Air pollution forecasting application based on deep learning model and optimization algorithm

Azim Heydari1 · Meysam Majidi Nezhad1 · Davide Astiaso Garcia2 · Farshid Keynia3 · Livio De Santoli1

Received: 3 November 2020 / Accepted: 1 April 2021 / Published online: 14 April 2021 © The Author(s) 2021

Abstract
Air pollution monitoring is constantly increasing, giving more and more attention to its consequences on human health. Since Nitrogen dioxide (NO2) and sulfur dioxide (SO2) are the major pollutants, various models have been developed on predicting their potential damages. Nevertheless, providing precise predictions is almost impossible. In this study, a new hybrid intelligent model based on long short-term memory (LSTM) and multi-verse optimization algorithm (MVO) has been developed to predict and analysis the air pollution obtained from Combined Cycle Power Plants. In the proposed model, long short-term memory model is a forecaster engine to predict the amount of produced NO2 and SO2 by the Combined Cycle Power Plant, where the MVO algorithm is used to optimize the LSTM parameters in order to achieve a lower forecasting error. In addition, in order to evaluate the proposed model performance, the model has been applied using real data from a Combined Cycle Power Plant in Kerman, Iran. The datasets include wind speed, air temperature, NO2, and SO2 for five months (May–September 2019) with a time step of 3-h. In addition, the model has been tested based on two different types of input parameters: type (1) includes wind speed, air temperature, and different lagged values of the output variables (NO2 and SO2); type (2) includes just lagged values of the output variables (NO2 and SO2). The obtained results show that the proposed model has higher accuracy than other combined forecasting benchmark models (ENN-PSO, ENN-MVO, and LSTM-PSO) considering different network input variables.

Graphic abstract

Keywords Nitrogen dioxide · Sulfur dioxide · Air pollution forecasting · Long short-term memory network · Multi-verse optimization
Introduction

High levels of air pollutants are added by industries, vehicles and different natural/anthropogenic sources leading to air quality degradation. As discussed in (Linares et al. 2018), institutions, like World Health Organization (WHO) or the European Environment Agency have reported that being exposed to pollutants increases the risk of early death (Garcia and Aznarte 2020). Air pollution causes harmful effects on the human’s and other animals’ health, together with damages to plants and monuments. In addition, particulate air pollution is directly related to cardiovascular (Brook et al. 2004) and respiratory (Gorai et al. 2014) disorders.

Nitrogen dioxide (NO$_2$) commonly is derived from fossil fuel combustion and causes bronchitis, pneumonia, emphysema, etc., when entered the alveoli. Sulfur dioxide (SO$_2$) comes from the fuel combustion, including sulfur, like coal and petroleum. At high levels, it irritates the respiratory tract making breathing difficult (Chen et al. 2012). Sulfur dioxide (SO$_2$) is one of highly reactive gases called “oxides of sulfur” that is emitted into the atmosphere by fossil fuels (coal and petroleum products), power and industrial plants, and industrial processes, like steel and mining, burning fuels consisting of high sulfur amount because of transportation vehicles, including locomotives, ships, pulp industries, and natural sources, like volcanic emissions (Andersson et al. 2013).

Global energy consumption is constantly increasing (from 3728 Mtoe in 1965 to 12,928 Mtoe in 2014) due to an increase in population and economic growth (Aydin 2015a). The use of different types of energy sources has increased the global primary energy consumption provided by fossil sources, particularly oil, coal, and natural gas, up to 87%. The increase in energy demand for the next years (Aydin 2014) would be met not only by the increase in renewable energy source, but also by fossil fuels, mainly oil and gas (Aydin 2015b). Countries with the highest energy consumption currently constitute about 62% of the world’s energy consumption. Thus, modeling their energy consumption for obtaining an estimated model of future world energy consumption is of great importance (Gokhan Aydin et al. 2016). Cities as high-density urban areas consume a high amount of energy. They occupy only 2% of the land, but use about 75% of the global energy consumption, and are responsible for 80% of the world’s greenhouse gas emission (Feng and Zhang 2012). Moreover, due the increase in fossil fuels use from the Industrial Revolution, greenhouse gases have significantly increased. Consequentially, global warming and climate changes have been the main concerns, particularly during the past two decades. The effect of global warming-related consequences affecting the global economy has been widely studied since the 1990s. International organizations aimed at decreasing the adverse effects of global warming using intergovernmental and binding laws. Carbon dioxide (CO$_2$) has been known as the most important greenhouse gases in the Earth’s atmosphere. The energy sector has been the first by the direct burning of fuels, by which a high amount of CO$_2$ is emitted. CO$_2$ from energy indicates nearly 60% of the emissions of anthropogenic greenhouse gas, which varies highly by country, because of diverse national energy structures (Köne and Büke 2010). Oil plays an important role in the world economy and natural gas (NG) has become a direct competitor for the first because of its environmental benefits and its effect on electricity production. In the next two decades, there will be a rise in energy demand provided by fossil fuels, and oil still plays a main role complemented with the increasing effect of NG mostly on electricity production (Aydin 2014).

In the current century, the rarity of the fossil fuels to generate electricity and the high population make energy as one of the basic needs of developing human societies. However, increased awareness and concern regarding the environmental issues resulted in the increased willingness for using energy-efficient technologies. Thus, energy is one of the most important global challenges for all countries (Mohammadi et al. 2018). The highest approved oil reserves, such as non-conventional oil deposits can be found in Venezuela, Saudi Arabia, Canada and Iran (20%, 18%, 13%, 10.6% of global reserves, respectively) (Tofigh and Abedian 2016). Based on Oil & Gas Journal, from January 2013, Iran has been estimated to have 155 billion barrels of oil reserves, over 10% of the global total reserves and 13% of Organization of the Petroleum Exporting Countries (OPEC) reserves (Tofigh and Abedian 2016) that remains approximately for 94 years (Nejat et al. 2013). The highest proved NG reserves can be found in Russia, Iran, and Qatar (24%, 16.8%, and 12.5% of global reserves), respectively. Also, Iran has been the third-largest NG reserves in Asia because of the South Pars field development. Nonetheless, it will be the seventh NG producer by 2035 (Tofigh and Abedian 2016).

