: 2021 1st International Conference on Electronic and Electrical Engineering and Intelligent System (ICE3IS), pp 1–6. https://doi.org/10.1109/ICE3IS54102.2021.9649763

Qi Z, Wang T, Song G, Hu W, Li X, Zhang Z (2018) Deep air learning: interpolation, prediction, and feature analysis of fine-grained air quality. IEEE Trans Knowl Data Eng 30(12):2285–2297. https://doi.org/10.1109/TKDE.2018.2823740

Rahimpour A, Amanollahi J, Tzanis CG (2021) Air quality data series estimation based on machine learning approaches for urban environments. Air Qual Atmos Health 14:191–201. https://doi.org/10.1007/s11869-020-00925-4

Samal KKR, Panda AK, Babu KS, Das SK (2021) Multi-output TCN autoencoder for long-term pollution forecasting for multiple sites. Urban Clim 39:100943. https://doi.org/10.1016/j.uclim.2021.100943

Samal KKR, Babu KS, Das SK (2021) Multi-directional temporal convolutional artificial neural network for PM2.5 forecasting with missing values: a deep learning approach. Urban Clim 36:100800. https://doi.org/10.1016/j.uclim.2021.100800

Schürholz D, Kubler S, Zaslavsky A (2020) Artificial intelligence-enabled context-aware air quality prediction for smart cities. J Clean Prod 271:121941. https://doi.org/10.1016/j.jclepro.2020.121941

Schürholz D, Nurgazy M, Zaslavsky A, Jayaraman PP, Kubler S, Mitra K, Saguna S (2019) Myaqi: Context-aware outdoor air pollution monitoring system. In: Proceedings of the 9th International Conference on the Internet of Things. IoT 2019, pp 1–8. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3365871.3365884D

Semlali BB, Amrani CE, Ortiz G, Boubeta-Puig J, de Prado AG (2021) SAT-CEP-monitor: an air quality monitoring software architecture combining complex event processing with satellite remote sensing. Comput Electr Eng 93:107257. https://doi.org/10.1016/j.compeleceng.2021.107257

Seng D, Zhang Q, Zhang X, Chen G, Chen X (2021) Spatiotemporal prediction of air quality based on LSTM neural network. Alex Eng J 60(2):2021–2032. https://doi.org/10.1016/j.aej.2020.12.009

Septiawan WM, Endah SN (2018) Suitable recurrent neural network for air quality prediction with back-propagation through time. In: 2018 2nd International Conference on Informatics and Computational Sciences (ICICoS), pp 1–6. https://doi.org/10.1109/ICICOS.2018.8621720

Sethi JK, Mittal M (2021) An efficient correlation based adaptive lasso regression method for air quality index prediction. Earth Sci Inform. https://doi.org/10.1007/s12145-021-00618-1

Sharma E, Deo RC, Prasad R, Parisi AV (2020) A hybrid air quality early-warning framework: an hourly forecasting model with online sequential extreme learning machines and empirical mode decomposition algorithms. Sci Total Environ 709:135934. https://doi.org/10.1016/j.scitotenv.2019.135934

Sharma M, Jain S, Mittal S, Sheikh TH (2021) Forecasting and prediction of air pollutants concentrates using machine learning techniques: the case of India. IOP Conf Ser 1022(1):012123. https://doi.org/10.1088/1757-899x/1022/1/012123

Shawabkeh A, Al-Beqain F, Redan A, Salem M (2018) Benzene air pollution monitoring model using ann and svm. In: 2018 Fifth HCT Information Technology Trends (ITT), pp 197–204. https://doi.org/10.1109/CTIT.2018.8649497

Shishegaran A, Saeedi M, Kumar A, Ghiasinejad H (2020) Prediction of air quality in tehran by developing the nonlinear ensemble model. J Clean Prod 259:120825. https://doi.org/10.1016/j.jclepro.2020.120825

Soh P-W, Chang J-W, Huang J-W (2018) Adaptive deep learning-based air quality prediction model using the most relevant spatial-temporal relations. IEEE Access 6:38186–38199. https://doi.org/10.1109/ACCESS.2018.2849820

