ARTICLE

The Influence of Air Pollution Concentrations on Solar Irradiance Forecasting Using CNN-LSTM-mRMR Feature Extraction

Ramiz Gorkem Birdal[*]

Department of Computer Engineering, Istanbul University-Cerrahpasa, Istanbul, 34300, Turkey

*Corresponding Author: Ramiz Gorkem Birdal. Email: ramizgorkem.birdal@iuc.edu.tr

ABSTRACT

Maintaining a steady power supply requires accurate forecasting of solar irradiance, since clean energy resources do not provide steady power. The existing forecasting studies have examined the limited effects of weather conditions on solar radiation such as temperature and precipitation utilizing convolutional neural network (CNN), but no comprehensive study has been conducted on concentrations of air pollutants along with weather conditions. This paper proposes a hybrid approach based on deep learning, expanding the feature set by adding new air pollution concentrations, and ranking these features to select and reduce their size to improve efficiency. In order to improve the accuracy of feature selection, a maximum-dependency and minimum-redundancy (mRMR) criterion is applied to the constructed feature space to identify and rank the features. The combination of air pollution data with weather conditions data has enabled the prediction of solar irradiance with a higher accuracy. An evaluation of the proposed approach is conducted in Istanbul over 12 months for 43791 discrete times, with the main purpose of analyzing air data, including particular matter (PM10 and PM25), carbon monoxide (CO), nitric oxide (NOX), nitrogen dioxide (NO_2), ozone (O_3), sulfur dioxide (SO_2) using a CNN, a long short-term memory network (LSTM), and MRMR feature extraction. Compared with the benchmark models with root mean square error (RMSE) results of 76.2, 60.3, 41.3, 32.4, there is a significant improvement with the RMSE result of 5.536. This hybrid model presented here offers high prediction accuracy, a wider feature set, and a novel approach based on air concentrations combined with weather conditions for solar irradiance prediction.

KEYWORDS

Forecasting solar irradiance; air pollution; convolutional neural network; long short-term memory network; mRMR feature extraction

1 Introduction

A polluted environment can affect the performance of renewable energy technologies such as solar panels and wind turbines. The presence of particulate matter and pollutants in the air can negatively affect the efficiency of solar panels by reducing the amount of sunlight they receive. As well, air pollution can affect wind patterns and the performance of wind turbines. Regulatory frameworks, public perception, and market dynamics can all be adversely affected by air pollution, which can influence the efficiency of renewable energy technologies.

Particulate matter and pollutants in the atmosphere can scatter sunlight in various directions. A diffuse sky radiation component can result from this scattering, reducing direct sunlight reaching the surface of the Earth. Also some certain pollutants and aerosols, such as black carbon (soot) and brown carbon, can absorb sunlight. A result of this absorption is a warming of the atmosphere, and a reduction in sunlight reaching the ground. Besides, air pollution can have an impact on cloud properties, such as cloud formation and optical thickness. The transmission of solar radiation through the atmosphere is affected by changes in cloud properties. Furthermore, airborne particles can settle on solar panels, reducing their transparency and the amount of sunlight that can be absorbed by the photovoltaic cells. This can lead to decreased efficiency in solar energy conversion. Observing the relationship between air pollution and solar irradiance is crucial for accurately assessing solar energy generation potential and optimizing solar power systems' performance.

In recent years, there has been a steady increase in demand for clean energy in the world since the pandemic [1,2]. Renewable energy as solar can save a lot of greenhouse gas emissions because it is one of the cleanest sources of energy [3–5]. Renewable energy planning has therefore become dependent on forecasts of solar irradiance, wind, and precipitation. It has been demonstrated in the literature that solar irradiance is estimated using solar geometry interactions between solar altitude angels, weather factors, etc., [6,7]. Furthermore, some studies are being conducted without a clear understanding of solar geometry [8,9]. At the same time, early iterations of data-driven models were built utilizing Artificial Neural Network (ANN) approaches [10–12]. For a detailed analysis, the aim was to increase the depth of artificial learning. Unlike their ANN counterparts, deep learning methods enables to processing of large data sets effectively [13]. As it relates to forecasting solar energy, some researchers utilized long short-term memory (LSTM) algorithm for prediction [14–16]. A study in another journal examines the differences in the identification phase of feed-forward and LSTM neural predictors [17]. The LSTM model was also used in a study that combined meteorological features and previous irradiance trends [18]. Another variant of deep stacked long short-term which proposes a dropout and early stop regularization strategy based on the Memory Network approach is utilized [19]. Furthermore, different models conducted in the literature were used to analyze the concentration of some polluting parameters [20]. In the field of deep learning, some studies have implemented a deep learning algorithm to forecast solar activity [21]. Studies in recent years took a different approach by collecting ground-level cloud images and establishing a relationship between images and insolation [22–25]. Using the local cloud cover (LCC) numerical feature in combination with the cloud feature in the sky image, some researchers improved the forecasting accuracy of horizontal irradiance [26]. The mRMR was used by some researchers to reduce the dimension of 100 deep features for CNN architecture by using mRMR [27–29]. At the same time mRMR feature selection algorithm is used with decision tree (DT) [30], k-nearest neighbors (kNN) [31], linear discriminant analysis (LDA) [32], linear regression (LR) [33].