The air quality forecast issue is addressed by time series assessment, in which the model selection and parametrization are highly regarded. Recently, using machine learning approaches, such as neural networks, supported via deep learning are producible (Valput et al. 2019). Gennaro et al. (2013), proposed the artificial neural network (ANN) to forecast daily PM$_{10}$ concentration at regional as well as urban background areas. Feng et al. (2015), presented air mass trajectory assessment as well as wavelet transform for improving ANN prediction accuracy of daily mean PM$_{2.5}$ concentration and it was applied for recognizing marked corridors, whereas wavelet transform was used for coping with the PM$_{2.5}$ concentration fluctuation effectively. In (Sun and Sun
In (Madaan 2019), authors provided a novel forecasting model according to Bidirectional Long Short-Term Memory Networks for the prediction of air quality via estimation of the concentration levels for different pollutants (NO2), particulate matter (PM2.5 and PM10), by which threat level can be classified for them in the next 24 h. Prasad et al. (2016), designed adaptive neuro-fuzzy inference system (ANFIS) in order to predict the daily air pollution levels of SO2, NO2, CO, ozone (O3), and PM10 in the climate of a Megacity (Howrah). In addition, Li et al. (2018) provided an intelligent model for air pollutant concentration prediction according to the weighted extreme learning machine (WELM) as well as the adaptive neuro-fuzzy inference system (ANFIS). García and Aznarte (2020) have proposed a forecasting model based on Shapley additive explanations to predict NO2 time series. Li and Jin (2018) presented a combined intelligent model based on fuzzy synthetic evaluation for the early warning monitoring of air pollutants in China. Sen et al. (2016) proposed an ARIMA forecasting model for predicting energy consumption and greenhouse gas (GHG) emission. The model applied for an Indian pig iron manufacturing organization. Ding et al. (2017) provided a new grey multivariable model to predict CO2 emission form fuel combustion in China. Wang and Ye (2017) presented a forecasting method based on nonlinear grey multivariable method for predicting Chinese carbon emission from fossil energy use. Furthermore, a discrete grey forecasting model for energy-related CO2 emissions prediction in China from 2011 to 2015 implemented by (Ding et al. 2020). A hybrid intelligent prediction model based on semi-experimental regression approach as well as ANFIS for air pollution forecasting has been applied by (Zeinalnezhad et al. 2020). Maleki et al. (2019) applied a forecasting model based on ANN for criteria air pollutant concentrations forecasting such as O3, NO2, SO2, PM10, PM2.5, CO, AQI, and AQHI. In (Say and Yücel 2006), Turkey’s energy sector reviewed from 1970 to 2002. The total energy consumption (TEC) was modeled through the economic growth (proxied by gross national product—GNP) as well as population growth addition, a regression model can predict the TEC according to principal component analysis (PCA) and least squares support vector machine (LSSVM) optimized with cuckoo search algorithm regarding the prediction of PM2.5 concentrations. In (Madaan 2019), authors provided a novel forecasting model according to Bidirectional Long Short-Term Memory Networks for the prediction of air quality via estimation of the concentration levels for different pollutants (NO2), particulate matter (PM2.5 and PM10), by which threat level can be classified for them in the next 24 h. Prasad et al. (2016), designed adaptive neuro-fuzzy inference system (ANFIS) in order to predict the daily air pollution levels of SO2, NO2, CO, ozone (O3), and PM10 in the climate of a Megacity (Howrah). In addition, Li et al. (2018) provided an intelligent model for air pollutant concentration prediction according to the weighted extreme learning machine (WELM) as well as the adaptive neuro-fuzzy inference system (ANFIS). García and Aznarte (2020) have proposed a forecasting model based on Shapley additive explanations to predict NO2 time series. Li and Jin (2018) presented a combined intelligent model based on fuzzy synthetic evaluation for the early warning monitoring of air pollutants in China. Sen et al. (2016) proposed an ARIMA forecasting model for predicting energy consumption and greenhouse gas (GHG) emission. The model applied for an Indian pig iron manufacturing organization. Ding et al. (2017) provided a new grey multivariable model to predict CO2 emission form fuel combustion in China. Wang and Ye (2017) presented a forecasting method based on nonlinear grey multivariable method for predicting Chinese carbon emission from fossil energy use. Furthermore, a discrete grey forecasting model for energy-related CO2 emissions prediction in China from 2011 to 2015 implemented by (Ding et al. 2020). A hybrid intelligent prediction model based on semi-experimental regression approach as well as ANFIS for air pollution forecasting has been applied by (Zeinalnezhad et al. 2020). Maleki et al. (2019) applied a forecasting model based on ANN for criteria air pollutant concentrations forecasting such as O3, NO2, SO2, PM10, PM2.5, CO, AQI, and AQHI. In (Say and Yücel 2006), Turkey’s energy sector reviewed from 1970 to 2002. The total energy consumption (TEC) was modeled through the economic growth (proxied by gross national product—GNP) as well as population growth as the main factors for determining the energy consumption in developing countries. Also, they reviewed the relationship between the TEC and total CO2 (TCO2) emission. Accordingly, they modeled, the potent association between TEC and TCO2 (R2 = 0.998) using regression analysis. In addition, a regression model can predict the TEC according to the population as well as the GNP with high confidence (R2 = 0.996). Kumar et al. (2017) provided American Meteorological Society/Environmental Policy Agency Regulatory Model (AERMOD) in order to predict short-term air quality

considering weather forecasting based on WRF method. The comprehensive emission inventory was provided regarding the sources in Chembur, Mumbai. In (Kumar et al. 2016), vehicular pollution modeling presented with AERMOD using simulated wheater meteorology through Weather Research and Forecasting method. NOx and PM levels were 3.6 and 1.45 times greater in peak time compared with off-peak and evening peak, respectively. Tao et al. (2019), proposed a new short-term forecasting model considering deep learning to PM2.5 concentration. The proposed model has been applied for the Beijing PM2.5 dataset. Liu et al. (2020) presented a new wind-sensitive attention structure using the long short-term memory (LSTM) neural network method for forecasting the air pollution—PM2.5 level. A combined air quality early warning model that includes estimation, forecasting, and assessment is provided by (Jiang et al. 2019). Hähnle et al. (2020) developed a deep-learning framework to monitor and forecast air-pollution that can train across various model domains. In (Ding et al. 2021), a new selection model considering the cooperative data index has been developed to characterize the optimum forecasting group from different forecasts. Agarwal et al. (2020) provided a prediction method based on artificial neural networks for forecasting PM10, PM2.5, NO2, and O3 pollutant level for the current day and the next four days in an area with high pollution.

In the light of the abovementioned state of the art, the main contributions and novelty of this paper are presented as follows:

(a) The air pollution from manufacturing industry production has been analyzed in different months based on various input variables.

(b) Since input variables plays very important roles in the performance of the forecasting model, the best features as input variable have been selected using mutual information model in this paper.

(c) A novel forecasting application (long short-term memory network and multi-verse optimization algorithm) has been proposed to predict the amount of produced NO2 and SO2 by the Combined Cycle Power Plant. In this model, the LSTM parameters are optimized using the MVO metaheuristic algorithm.