Sonawani S, Patil K, Chumchu P (2021) NO_2 pollutant concentration forecasting for air quality monitoring by using an optimised deep learning bidirectional GRU model. Int J Comput Sci Eng 24(1):64–73. https://doi.org/10.1504/ijcse.2021.113652

Song X, Huang J, Song D (2019) Air quality prediction based on LSTM-kalman model. In: 2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC), pp 695–699. https://doi.org/10.1109/ITAIC.2019.8785751

Sotomayor-Olmedo A, Aceves-Fernandez MA, Gorrostieta-Hurtado E, Pedraza-Ortega JC, Vargas-Soto JE, Ramos-Arreguin JM, Villaseñor-Carillo U (2011) Evaluating trends of airborne contaminants by using support vector regression techniques. In: CONIELECOMP 2011, 21st International Conference on Electrical Communications and Computers, pp 137–141. https://doi.org/10.1109/CONIELECOMP.2011.5749350

Srikamdee S, Onpans J (2019) Forecasting daily air quality in northern thailand using machine learning techniques. In: 2019 4th International Conference on Information Technology (InCIT), pp 259–263. https://doi.org/10.1109/INCIT.2019.8912072

Srivastava C, Singh S, Singh AP (2018) Estimation of air pollution in delhi using machine learning techniques. In: 2018 International Conference on Computing, Power and Communication Technologies (GUCON), pp 304–309. https://doi.org/10.1109/GUCON.2018.8675022

Sun X, Xu W, Jiang H, Wang Q (2020) A deep multitask learning approach for air quality prediction. Ann Oper Res 303:51–79. https://doi.org/10.1007/s10479-020-03734-1

Sunori SK, Negi PB, Maurya S, Mittal A, Bhakuni AS, Juneja P(2021) Neural network and machine learning based prediction model design for air pollutant PM10. In: 2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA), pp 1128–1133. https://doi.org/10.1109/ICECA52323.2021.9676128

Su Y, Xie H (2020) Prediction of aqi by bp neural network based on genetic algorithm. In: 2020 5th International Conference on Automation, Control and Robotics Engineering (CACRE), pp 625–629. https://doi.org/10.1109/CACRE50138.2020.9230036

Tella A, Balogun A-L, Adebisi N, Abdullah S (2021) Spatial assessment of PM10 hotspots using Random Forest, K-Nearest Neighbour and Naïve Bayes. Atmos Pollut Res 12(10):101202. https://doi.org/10.1016/j.apr.2021.101202

Thanavanich T, Yaibuates M, Suchaya P (2021) Improving the accuracy of forecasting PM2.5 concentrations with hybrid neural network model. In: 2021 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunication Engineering, pp 18–22. https://doi.org/10.1109/ECTIDAMTNCON51128.2021.9425724

Thaweephol K, Wiwatwattana N (2019) Long short-term memory deep neural network model for PM2.5 forecasting in the bangkok urban area. In: 2019 17th International Conference on ICT and Knowledge Engineering (ICT KE), pp 1–6. https://doi.org/10.1109/ICTKE47035.2019.8966854

Verma P, Reddy SV, Ragha L, Datta D (2021) Comparison of time-series forecasting models. In: 2021 International Conference on Intelligent Technologies (CONIT), pp 1–7. https://doi.org/10.1109/CONIT51480.2021.9498451

Wahid H, Ha QP, Duc H, Azzi M (2013) Neural network-based meta-modelling approach for estimating spatial distribution of air pollutant levels. Appl Soft Comput 13(10):4087–4096. https://doi.org/10.1016/j.asoc.2013.05.007

Wang Y, Kong T (2019) Air quality predictive modeling based on an improved decision tree in a weather-smart grid. IEEE Access 7:172892–172901. https://doi.org/10.1109/ACCESS.2019.2956599

Wang J, Song G (2018) A deep spatial-temporal ensemble model for air quality prediction. Neurocomputing 314:198–206. https://doi.org/10.1016/j.neucom.2018.06.049