This study is organized as follows: The materials and methods section gives a short information about the dataset and the methods utilized. The results and discussion section presents the output results of the suggested approach. Conclusion remarks are given in the conclusions section.

2 Structure

2.1 Methods, Instruments and Data

The data used in the study was provided by the Istanbul Air Quality Monitoring Network which included hourly values for a total of 356 days between January 01, 2022, and January 01, 2023. Fig. 1 shows dust-related PM10 and PM25 values, while Fig. 2 shows some values for greenhouse

gases. Moreover, Fig. 3 shows rain, temperature, and wind speed values related to weather. A chart representing the solar irradiation value throughout the year can be seen in Fig. 4.
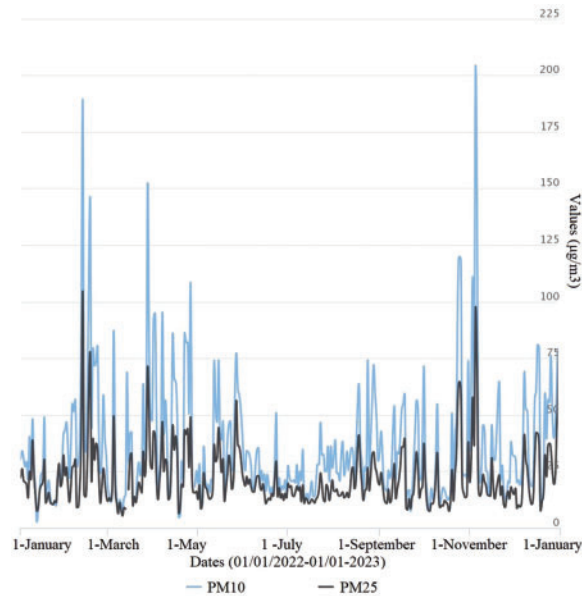

Figure 1: Measurements related to dust (x-axis: time|y-axis: µg/m³)

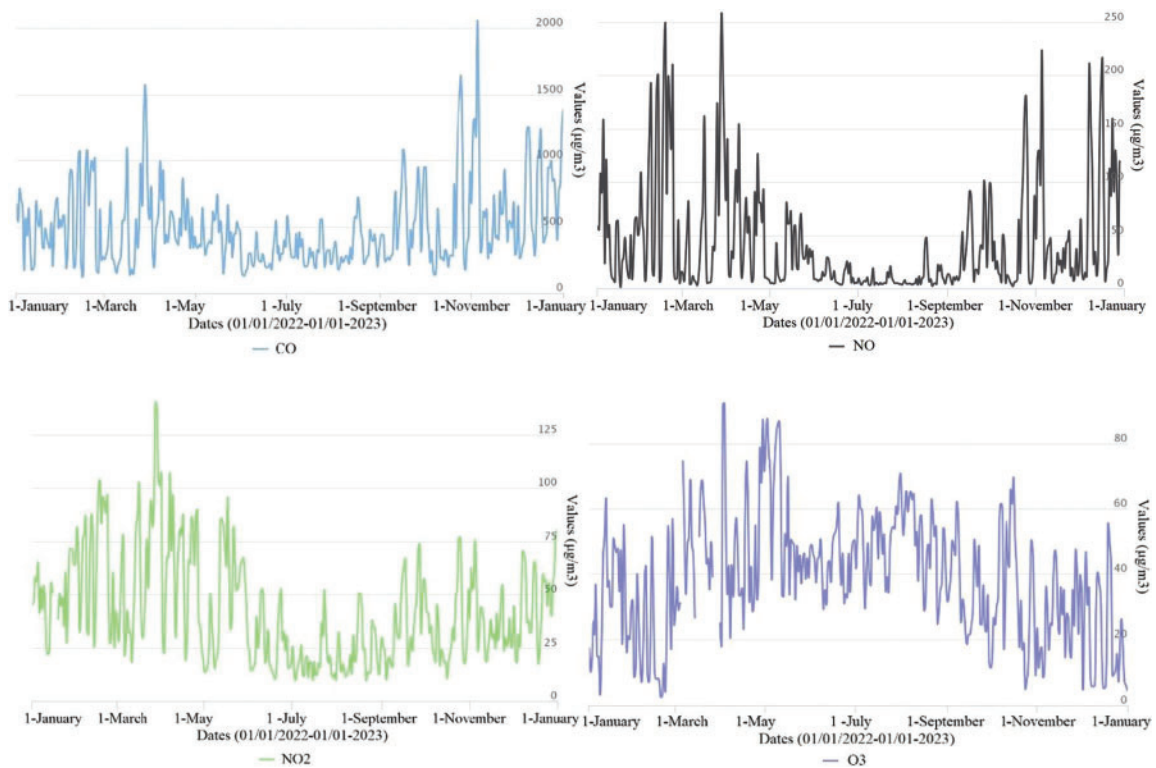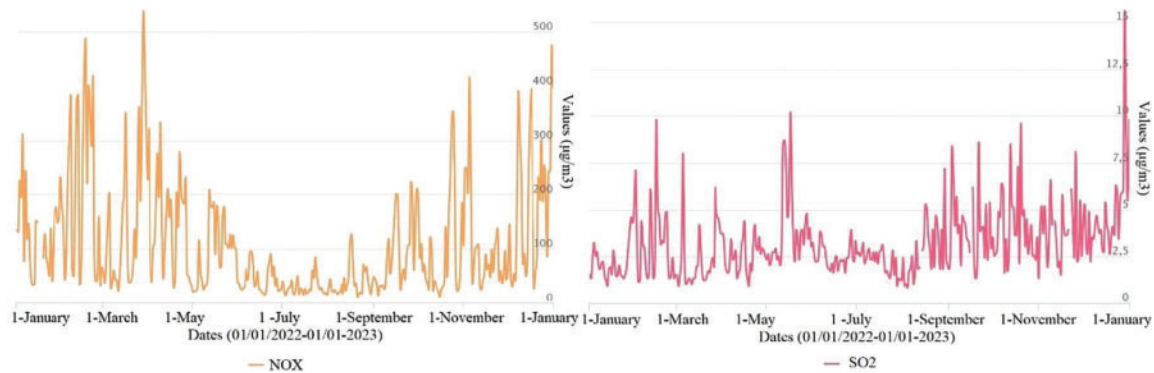**Figure 2:** (Continued)