(d) The proposed forecasting application has been successfully verified on the real datasets. In addition, the forecasting model is compared with the other valid benchmark models such as (LSTM-PSO, ENN-MVO, and ENN-PSO).

(e) The best advantage of the proposed prediction application compared with the methods presented in recent studies is the use of a combined deep learning method. In particular, in the proposed application, deep learning models and optimization algorithms have been used
and integrated with the method of mutual information to select the best features as the model inputs.

This paper is organized as follows: Sect. 2 explains the case study and artificial intelligence methods. Section 3 describes the proposed air pollution forecasting application. The results and discussion provided in Sect. 4. Finally, Sect. 5 presents research conclusions.

Materials and methods

After a brief explanation of the case study and data gathering, this section explains the artificial intelligence methods developed in the paper.

Case study

The Kerman Combined Cycle Power Plant is located in Kerman, Iran. The latitude and longitude are 56.7904° and 30.2091°, respectively.

The Kerman Combined Cycle Power Plant, with a rated power equivalent to 1912 (MW), includes eight gas units, each with a rated power of 159 (MW) along with 4 steam units, each with a rated power of 160 (MW). Each steam contains two vertical heat recovery steam generator. The exhaust steam is transferred to a steam turbine from two boilers, each boiler consisting of two high and low pressure drums as well as a deaerator drum.

The main fuel of this power plant is natural gas, while its support fuel is gas oil (diesel), which is stored in 20 million liter tanks. In this regard, gas at 70% and diesel at 30% are the fuel of this power plant. Figure 1 shows the Kerman Combined Cycle Power Plant. The data has been collected based on the different stacks output. Moreover, the proposed forecasting model has been tested by two different datasets: (Case A) is the NO₂ and SO₂ produced by stack 1; (Case B) is the NO₂ and SO₂ produced by stack 2. Based on these two cases, we can analyze the level of air pollution and better evaluate the proposed model. The datasets including the wind speed, air temperature, NO₂, and SO₂ for five months (May–September 2019) with a time step of 3-h. The location of Kerman Combined Cycle Power Plant is indicated in Fig. 1. In addition, Fig. 2 represents the Windrose diagram based on wind speed and wind direction data.

Mutual Information

Various indicators are developed to calculate statistical dependence, including Mutual information (MI). Shannon, as the creators of the entropy (Shannon 1948), proposed the following formula to calculate the Shannon entropy (which is also known as information content).

\[ H(X) = - \sum_{i=1}^{N} p(x_i) \log[p(x_i)] \]  

(1)

Yang et al. (2000) determined the MI between two randomly selected variables.

\[ \text{MI}(X, Y) = H(X) + H(Y) - H(X, Y) \]  

(2)

Here \( H(X), H(Y), \) and \( H(X, Y) \) are equal to the entropy of \( X, Y, \) and their common entropy, as follows:

\[ H(X, Y) = - \sum_{x \in X} \sum_{y \in Y} p_{X,Y}(x, y) \]  

(3)

Selecting appropriate variables is a crucial step in developing models. Analyzing the previously observed trends is a valuable source for identifying important factors. Babel and
colleagues argued that access to mutual information is a vital source to identify appropriate factors (Babel et al. 2015). In the present research, MI was used to select the best features of the model and to evaluate the appropriateness of used lagged values. Table 1 pinpoints strong correlations between NO$_2$ and SO$_2$ and all the other indicators.

**Elman neural network**

Elman neural network (ENN) is a kind of self-recursive neural network that is usually composed of two hidden layers and output (Elman 1990). This network is combined by a hidden layer and an output layer. In Elman neural network, due to the recurrent nature, the output of the hidden layer will be returned by the feedback loop to input of the hidden layer. The most important part of neural network is determining the most optimal values of parameters in the network. Network parameters are the weights and biases. Since the Elman neural network has recursive feature, if two different Elman networks with the same weight and biases and with a specified period and specific input enter into the network, they may produce a different output. Figure 3 shows the Elman neural network and recursive loops. This network, like other neural networks for optimal calculation of the parameters and training the network needs a way to train the network to obtain the optimum value of parameters.

**Long short-term memory neural network**

LSTM is a unique recurrent neural network (RNN) method with the RNN features, in which a series of memory cells are used for coping with the arbitrary input data and enhancing the learning process of time series. Also, it captures the long-term dependence of the input information for preventing the

### Table 1 The correlation coefficient of NO$_2$ and SO$_2$ with other indicators

| Different cases | Output variables | Winds speed | Air temperature | NO$_2$$_{-1}$ | NO$_2$$_{-2}$ | NO$_2$$_{-3}$ | SO$_2$$_{-1}$ | SO$_2$$_{-2}$ | SO$_2$$_{-3}$ |
|-----------------|-----------------|-------------|-----------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Case A          | NO$_2$          | 0.2453      | 0.1015          | 0.8177       | 0.7411       | 0.7027       | –            | –            | –            |
|                 | SO$_2$          | 0.232       | 0.1005          | –            | –            | –            | 0.7026       | 0.7005       | 0.6891       |
| Case B          | NO$_2$          | 0.2201      | 0.3058          | 0.9631       | 0.8153       | 0.7558       | –            | –            | –            |
|                 | SO$_2$          | 0.1519      | 0.1225          | –            | –            | –            | 0.7043       | 0.7009       | 0.7001       |
gradient disappearing of data transmission, leading to enhancing its ability for capturing the dynamic alterations of the time series (Yuan et al. 2019). Hochreiter and Schmidhuber (1997) developed the long short-term memory (LSTM) architecture and it was improved by Gers et al. (1999) using an extra forget gate. It is the most effective RNN architecture and is widely used. In the LSTM method, gate units, output and input units, and memory cells, are combined (Yuan et al. 2019).

Gate units:

\[ g_i = \sigma(W_{ix} \cdot X_t + W_{ih} \cdot h_{t-1} + b_i) \]  
\[ g_f = \sigma(W_{fx} \cdot X_t + W_{fh} \cdot h_{t-1} + b_f) \]  
\[ g_o = \sigma(W_{ox} \cdot X_t + W_{oh} \cdot h_{t-1} + b_o) \]  
\[ C_{in} = \beta(W_{cx} \cdot X_t + W_{ch} \cdot h_{t-1} + b_c) \]  
\[ C_t = g_f \cdot C_{t-1} + g_i \cdot C_{in} \]  
\[ h_t = g_o \cdot \theta(C_t) \]  
\[ y = \varphi(W_y \cdot h_t + b_y) \]

where \( \sigma(*) \), \( \beta(*) \), and \( \theta(*) \) indicate nonlinear activation functions, respectively; \( \varphi(*) \) represents the output unit function; the weight coefficient \( W \) as well as bias coefficient \( b \) are applied for establishing the relationship among the units made the model. Figure 4 shows the main structure of the LSTM network.