Wang J, Zhang X, Guo Z, Lu H (2017) Developing an early-warning system for air quality prediction and assessment of cities in China. Expert Syst Appl 84(C):102–116. https://doi.org/10.1016/j.eswa.2017.04.059

Wang J, Li J, Wang X, Wang J, Huang M (2021) Air quality prediction using ct-LSTM. Neural Comput Appl 33:4779–4792. https://doi.org/10.1007/s00521-020-05535-w

Wu Z, Zhang S (2019) Study on the spatial-temporal change characteristics and influence factors of fog and haze pollution based on gam. Neural Comput Appl 31(5):1619–1631. https://doi.org/10.1007/s00521-018-3532-z

Yadav V, Nath S (2018) Daily prediction of PM10 using radial basis function and generalized regression neural network. In: 2018 Recent Advances on Engineering, Technology and Computational Sciences (RAETCS), pp 1–5. https://doi.org/10.1109/RAETCS.2018.8443887

Yan R, Liao J, Yang J, Sun W, Nong M, Li F (2021) Multi-hour and multi-site air quality index forecasting in beijing using CNN, LSTM, CNN-LSTM, and spatiotemporal clustering. Expert Syst Appl 169:114513. https://doi.org/10.1016/j.eswa.2020.114513

Yang R, Hu X, He L (2020) Prediction of shanghai air quality index based on bp neural network optimized by genetic algorithm. In: 2020 13th International Symposium on Computational Intelligence and Design (ISCID), pp. 205–208. https://doi.org/10.1109/ISCID51228.2020.00052

Yan L, Wu Y, Yan L, Zhou M (2018) Encoder-decoder model for forecast of PM2.5 concentration per hour. In: 2018 1st International Cognitive Cities Conference (IC3), pp 45–50 . https://doi.org/10.1109/IC3.2018.00020

Yao L, Lu N, Jiang S (2012) Artificial neural network (ANN) for multi-source PM2.5 estimation using surface, modis, and meteorological data. In: 2012 International Conference on Biomedical Engineering and Biotechnology, pp 1228–1231. https://doi.org/10.1109/iCBEB.2012.81

Yarragunta S, Nabi MA, P J, S R (2021) Prediction of air pollutants using supervised machine learning. In: 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), pp. 1633–1640. https://doi.org/10.1109/ICICCS51141.2021.9432078

Yi X, Duan Z, Li R, Zhang J, Li T, Zheng Y (2020) Predicting fine-grained air quality based on deep neural networks. IEEE Trans Big Data. https://doi.org/10.1109/TBDATA.2020.3047078

Yi H, Xiong Q, Zou Q, Xu R, Wang K, Gao M (2019) A novel random forest and its application on classification of air quality. In: 2019 8th International Congress on Advanced Applied Informatics (IIAI-AAI), pp 35–38. https://doi.org/10.1109/IIAI-AAI.2019.00018

Yi X, Zhang J, Wang Z, Li T, Zheng Y (2018) Deep distributed fusion network for air quality prediction. In: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. KDD '18, pp 965–973. Association for Computing Machinery, New York, NY, USA . https://doi.org/10.1145/3219819.3219822

Zhang Y, Wang Y, Gao M, Ma Q, Zhao J, Zhang R, Wang Q, Huang L (2019) A predictive data feature exploration-based air quality prediction approach. IEEE Access 7:30732–30743. https://doi.org/10.1109/ACCESS.2019.2897754

Zhang K, Thé J, Xie G, Yu H (2020) Multi-step ahead forecasting of regional air quality using spatial-temporal deep neural networks: a case study of Huaihai economic zone. J Clean Prod 277:123231. https://doi.org/10.1016/j.jclepro.2020.123231

Zhang Q, Wu S, Wang X, Sun B, Liu H (2020) A PM2.5 concentration prediction model based on multi-task deep learning for intensive air quality monitoring stations. J Clean Prod 275:122722. https://doi.org/10.1016/j.jclepro.2020.122722

Zhang L, Liu P, Zhao L, Wang G, Zhang W, Liu J (2021) Air quality predictions with a semi-supervised bidirectional LSTM neural network. Atmos Pollut Res 12(1):328–339. https://doi.org/10.1016/j.apr.2020.09.003