Figure 2: Values related to some greenhouse gases & aerosols (x-axis: time|y-axis: $\mu g/m^3$)
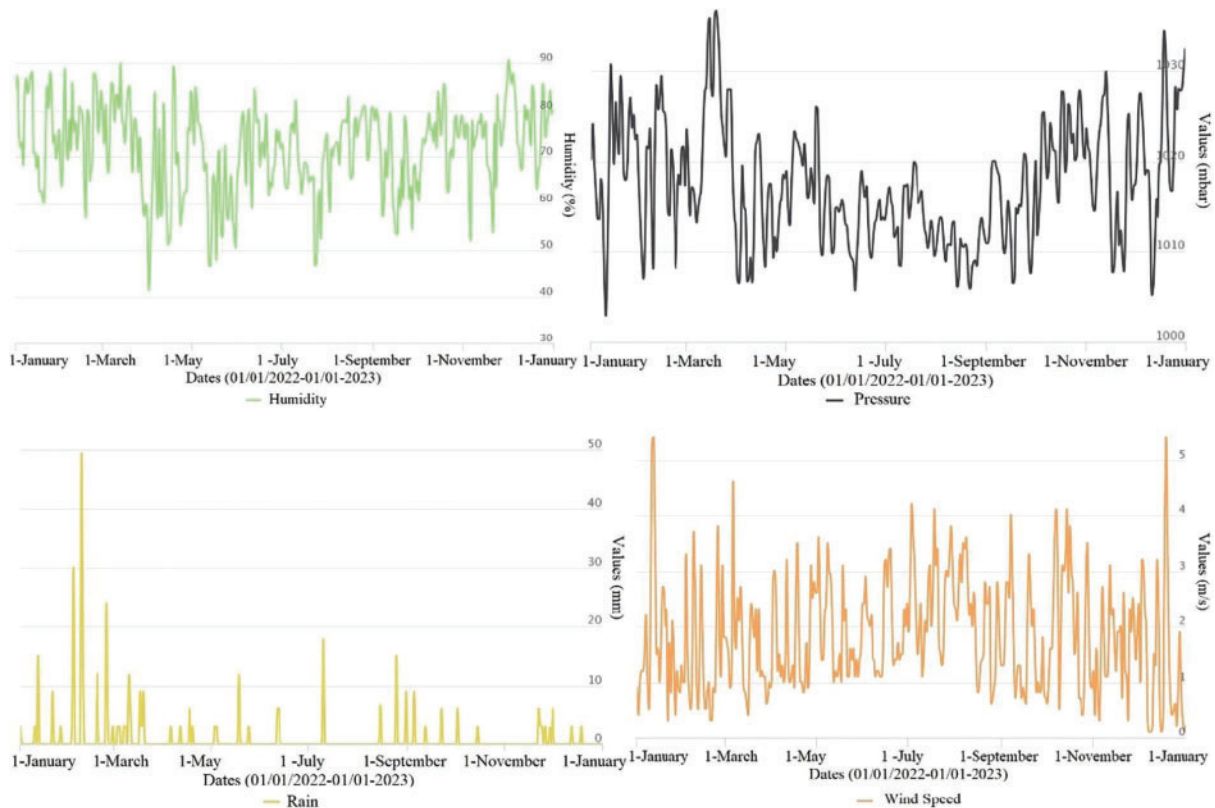