**Multi-verse optimization algorithm**

As a nature-based algorithm, the multi-verse optimization algorithm (MVO) is developed by Mirjalili et al. (2016). This algorithm is the main inspiration of developing the theory of multi-verse in astrophysics. Based on the MVO, several big bangs form various universes, while white holes, black holes, and wormholes connect these universes. Mirjalili argued that matters of the MVO move from a universe to another by white/black holes so that black and white holes attract and emit matters, respectively. Wormholes connect two sides of a universe. The major terms of this theory are as follows: each universe is a solution, while each solution is comprised of a series of objects, generations or iterations that are used to demonstrate the time, and the inflation rate is used to demonstrate the value...
of each object in a particular universe. In this theory, a solution
is equal to a universe with various white/black/wormholes. To
enhance the value of objects, white holes are assumed to be
more likely in a particular solution, which indicates a higher
value. On the other hand, black holes are more likely to be
formed in objects with the worst values, which causes the
 transmission of values from variables with acceptable solu-
tions. This enhances the likelihood of improving weak solu-
tions, which in turn translates into an improved mean value of
all solutions. In Eqs. 3 and 4 the core structure of the algorithm
is described:

\[
x_i^j = \begin{cases} x_i^j & r_1 < \text{NI}(U_i) \\ x_i^j + \text{TDR} \times ((\text{UB}_i - \text{LB}_b) \times r_2 \times \text{LB}_b) & r_1 < 0.5[r_2 < \text{WEP}] \\ x_i^j - \text{TDR} \times ((\text{UB}_i - \text{LB}_b) \times r_3 \times \text{LB}_b) & r_1 \geq 0.5[r_2 < \text{WEP}] \\ x_i^j & r_2 \geq \text{WEP} \end{cases}
\]  

(12)

where \(X_i^j\) represents the \(j\)th object of the \(i\)th universe, \(r_1\) is
a random number in a predefined spectrum ranging from
zero to one, \(\text{NI}(U_i)\) is equal to the normalized inflation rate
of the \(i\)th universe and \(X_k^j\) represents the \(j\)th object of the \(k\)th
universe.

\[
x_i^j = \begin{cases} x_i^j & r_1 < \text{NI}(U_i) \\ x_i^j + \text{TDR} \times ((\text{UB}_i - \text{LB}_b) \times r_2 \times \text{LB}_b) & r_1 < 0.5 \text{ and } r_2 < \text{WEP} \\ x_i^j - \text{TDR} \times ((\text{UB}_i - \text{LB}_b) \times r_3 \times \text{LB}_b) & r_1 \geq 0.5 \text{ and } r_2 < \text{WEP} \\ x_i^j & r_2 \geq \text{WEP} \end{cases}
\]  

(13)

where \(x_i^j\) is the \(j\)th centroid of the best universe obtained so
far, \(\text{UB}\) represents the upper bound, \(\text{LB}\) equals the minimum
bound, \(\text{TDR}\) and \(\text{WEP}\) are coefficients, \(r_2\), \(r_3\), and \(r_4\) represent
random values ranging from zero to one.

Also, the algorithm of MVO represents the ideal solution
to optimize and apply it to influence other solutions. In the
original research, the authors argued that wormholes can be
found in all universes. Again, in turn, it enhances the likeli-
hood of having access to better solutions and maintaining the
ideal solution which is reached in the optimization process. At
the end of optimization, the ideal solution can be achieved as
a global optimum for a particular problem. The prerequisite
of solving the abovementioned equations is exchanging vari-
bles between various solutions. It worth noting that if these
equations be applied in similar patterns (either exploitative or
exploratory) yield similar results. The MVO contains the
following processes to appropriately focus on various patterns
during the optimization pattern, which is as follows:

\[
\text{WEP} = \text{Min} + \text{Iteration} \times \left( \frac{\text{Max} - \text{Min}}{L} \right)
\]  

(13)

\[
\text{TDR} = 1 - \frac{\text{Iteration}}{\text{Max Iteration}} \times \frac{1}{p}
\]  

(14)

where \(p\) is the exploitation component. Two types of adap-
tive variables are available in MVO: WEP and TDR. WEP
increases based on the frequency of iterations to increase
exploitation. To enhance the precision of exploitation/local
in the process of finding the best solution, TDR should be
increased in various iterations. Hence, MVO can be consid-
ered as a revolutionary algorithm to exchange matters. This
issue indicates the crossover operator, a known revolutionary
operator to find the optimum solution. It leads to an unex-
pected alternation of universes, enhances the exploration,
and keeps the diverseness of the universes while performing
the iterations. After identifying the best universe, each
universe takes a series of variables in a random process. This
issue indicates the mutation, a revolutionary algorithm. In
turn, the mutation operator results in slight alternations in
best solutions and exploitation. Elitism is a revolutionary
operator to maintain the best solution achieved during the
optimization process. The elitism achieves through finding
the best universe (Fathy and Rezk 2018).

### Proposed air pollution forecasting application

The proposed strategy acts in such a way that the data are
firstly evaluated basing on the mutual information statisti-
cal method; the parameters that have a good correlation
move to the next stage; then, input variables (best features)
are classified to predict two output variables (NO\textsubscript{2} and
SO\textsubscript{2}).

After selecting the feature, the future amount of NO\textsubscript{2} and
SO\textsubscript{2} will be predicted by using combined intelligent fore-
casting application. The proposed air forecasting application
phases are as follows:

**Phase 1** Determining the effect coefficient of effective
indicators on the NO\textsubscript{2} and SO\textsubscript{2} emissions, by using the
Pearson correlation.
Phase 2 Designing the main engine forecaster based on long short-term memory neural network considering the obtained best features in phase 1.

Phase 3 Optimizing the parameters of LSTM and ENN neural network, by using the different metaheuristic optimization algorithms such as multi-verse optimization algorithm and particle swarm optimization algorithm.

Fig. 5 The overall framework of the proposed hybrid forecasting application
Phase 4  Testing and training of the combined models, by using the total data as well as the classified data in phase 1.

Figure 5 presents the overall framework of the proposed forecasting application.

Results and discussions

This section summarizes the error measurement indicators and the “NO₂ and SO₂ forecasting results” section.