Zhang Z, Zeng Y, Yan K (2021) A hybrid deep learning technology for PM2.5 air quality forecasting. Environ Sci Pollut Res 312:39409–39422. https://doi.org/10.1007/s11356-021-12657-8

Zhang Z, Zhang S, Zhao X, Chen L, Yao J (2022) Temporal difference-based graph transformer networks for air quality PM2.5 prediction: a case study in China. Front Environ Sci 10:924986. https://doi.org/10.3389/fenvs.2022.924986

Zhang S, Li X, Li Y, Mei J (2018) Prediction of urban PM2.5 concentration based on wavelet neural network. In: 2018 Chinese Control And Decision Conference (CCDC), pp 5514–5519. https://doi.org/10.1109/CCDC.2018.8408092

Zhang C, Yuan D (2015) Fast fine-grained air quality index level prediction using random forest algorithm on cluster computing of spark. In: 2015 IEEE 12th Intl Conf on Ubiquitous Intelligence and Computing and 2015 IEEE 12th Intl Conf on Autonomic and Trusted Computing and 2015 IEEE 15th Intl Conf on Scalable Computing and Communications and Its Associated Workshops (UIC-ATC-ScalCom), pp 929–934 . https://doi.org/10.1109/UIC-ATC-ScalCom-CBDCom-IoP.2015.177

Zhao G, Huang G, He H, He H, Ren J (2019) Regional spatiotemporal collaborative prediction model for air quality. IEEE Access 7:134903–134919. https://doi.org/10.1109/ACCESS.2019.2941732

Zhao Z, Qin J, He Z, Li H, Yang Y, Zhang R (2020) Combining forward with recurrent neural networks for hourly air quality prediction in Northwest of China. Environ Sci Pollut Res. https://doi.org/10.1007/s11356-020-08948-1

Zhao C, van Heeswijk M, Karhunen J (2016) Air quality forecasting using neural networks. In: 2016 IEEE Symposium Series on Computational Intelligence (SSCI), pp 1–7 . https://doi.org/10.1109/SSCI.2016.7850128

Zheng H, Cheng Y, Li H (2020) Investigation of model ensemble for fine-grained air quality prediction. China Commun 17(7), 207–223. https://doi.org/10.23919/J.CC.2020.07.015

Zhenghua W, Zhihui T (2017) Prediction of air quality index based on improved neural network. In: 2017 International Conference on Computer Systems, Electronics and Control (ICCSEC), pp 200–204. https://doi.org/10.1109/ICCSEC.2017.8446883

Zheng Y, Liu F, Hsieh H-P (2013) U-air: when urban air quality inference meets big data. In: Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. KDD '13, pp 1436–1444. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/2487575.2488188

Zhou Y, Chang F-J, Chang L-C, Kao I-F, Wang Y-S (2019) Explore a deep learning multi-output neural network for regional multi-step-ahead air quality forecasts. J Clean Prod 209:134–145. https://doi.org/10.1016/j.jclepro.2018.10.243

Zhoul L, Chenl M, Ni Q (2020) A hybrid prophet-LSTM model for prediction of air quality index. In: 2020 IEEE Symposium Series on Computational Intelligence (SSCI), pp 595–601. https://doi.org/10.1109/SSCI47803.2020.9308543

Zhu S, Sun J, Liu Y, Lu M, Liu X (2020) The air quality index trend forecasting based on improved error correction model and data preprocessing for 17 port cities in China. Chemosphere 252:126474. https://doi.org/10.1016/j.chemosphere.2020.126474

Zhu H, Hu J (2019) Air quality forecasting using svr with quasi-linear kernel. In: 2019 International Conference on Computer, Information and Telecommunication Systems (CITS), pp 1–5. https://doi.org/10.1109/CITS.2019.8862114

Zhu W, Liu X, Zhu Z (2020) Optimized air quality prediction model based on neural network. In: 2020 International Conference on Computer Engineering and Application (ICCEA), pp 565–568. https://doi.org/10.1109/ICCEA50009.2020.00124