Figure 3: Associated values with weather conditions

2.2 Convolutional Neural Network (CNN)

Over the years, CNN's achieved capability to extract features has allowed it to be successfully applied in a wide variety of fields. Therefore, the spatial characteristics of data captured efficiently can reveal a wealth of information about solar radiation. The influence of weather and air pollution on the concentration change of solar irradiance was revealed by extracting the spatial features of solar irradiance (SI). As shown in Fig. 5, the raw data utilized in the paper includes PM10, PM25, CO, NO, NO_2, NOX, O_3, SO_2, humidity (H%), rain (mm), atmospheric pressure (mbar), wind speed (m/s), air temperature (°C) and solar irradiance (W/m^2).
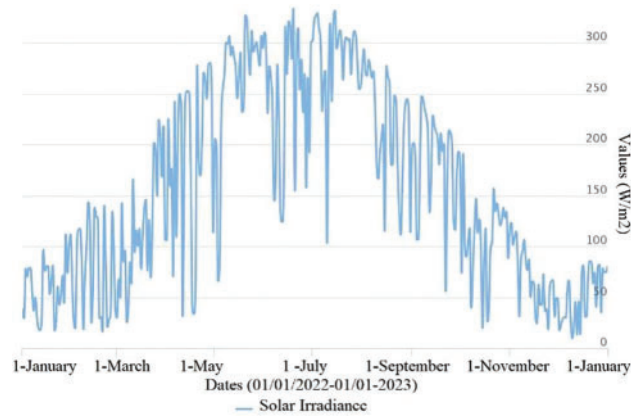
Figure 4: Solar irradiance values (x-axis: time|y-axis: W/m²)
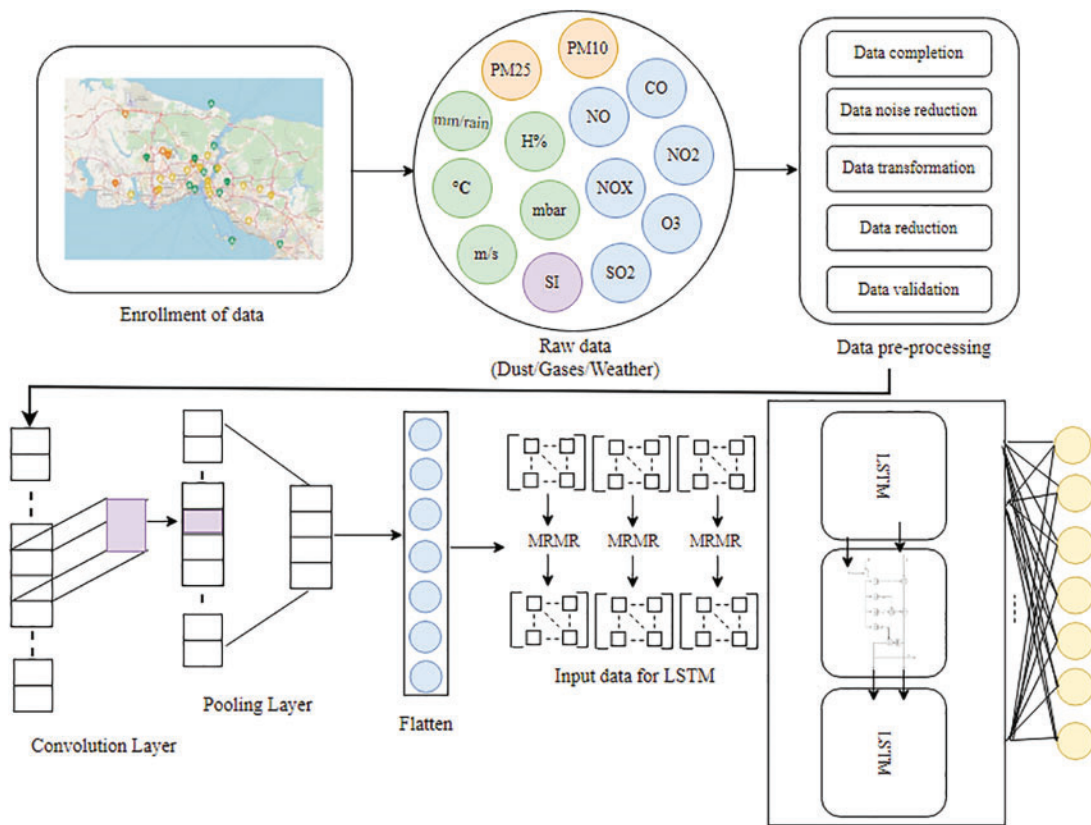

Figure 5: Proposed framework

CNNs differ from traditional neural networks in how they are composed of three layers: convolution, pooling, and full connection layers. CNN's proposed feature extraction block consists of three layers of convolution in one-dimensional (1D) in the study. The convolution and pooling layers are computed as:

$$y_i = f(w_i * x_i + b_i) \tag{1}$$

$$\gamma(y_i, y_{i-1}) = max(y_i, y_{i-1}) \tag{2}$$

$$p_i = \gamma(y_i, y_{i-1}) + \delta_i \tag{3}$$

where x_i is the input to the convolution layer, representing the pre-processed concentration values of PM10, PM25, CO, NO, NO$_2$, NOX, O$_3$, SO$_2$, humidity (H%), rain (mm), atmospheric pressure (mbar), wind speed (m/s), air temperature (°C) and solar irradiance (W/m^2). y_i is the feature mapping of the ith output, w_i is the weight matrix, $*$ indicates the dot product, b_i and δ_i are bias vectors, $f(.)$ denotes the activation function, $\gamma(.)$ denotes the sampling function, p_i represents the output of the pooling layer after the convolution operation. Convolution layers were separated by MaxPooling and Rectified Linear Unit (ReLU) as seen in Table 1.