Error measurement indicators

After optimizing the neural network parameters, the hybrid forecasting models are trained and tested. The optimal parameter values should be integrated in the neural network. The network should be trained, using some of the input data (in this study, 80% of the data is used for training). Afterwards, different error indicators are used to assess the trained network. In this study, to evaluate the performance of forecasting models, three error indicators were incorporated, i.e., RMSE (Eq. 15), MAE (Eq. 16), and MAPE (Eq. 17).

\[
RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^{m} (x_{acti} - x_{fori})^2}
\] (15)

\[
MAE = \frac{1}{m} \sum_{i=1}^{m} |x_{acti} - x_{fori}|
\] (16)

\[
MAPE = \frac{100}{m} \sum_{i=1}^{m} \left| \frac{x_{acti} - x_{fori}}{x_{acti}} \right|
\] (17)

NO₂ and SO₂ forecasting results

Since the environmental pollutants play a very important role in global warming as well as climate change, predicting their growth rate in the near future will help to manage the control of these pollutants. Nowadays, artificial intelligent methods have a very acceptable accuracy in predictive problems. As discussed previously, the aim of this paper is to predict the amount of produced NO₂ and SO₂ by the Combined Cycle Power Plant. In this study, various intelligent models (ENN-PSO, ENN-MVO, and LSTM-PSO) are applied as predictive intelligence models. Table 2 shows the predicted results of the proposed model and the other provided model in this study using the data of Case A from May to September in two different types of input data, respectively. In this Table, two types of input variables are considered. It means, type (1) includes: wind speed, air temperature, and three lagged values of each output variables (NO₂−1, NO₂−2, and NO₂−3 for NO₂ forecasting; SO₂−1, SO₂−2, and SO₂−3 for SO₂ forecasting), type (2) includes: just lagged values of each output variables.

As indicated in Table 2, the performance of the proposed approach is very better than the other combined forecasting models for two output variables (NO₂ and SO₂). In addition, to better monitor the results of the Table, the forecasted value and real value of two different months have been selected and are shown in Fig. 6.

Furthermore, the two output variables (NO₂ and SO₂) have been predicted based on two different types of data for Case B. The results are shown in Table 3.

As shown in Tables 2 and 3, the proposed model (ML-LSTM-MVO) is presented to predict NO₂ and SO₂. The results illustrate that the efficiency and stability of the proposed model is better than the other provided model for both outputs forecasting. Based on the accuracy of the proposed forecasting model, the air pollution produced by a Combined Cycle Power Plant can be predicted and managed. Additionally, accurate and reliable forecasting methods can be used by policy makers as useful tool to accurately define strategies, plans and rules for decreasing atmospheric pollution.
Table 2 Performance of the proposed model and other combined forecasting models to predict NO$_2$ and SO$_2$ based on type (1) and (2) input variables (Case A)

| Output variable | Input variables                                      | Time | Error indicators |
|-----------------|------------------------------------------------------|------|-----------------|
|                 |                                                      |      | ENN-PSO         | ENN-MVO         | LSTM-PSO         | Proposed model   |
|                 |                                                      |      | RMSE  | MAE  | MAPE | RMSE  | MAE  | MAPE | RMSE  | MAE  | MAPE | RMSE  | MAE  | MAPE |
| Data type (1)   |                                                      |      |       |      |      |       |      |      |       |      |      |       |      |      |
| NO$_2$          | Wind speed, air temperature, NO$_{2,-1}$, NO$_{2,-2}$, NO$_{2,-3}$ | May  | 0.0454 | 0.0413 | 25.0994 | 0.0484 | 0.0241 | 25.1845 | 0.0265 | 0.0215 | 18.6828 | 0.0310 | 0.0263 | 17.2091 |
|                 |                                                      | June | 0.1812 | 0.1281 | 18.7259 | 0.1884 | 0.1307 | 18.1022 | 0.1743 | 0.1267 | 18.6185 | 0.1701 | 0.1227 | 16.2374 |
|                 |                                                      | July  | 0.2668 | 0.1416 | 19.6184 | 0.2857 | 0.1129 | 19.6746 | 0.2882 | 0.1394 | 15.919 | 0.2836 | 0.1106 | 12.2081 |
|                 |                                                      | August | 0.1672 | 0.1468 | 19.7783 | 0.1612 | 0.1205 | 17.2039 | 0.1693 | 0.1478 | 19.8978 | 0.1690 | 0.1477 | 19.1738 |
|                 |                                                      | September | 0.2247 | 0.1385 | 24.2174 | 0.1509 | 0.1017 | 15.4708 | 0.2264 | 0.1511 | 17.0821 | 0.2355 | 0.1609 | 13.8128 |
| SO$_2$          | Wind speed, Air Temperature, SO$_{2,-1}$, SO$_{2,-2}$, SO$_{2,-3}$ | May  | 0.3660 | 0.3595 | 18.3843 | 0.0501 | 0.0223 | 19.5786 | 0.0281 | 0.0227 | 23.1595 | 0.0178 | 0.0146 | 18.3162 |
|                 |                                                      | June  | 0.1905 | 0.0932 | 24.3375 | 0.1899 | 0.0887 | 19.0052 | 0.2047 | 0.1022 | 21.6793 | 0.1892 | 0.0841 | 19.6709 |
|                 |                                                      | July  | 0.3901 | 0.3388 | 13.7082 | 0.3405 | 0.1282 | 14.5558 | 0.3459 | 0.1493 | 14.3872 | 0.3424 | 0.1259 | 12.8842 |
|                 |                                                      | August | 0.2619 | 0.2333 | 23.7555 | 0.1110 | 0.0874 | 14.6584 | 0.0999 | 0.0794 | 18.3193 | 0.0938 | 0.0735 | 13.6746 |
|                 |                                                      | September | 0.1746 | 0.1522 | 22.6640 | 0.0438 | 0.0312 | 18.4149 | 0.0339 | 0.0274 | 24.4306 | 0.0327 | 0.0266 | 15.1418 |
| Data type (2)   |                                                      |      |       |      |      |       |      |      |       |      |      |       |      |      |
| NO$_2$          | NO$_{2,-1}$, NO$_{2,-2}$, NO$_{2,-3}$                | May  | 0.0454 | 0.0425 | 26.1412 | 0.0604 | 0.0248 | 24.8419 | 0.0270 | 0.0223 | 22.0448 | 0.0308 | 0.0261 | 19.0186 |
|                 |                                                      | June  | 0.1685 | 0.1206 | 19.6511 | 0.2017 | 0.1342 | 22.3796 | 0.1724 | 0.1249 | 17.2096 | 0.1706 | 0.1233 | 16.2536 |
|                 |                                                      | July  | 0.2791 | 0.1102 | 22.0805 | 0.2933 | 0.1233 | 21.1644 | 0.2821 | 0.1167 | 14.8509 | 0.2828 | 0.1109 | 13.0082 |
|                 |                                                      | August | 0.1735 | 0.1526 | 20.5408 | 0.1861 | 0.1173 | 21.7062 | 0.1605 | 0.1390 | 18.6735 | 0.1591 | 0.1376 | 18.4700 |
|                 |                                                      | September | 1.3237 | 1.3099 | 16.7179 | 0.2428 | 0.1389 | 17.0180 | 0.2294 | 0.1538 | 16.5394 | 0.2211 | 0.1447 | 15.9608 |
| SO$_2$          | SO$_{2,-1}$, SO$_{2,-2}$, SO$_{2,-3}$                | May  | 0.1105 | 0.1089 | 19.9810 | 0.0312 | 0.0175 | 17.4798 | 0.0177 | 0.0145 | 25.2121 | 0.0178 | 0.0145 | 19.5876 |
|                 |                                                      | June  | 0.2050 | 0.1120 | 18.9381 | 0.2214 | 0.1284 | 19.8820 | 0.1943 | 0.0926 | 18.7956 | 0.1890 | 0.0852 | 19.8256 |
|                 |                                                      | July  | 0.5772 | 0.3179 | 24.2852 | 0.3412 | 0.1252 | 19.4358 | 0.3492 | 0.1367 | 22.7681 | 0.3424 | 0.1260 | 15.0703 |
|                 |                                                      | August | 0.1061 | 0.0841 | 18.6300 | 0.1155 | 0.0819 | 17.5991 | 0.0939 | 0.0736 | 18.2375 | 0.0944 | 0.0740 | 16.9249 |
|                 |                                                      | September | 0.2888 | 0.2860 | 24.2992 | 0.0689 | 0.0388 | 21.1896 | 0.0331 | 0.0269 | 25.6128 | 0.0327 | 0.0267 | 20.2243 |
Figure 7 shows the MAPE error of the provided models for NO2 and SO2 forecasting.