Table 1: Construction of the model

Proposed Model		
Convolution 1	Kernels	96
	Size of receptive field	3
MaxPooling	–	–
ReLU	–	–
Convolution 1	Kernels	48
	Size of receptive field	3
MaxPooling	–	–
ReLU	–	–
Convolution 1	Kernels	24
	Size of receptive field	3
MaxPooling	–	–
ReLU	–	–
Dropout & MRMR Layer	–	0.25
LSTM1	Hidden notes	20
	Return sequence	True
LSTM2	Hidden notes	20
	Return sequence	True
LSTM3	Hidden notes	20
	Return sequence	False
Fully connected	Hidden notes	20
Output	Hidden notes	1/2/6

A pooling layer is typically included after the convolutional layer to mitigate the limitation of the invariance of the created feature map, whilst the activation function is used to improve the model's capacity to learn complex structures. The LSTM network structure lies behind this step in the system's flow.

2.3 Long Short-Term Memory (LSTM)

As opposed to traditional neural networks, which suffer from vanishing gradients, Long Short-Term Memory (LSTM) is a recurrent neural network designed to overcome this problem. It occurs

when gradients used to update a network's weights become very small during backpropagation, making it difficult for the network to learn from long-term dependencies in the data. In addition to capturing both local and temporal variations in solar irradiance, the CNN-LSTM model can also anticipate future conditions by integrating spatial awareness and temporal memory. It combines the extracted features from the CNN with the temporal information learned by the LSTM. Input sequences to the LSTM layer can include the output from the CNN and other relevant time-dependent features. Since LSTM networks are able to remember information for long periods of time, they can capture sequential patterns effectively. The flow of information into and out of each memory cell is controlled by three types of gates: input gates, output gates, and forget gates. During training, the LSTM network learns to selectively store or discard information in the memory cells based on the input data and the current state of the network. The input gate determines which information to store in the memory cells, while the forget gate determines which information to discard. The output gate then determines which information to output from the memory cells to the next layer of the network. The structure diagram of LSTM is seen in Fig. 6.
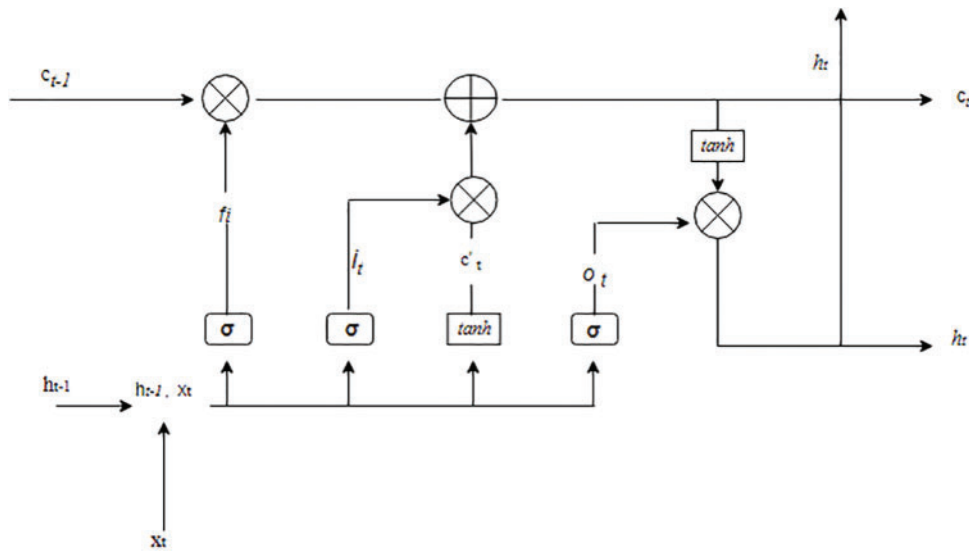

Figure 6: Structure of LSTM

Variables:

$x_t \in \mathfrak{R}^d$: input vector to the LSTM unit

$f_t \in (0, 1)^h$: forget gate's activation vector

$i_t \in (0, 1)^h$: input/update gate's activation vector

$o_t \in (0, 1)^h$: output gate's activation vector

$h_t \in (-1, 1)^h$: hidden state vector also known as output vector of the LSTM unit

$c'_t \in (-1, 1)^h$: cell input activation vector

$c_t \in \mathfrak{R}^h$: cell state vector

$W \in \mathfrak{R}^{hxd}, U \in \mathfrak{R}^{hxh}$ and $b \in \mathfrak{R}^h$: weight matrices and bias vector parameters which need to be learned during training where the superscripts d and h h refer to the number of input features and the number of hidden units, respectively.