Conclusion

As a globally important phenomenon, more attention is paying to air pollution, mainly because predicting the polluted days both can prevent negative health outcomes and provides the necessary information to increase policy-makers’ awareness. Identifying factors that contribute to air pollution and their trend over time has a crucial role in developing effective models aimed to reduce air pollution. In this study, a forecasting method more efficient and effective has been developed, in order to predict air pollution from manufacturing industry. The major contribution of the paper is combining the long short-term memory based on optimizing the hyper parameters of the LSTM using multi-verse optimization metaheuristic algorithm to predict NO2 and SO2. In addition, the proposed method is tested with the real manufacturing-related air pollution dataset from Combined Cycle Power Plant in Kerman, Iran from May to September 2019. Also, the forecasting performances of the proposed method have been compared with some benchmark methods (LSTM-PSO, ENN-MVO, and ENN-PSO). According to the results section, the proposed method (MI-LSTM-MVO) has better performances than other methods considering different months and network input variables.

Nonetheless, despite the homogenous representation of meteorology, forecasts are well linked to the observations. Compared with the uncertainty of the emission inventories, such ANN-oriented forecasting model is rapid and less resource-intensive to generate daily forecasts. The forecasting results are applicable for scientific purposes and also to take short-term corrective plans for air quality control in cities with high pollution. It can be replicated in different cities following the needed setup, optimization, and validation for emergency planning and short-term air quality control.
Table 3: Performance of the proposed model and other combined forecasting models to predict NO₂ and SO₂ based on type (1) and (2) input variables (Case B)

| Output variable | Input variables | Time  | Error indicators |
|-----------------|-----------------|-------|------------------|
|                 |                 |       | ENN-PSO          | ENN-MVO          | LSTM-PSO         | Proposed model   |
|                 |                 |       | RMSE  | MAE  | MAPE | RMSE  | MAE  | MAPE | RMSE  | MAE  | MAPE |
| NO₂             | Wind speed, air temperature NO₂₋₁, NO₂₋₂, NO₂₋₃ | May   | 0.0073 | 0.0056 | 19.3811 | 0.0440 | 0.0168 | 19.9767 | 0.0286 | 0.0242 | 18.3681 | 0.0103 | 0.0087 | 15.3048 |
|                 |                 | June  | 0.1714 | 0.1248 | 16.3649 | 0.1679 | 0.1214 | 16.2431 | 0.1710 | 0.1210 | 17.1220 | 0.1641 | 0.1149 | 15.8799 |
|                 |                 | July  | 0.4116 | 0.2112 | 29.5026 | 0.3815 | 0.2025 | 25.2560 | 0.4025 | 0.2223 | 20.2054 | 0.4061 | 0.2137 | 15.4447 |
|                 |                 | August| 0.1617 | 0.1360 | 23.5549 | 0.1144 | 0.0911 | 21.1828 | 0.1916 | 0.1697 | 21.7340 | 0.1802 | 0.1575 | 20.0865 |
|                 |                 | September| 0.1374 | 0.0763 | 17.5957 | 0.0489 | 0.0381 | 18.4797 | 0.0426 | 0.0314 | 19.1170 | 0.0546 | 0.0465 | 17.4011 |
| SO₂             | Wind speed, air temperature SO₂₋₁, SO₂₋₂, SO₂₋₃ | May   | 0.0202 | 0.0163 | 30.4026 | 0.0601 | 0.0310 | 18.0344 | 0.0217 | 0.0171 | 14.4162 | 0.0188 | 0.0154 | 13.3942 |
|                 |                 | June  | 0.1733 | 0.0932 | 15.9190 | 0.1575 | 0.0769 | 15.4098 | 0.1570 | 0.0810 | 15.5723 | 0.1562 | 0.0746 | 12.4922 |
|                 |                 | July  | 0.4537 | 0.2484 | 20.5735 | 0.3912 | 0.2135 | 21.6297 | 0.4776 | 0.2582 | 19.8659 | 0.4652 | 0.2410 | 19.1790 |
|                 |                 | August| 0.0961 | 0.0798 | 28.1677 | 0.1272 | 0.0777 | 23.8905 | 0.0812 | 0.0617 | 25.2105 | 0.0791 | 0.0583 | 15.2904 |
|                 |                 | September| 0.6125 | 0.2457 | 19.8431 | 0.0528 | 0.0201 | 18.6855 | 0.0118 | 0.0104 | 19.1881 | 0.0127 | 0.0126 | 18.4804 |
| NO₂             | NO₂₋₁, NO₂₋₂, NO₂₋₃ | May   | 0.0473 | 0.0464 | 19.4336 | 0.0112 | 0.0085 | 15.8871 | 0.0134 | 0.0119 | 17.4205 | 0.0103 | 0.0087 | 15.6168 |
|                 |                 | June  | 0.1838 | 0.1374 | 25.8230 | 0.2275 | 0.1529 | 21.7123 | 0.1673 | 0.1165 | 16.6511 | 0.1647 | 0.1161 | 15.9170 |
|                 |                 | July  | 0.6432 | 0.4242 | 27.6977 | 0.3864 | 0.2163 | 22.4953 | 0.4116 | 0.2207 | 24.2079 | 0.4062 | 0.2136 | 20.3056 |
|                 |                 | August| 0.1194 | 0.0992 | 17.6530 | 0.1027 | 0.0767 | 19.5364 | 0.1843 | 0.1618 | 20.6773 | 0.1777 | 0.1548 | 19.7355 |
|                 |                 | September| 0.4945 | 0.1031 | 18.7608 | 0.0890 | 0.0368 | 24.0528 | 0.0527 | 0.0442 | 17.9251 | 0.0545 | 0.0465 | 17.3959 |
| SO₂             | SO₂₋₁, SO₂₋₂, SO₂₋₃ | May   | 0.0956 | 0.0815 | 17.1580 | 0.0202 | 0.0163 | 13.9245 | 0.0178 | 0.0139 | 21.3474 | 0.0189 | 0.0154 | 13.3872 |
|                 |                 | June  | 0.1892 | 0.1243 | 21.8019 | 0.1742 | 0.0931 | 17.5050 | 0.1711 | 0.0901 | 16.0800 | 0.1562 | 0.0747 | 15.4610 |
|                 |                 | July  | 0.9621 | 0.6573 | 25.6021 | 0.2680 | 0.1402 | 22.3912 | 0.4634 | 0.2431 | 18.3560 | 0.4651 | 0.2411 | 19.6895 |
|                 |                 | August| 0.0811 | 0.0624 | 24.9351 | 0.1340 | 0.0882 | 24.6836 | 0.0798 | 0.0596 | 15.4527 | 0.0791 | 0.0583 | 12.5204 |
|                 |                 | September| 0.2094 | 0.2072 | 18.0764 | 0.0935 | 0.0187 | 19.7296 | 0.0104 | 0.0103 | 19.3974 | 0.0126 | 0.0126 | 15.8316 |
Air pollution forecasting application based on deep learning model and optimization algorithm