An output vector between 0 and 1 is generated by passing the current input and the previous hidden state through a sigmoid function. The hidden state is the output of the LSTM network at each time step. It is a combination of the current memory cell state and the output gate values. As part of the training process, the network parameters are updated according to an optimizing algorithm using backpropagation through time (BPTT), which is a method for computing the gradient of the loss in each time step.

2.4 Maximum Relevance Minimum Redundancy (mRMR)

One of the most important aspects of our approach is that we use the mRMR method for feature selection. A feature selection algorithm called mRMR creates a subset within which related data properties are extracted and unrelated ones are discarded. The algorithm calculates the similarities between each attribute of X and Y by taking into account the mutual information between the two attributes.

$$I(X, Y) = \sum_{y \in Y} \sum_{x \in X} p(x, y) \log\left(\frac{p(x, y)}{p_1(x) p_2(y)}\right) \tag{4}$$

There are two marginal probability distribution functions, $p_1(x)$ and $p_2(x)$, and $p(x, y)$ represents the combined probability distribution function.

As a convenient and easy way to express each attribute f_i, we will define it as a vector with N observations.

$$f_i = [f_i^1, f_i^2, \ldots f_i^N] \tag{5}$$

It is expressed in the form $I(F_i, F_j)$ where i = 1,2,...,d; j = 1,2,...,d represents the number of attributes in the dataset. The first condition is that S is set of attributes to be selected, |S| is number of elements of a cluster. As a second condition, S must comply with a minimum level of redundancy.

$$min W, W = \frac{1}{|S|^2} = \sum_{Fi, Fj \in S} I(F_i, F_j) \tag{6}$$

The maximum relevance condition:

$$max V, V = \frac{1}{|S|} = \sum_{Fi \in S} I(F_i, H) \tag{7}$$

A similarity result is known as $I(F_i, H)$. Two conditions can be expressed as follows:

$$max(V - W) \tag{8}$$

$$max(V / W) \tag{9}$$

The attribute that provides a solution for equality (10) and (11) is selected in each subsequent step.

$$max_{F_i \in \Omega_S} I(H, F_i) \tag{10}$$

$$min_{F_i \in \Omega_S} \frac{1}{|S|} \sum_{F_{i \in S}} I(F_i, F_j) \tag{11}$$

As a result of combining equality (10) and (11) with their respective criteria (8) and (9), the following selection criteria have been generated as seen in Table 2.

Table 2: Criterias of mRMR

Criterias of mRMR	
Criteria	Equations
Mutual information difference (MID)	$\max_{F_i \in \Omega_S} \left[I(F_i, H) - \dfrac{1}{|S|} \sum_{F_i \in S} I(F_i, F_j) \right]$
Mutual quotient information (MIQ)	$\max_{F_i \in \Omega_S} \left[I(F_i, H) / \dfrac{1}{|S|} \sum_{F_i \in S} I(F_i, F_j) \right]$

3 Results and Discussion

The data of PM10 concentration is collected as a dependent variable and solar irradiance is collected as an independent variable for regression analysis. This data covers a total of 43,791 h of weather analysis results and is collected at a frequency that matches the dynamics of the relationship we are investigating. The scatter plot seen in Fig. 7 visualizes the relationship between PM10 and solar irradiance. The regression formulation utilized in the figure is this:

$$r = \frac{n \sum xy - \left(\sum x \sum y \right)}{\sqrt{n \left(\sum x^2 \right) - \left(\sum x \right) 2} \sqrt{n \left(\sum y^2 \right) - \left(\sum y \right)^2}} \tag{12}$$

r = 0.072843 is an output of this regression and presents their power of plain relationship to evaluate the harmony. After examining plain relations then forecasting step is implemented.
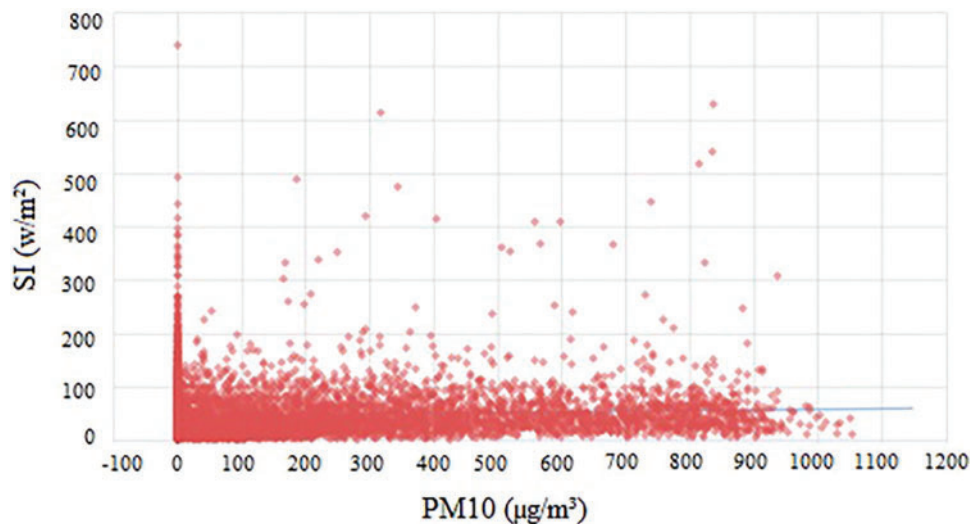

Figure 7: Regression analysis output between solar irradiance and PM10

There is also an mRMR analysis performed for each variable to determine how important it is for solar irradiance prediction which is seen in Table 3. PM25 concentration at time $t + 1$ has the greatest mutual information with solar irradiance at time t. Since mRMR considers irrelevant redundancies as well, other variables must have had a lower redundancy scores. According to the rankings, the next ranks are in the next positions, most likely because they appear to be redundant with SI.

Table 3: Prediction of SI: mRMR importance of input variables

	1	2	3	4	5	6	7	8	9	10
mRMR	PM25	PM10	CO	O_3	NO	NO_2	NOX	SO_2	MBAR	%H

The forecasting process is applied to the data between 01/01/2022 and 01/01/2023 to evaluate the goodness of fit of our model which is seen in Fig. 8. Understanding how much the solar irradiance is expected to change in this context will be revealed by interpreting the results for PM10 concentration. Mean absolute percentage error (MAPE = 6.93) confirms the statistical significance of the relationship, which is important for determining whether or not it is genuine.
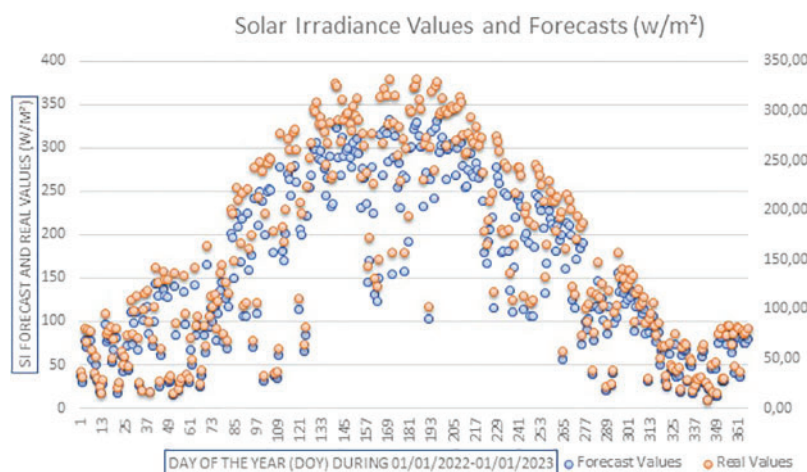

Figure 8: Solar Irradiance values and the forecast results obtained by the suggested model

The types of models in the literature are shown in Table 4 along with the corresponding input parameters, forecasting horizon, and time. The results of using the presented model for predicting solar irradiance values can also be seen in Table 5. A total of two error measures are used to evaluate the performance of the proposed models, including root mean square error (RMSE), and mean absolute percentage error (MAPE). Performance scores are derived based on the aforementioned metrics. The error measures RMSE and MAPE values of the Istanbul dataset by CNN-LSTM-mRMR approach are 5,251 and 6,863 in the fall & winter seasons, respectively. The values for the spring & summer seasons are a little bit higher as 7,654 and 8,659, respectively. Whole-year values are 5,536 and 6,935. The approach is robust based on these reliable scores.

Table 4: Solar irradiance forecasting model results in the literature

Authors	Input parameters	Size of local data required for training	RMSE (W/m^2)	Model
Qing et al. [8]	Wind speed Temperature Humidity	30 months	76.2	LSTM

(Continued)

Table 4 (continued)

Authors	Input parameters	Size of local data required for training	RMSE (W/m^2)	Model
Husein et al. [34]	Sky cover Temperature Humidity Wind speed Wind direction Precipitation Sky cover	168 months	60.3	LSTM
Yu et al. [35]	Irradiance Sky cover Temperature Precipitation Humidity Solar zenith angle Wind speed Wind direction	48 months	41.3	LSTM
Wang et al. [36]	Irradiance	101 months	32.4	CNN-LSTM
Cheng et al. [37]	Temperature Precipitation Humidity	2 months	NMAE 19%, 20%, 23%, respectively	LSTM
Proposed framework	Temperature Humidity Precipitation Atm. Pressure Wind speed PM10, PM25 NO, NOX NO$_2$, SO$_2$, O$_3$, CO	12 months	5,536	CNN-LSTM-mRMR

Table 5: Measures of prediction from different models compared

Country	Model	MAPE	RMSE
Solar irradiance values (in fall&winter)	CNN-LSTM-mRMR	6,863	5,251
Solar irradiance values (in spring&summer)	CNN-LSTM-mRMR	8,659	7,654
Solar irradiance values (whole year)	CNN-LSTM-mRMR	6.935	5,536