Funding  Open access funding provided by Università degli Studi di Roma La Sapienza within the CRUI-CARE Agreement.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Agarwal S, Sharma S, Suresh R, Rahman MdH, Vranckx S, Maiheu B, Blyth L, Janssen S, Gargava P, Shukl 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

Andersson SM, Martinsson BG, Friberg J, Brenninkmeijer CAM, Raathe-Schöch A, Hermann M, Van Velthoven PFI, Zahn A (2013) Composition and evolution of volcanic aerosol from eruptions of Kasatochi, Sarychev and Eyjafjallajökull in 2008–2010 based on CARIBIC observations. Atmos Chem Phys 13(4):1781–1796. https://doi.org/10.5194/acp-13-1781-2013

Aydin G (2014) Production modeling in the oil and natural gas industry: an application of trend analysis. Pet Sci Technol 32(5):555–564. https://doi.org/10.1080/10916466.2013.825271

Aydin G (2015a) Forecasting natural gas production using various regression models. Pet Sci Technol 33(15–16):1486–1492. https://doi.org/10.1080/10916466.2015.1076842

Aydin G (2015b) Regression models for forecasting global oil production. Pet Sci Technol 33(21–22):1822–1828. https://doi.org/10.1080/10916466.2015.1101474

Aydin G, Jang H, Topal E (2016) Energy consumption modeling using artificial neural networks: the case of the world’s highest consumers. Energy Sources Part B 11(3):212–219. https://doi.org/10.1080/15567249.2015.1075086

Babel MS, Badgjajar GB, Shinde VR (2015) Using the mutual information technique to select explanatory variables in artificial neural networks for rainfall forecasting. Meteorol Appl 616:610–616. https://doi.org/10.1002/met.1495

Fig. 7 Comparisons of MAPE error criteria for different forecasting models
Brook RD, Franklin B, Cascio W, Hong Y, Howard G, Lipsett M, Luepker R, Mittleman M, Samet J, Smith SC, Tager I (2004) Air pollution and cardiovascular disease: a statement for healthcare professionals from the expert panel on population and prevention science of the American heart association. Circulation 109(21):2655–2671. https://doi.org/10.1161/01.CIR.0000128587.30041.C8

Chen C, Zhao B, Weschler CJ (2012) Assessing the influence of indoor exposure to “outdoor ozone” on the relationship between ozone and short-term mortality in US communities. Environ Health Perspect 120(2):235–240. https://doi.org/10.1289/ehp.1103921

Ding S, Dang YG, Li XM, Wang JJ, Zhao K (2017) Forecasting Chinese CO₂ emissions from fuel combustion using a novel grey multivariable model. J Clean Prod 162:1527–1538. https://doi.org/10.1016/j.jclepro.2017.06.167

Ding S, Xu N, Ye J, Zhou W, Zhang X (2020) Estimating Chinese energy-related CO₂ emissions by employing a novel discrete grey prediction model. J Clean Prod 259:120793. https://doi.org/10.1016/j.jclepro.2020.120793

Ding Z, Chen H, Zhou L (2021) Optimal group selection algorithm in air quality index forecasting via cooperative information criterion. J Clean Prod 283:125248. https://doi.org/10.1016/j.jclepro.2020.125248

Elman JL (1990) Finding structure in time. Cogn Sci 14(2):179–211

Fathy A, Rezk H (2018) Multi-verse optimizer for identifying the optimal parameters of PEMFC model. Energy 143:634–644. https://doi.org/10.1016/j.energy.2017.11.014

Feng YY, Zhang LX (2012) Scenario analysis of urban energy saving and carbon abatement policies: A case study of Beijing city. China. Procedia Environ Sci 13:632–644. https://doi.org/10.1016/j.proenv.2012.01.055

Feng X, Li Q, Zhu Y, Hou J, Jin L, Wang J (2015) Artificial neural networks forecasting of PM_{1.0} pollution using air mass trajectory based geographic model and wavelet transformation. Atmos Environ 107:118–128. https://doi.org/10.1016/j.atmosenv.2015.02.030

García MV, Aznarte JL (2020) Ecological Informatics Shapley additive explanations for NO₂ forecasting. Eco Inform 56(2):101039. https://doi.org/10.1016/j.ecoinf.2019.101039

Gennaro G, Trizio L, Di A, Pey J, Pérez N, Cusack M, Alastuey A, Querol X (2013) Neural network model for the prediction of PM_{10} daily concentrations in two sites in the Western Mediterranean. Sci Total Environ 463–464:875–883. https://doi.org/10.1016/j.scitotenv.2013.06.003

Gers FA, Schmidhuber J, Cummins F (1999) Learning to forget: continuation prediction with LSTM. In: 9th International conference on artificial neural networks: ICANN, pp 850–855

Gorai AK, Tuluri F, Tchounwou PB (2014) A GIS based approach for assessing the association between air pollution and asthma in New York State, USA. Int J Environ Res Public Health 11(5):4845–4869. https://doi.org/10.3390/ijerph110504845

Hähnel P, Mareček J, Monteil J, O’Donncha F (2020) Using deep learning to extend the range of air pollution monitoring and forecasting. J Comput Phys 408:109278. https://doi.org/10.1016/j.jcp.2020.109278

Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9(8):1735–1780

Jiang P, Li C, Li R, Yang H (2019) An innovative hybrid air pollution early-warning system based on pollutants forecasting and Extncs evaluation. Knowl-Based Syst 164:174–192. https://doi.org/10.1016/j.knosys.2018.10.036

Köne AÇ, Bükê T (2010) Forecasting of CO₂ emissions from fuel combustion using trend analysis. Renew Sustain Energy Rev 14:2906–2915. https://doi.org/10.1016/j.rser.2010.06.006

Kumar A, Patil RS, Kumar A, Rakesh D (2016) Comparison of predicted vehicular pollution concentration with air quality standards for different time periods. Clean Technol Environ Policy 18(7):2293–2303. https://doi.org/10.1007/s10098-016-1147-6

Kumar A, Patil RS, Kumar A, Rakesh D (2017) Application of AERMOD for short-term air quality prediction with forecasted meteorology using WRF model. Clean Technol Environ Policy 19(7):1955–1965. https://doi.org/10.1007/s10098-017-1379-0

Li R, Jin Y (2018) The early-warning system based on hybrid optimization algorithm and fuzzy synthetic evaluation model. Inf Sci 435:296–319. https://doi.org/10.1016/j.ins.2017.12.040

Li Y, Jiang P, She Q, Lin G (2018) Research on air pollutant concentration prediction method based on self-adaptive neuro-fuzzy weighted extreme learning machine. Environ Pollut 241:1115–1127. https://doi.org/10.1016/j.envpol.2018.05.072

Linares C, Falcón I, Ortiz C, Díaz J (2018) An approach estimating the short-term effect of NO₂ on daily mortality in Spanish cities. Environ Int 116(2):18–28. https://doi.org/10.1016/j.envint.2018.04.002

Liu D-R, Lee S-J, Huang Y, Chiu C-J (2020) Air pollution forecasting based on attention-based LSTM neural network and ensemble learning. Expert Syst 37(3):1–16. https://doi.org/10.1111/exsy.12511

Madaan D (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

Maleki H, Sorooshian A, Goudarzi G, Baboli Z, Tahmasebi Y (2019) Air pollution prediction by using an artificial neural network model. Clean Technol Environ Policy 21(6):1341–1352. https://doi.org/10.1007/s10098-019-01709-w

Mirjalili S, Mirjalili SM, Hatamlou A (2016) Multi-Verse optimizer: a nature-inspired algorithm for global optimization. Neural Comput Appl 27(30):495–513. https://doi.org/10.1007/s00521-015-1870-7

Mohammadi M, Ghasempour R, Astaraci FR, Ahmadi E, Aligholian A, Toopshekan A (2018) Optimal planning of renewable energy resource for a residential house considering economic and reliability criteria. Electr Power Energy Syst 96:261–273. https://doi.org/10.1016/j.ep energy.2017.11.014

Nejad P, Kasir A, Jomehzadeh F, Behzad H, Saeed M, Majid MZA (2013) Iran’s achievements in renewable energy during fourth development program in comparison with global trend. Renew Sustain Energy Rev 22:561–570. https://doi.org/10.1016/j.rser.2013.01.042

Prasad K, Gorai AK, Goyal P (2016) Development of ANFIS models for air quality forecasting and input optimization for reducing the computational cost and time. Atmos Environ 128:246–262. https://doi.org/10.1016/j.atmosenv.2015.06.007

Say NP, Yücel M (2006) Energy consumption and CO₂ emissions in Turkey: empirical analysis and future projection based on an economic growth. Energy Policy 34:3870–3876. https://doi.org/10.1016/j.enpol.2005.08.024

Sen P, Roy M, Pal P (2016) Application of ARIMA for forecasting energy consumption and GHG emission: a case study of an Indian pig iron manufacturing organization. Energy 116:1031–1038. https://doi.org/10.1016/j.energy.2016.10.068

Shannon CE (1948) A mathematical theory of communication. Bell Syst Tech J 27(3):379–423

Sun W, Sun J (2017) Daily PM_{2.5} concentration prediction based on principal component analysis and LSSVM optimized by cuckoo search algorithm. J Environ Manag 188:144–152. https://doi.org/10.1016/j.jenvman.2016.12.011

Tao Q, Liu F, Li Y, Sidorov D (2019) Air pollution forecasting using a deep learning model based on 1D convnets and bidirectional GRU. IEEE Access 7:76690–76698. https://doi.org/10.1109/ACCESS.2019.2921578
Tofigh AA, Abedian M (2016) Analysis of energy status in Iran for designing sustainable energy roadmap. Renew Sustain Energy Rev 57:1296–1306. https://doi.org/10.1016/j.rser.2015.12.209
Valput D, Navares R, Aznarte JL (2019) Forecasting hourly NO₂ concentrations by ensembling neural networks and mesoscale models. Neural Comput Appl 32:9331–9342. https://doi.org/10.1007/s00521-019-04442-z
Wang ZX, Ye DJ (2017) Forecasting Chinese carbon emissions from fossil energy consumption using non-linear grey multivariable models. J Clean Prod 142:600–612. https://doi.org/10.1016/j.jclepro.2016.08.067
Yang HH, Van Vuuren S, Sharma S, Hermansky H (2000) Relevance of time–frequency features for phonetic and speaker-channel classification. Speech Commun 31(1):35–50. https://doi.org/10.1016/S0167-6393(00)00007-8
Yuan X, Chen C, Jiang M, Yuan Y (2019) Prediction interval of wind power using parameter optimized Beta distribution based LSTM model. Appl Soft Comput J 82:105550. https://doi.org/10.1016/j.asoc.2019.105550
Zeinalnezhad M, Ghomazadeh A, Klem J (2020) Air pollution prediction using semi-experimental regression model and adaptive neuro-fuzzy inference system. J Clean Prod 261:121218. https://doi.org/10.1016/j.jclepro.2020.121218

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Authors and Affiliations**

Azim Heydari¹ · Meysam Majidi Nezhad¹ · Davide Astiaso Garcia² · Farshid Keynia³ · Livio De Santoli¹

¹ Department of Astronautical, Electrical and Energy Engineering (DIAEE), Sapienza University, Rome, Italy
² Department of Planning, Design, and Technology of Architecture, Sapienza University of Rome, Rome, Italy
³ Department of Energy Management and Optimization, Institute of Science and High Technology and Environmental Sciences, Graduate University of Advanced Technology, Kerman, Iran