4 Conclusions

The present study investigates the effect of air pollutants on solar irradiance using a novel forecasting approach. A thorough analysis of the impact of weather conditions, greenhouse gases, and dust on solar energy systems is the backbone of the search. Solar irradiance could be enhanced to a great extent by improving air quality, as shown by MAPE and RMSE results, 6,935 and 5,536, respectively, and this speeds up the adoption of renewable energy. In other words, air pollution has a strong correlation with reducing effectiveness, particularly in the form of a decrease in the intensity

of sunlight, which can have a significant impact on those who invest their capital and wait a short period for a return on investment (ROI). The time of ROI increases, and so does the cost of investing in such energy systems. Furthermore, low effectiveness can discourage government support for renewable energy investments by limiting incentives, like tax credits and rebates, for investors to switch to a more affordable and environmentally friendly mode of green energy.

It is clear from the study that air pollution has a remarkable impact on solar energy systems. A clean environment of decreasing air pollution is prompting the transition to green energy. The nature of the relationship may be complex, and further analysis or additional variables may be needed for a comprehensive understanding. A future study should focus on additional air data, behavioral characteristics of another gas, and using this data to draw meaningful conclusions.

Time/day	Time/hour	CO ($\mu g/m^3$)	Pressure (mbar)	NO ($\mu g/m^3$)	NO_2 ($\mu g/m^3$)	NOX ($\mu g/m^3$)	Humidity (%)	O_3 ($\mu g/m^3$)	PM_{10} ($\mu g/m^3$)	PM_{25} ($\mu g/m^3$)	SO_2 ($\mu g/m^3$)
1.01.2022	00:00	412,2	1021,3	14,6	27,6	50,1	82,9	24,4	30,4	22,4	2,5
2.01.2022	00:00	643,2	1025,1	51,3	33,2	111,9	84,4	15,5	34	26,1	2,8
3.01.2022	00:00	357,2	1020,6	30,7	31,8	78,8	77,5	20,9	30,8	20,8	1,9
4.01.2022	00:00	389,6	1017,4	42,7	37,6	103	79,6	26,7	27,1	20,2	2,3
5.01.2022	00:00	362,3	1014,2	14	29,5	51	72,8	23,6	29,3	19,2	3,5
6.01.2022	00:00	529,5	1014,6	10,8	24,2	40,7	75,9	20,3	21,4	13,3	1,8
7.01.2022	00:00	724	1019,2	81,6	33,5	158,6	84,2	16,6	40,3	25	1,3
8.01.2022	00:00	632,2	1016,7	35,5	29	83,6	87	19	24,7	21,4	0,9
9.01.2022	00:00	550,6	1008,7	63,8	36,6	134,4	84,4	23,6	48,1	38,8	1,9
10.01.2022	00:00	286,3	1004,1	18,2	29,8	57,8	85,1	25	27,3	23,7	1,5
11.01.2022	00:00	519	1011,2	11,4	17,7	35	86,3	30,7	17,2	14,3	1,2
12.01.2022	00:00	424,7	1022,1	6,4	17,9	27,7	85,1	22,6	2,7	7,5	1,3
13.01.2022	00:00	681,5	1032,4	7	22,6	33,3	56,9	23,8	9,7	9,5	1,5
14.01.2022	00:00	706,9	1026,7	15,7	29,9	54,1	31	15,6	18,3	16,9	3
15.01.2022	00:00	871	1020,8	40,2	40,1	101,5	33,4	11,6	24	18,9	2
16.01.2022	00:00	535,7	1027,7	56,6	41,7	128,5	32	24,5	24	20,3	2,2
17.01.2022	00:00	738,7	1023,8	78,7	32,5	153,2	32,2	22,9	48,9	30,3	2,7
18.01.2022	00:00	784,1	1021,8	23,1	31,3	66,4	32,4	17,8	13	11,1	2,6
19.01.2022	00:00	710,2	1030,7	25,5	30,7	69,9	27,1	28,9	17,5	12,7	1,2
20.01.2022	00:00	801,9	1027,1	9,4	17,9	32,3	28,7	29,2	21,1	15,6	2
21.01.2022	00:00	779,6	1019,6	7,7	26,2	38	37,9	21,2	14	11,5	1,8
22.01.2022	00:00	685,7	1019	11,2	32,2	49,2	32,2	19,1	10,5	10,7	2,1

Please visit the link for all data: https://havakalitesi.ibb.gov.tr/STN/VWSTN_Reports/.

Author Contributions: The author confirms contribution to the paper as follows: Conceptualization, formal analysis, methodology and implementation, analysis and interpretation of results, writing—original draft, review and editing are made by Ramiz Görkem Birdal. The author reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: Data available on request from the authors.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] R. Heffron *et al.*, "Justice in solar energy development," *Sol. Energy*, vol. 218, pp. 68–75, 2021. doi: 10.1016/j.solener.2021.01.072.

[2] T. S. Ge *et al.*, "Solar heating and cooling: Present and future development," *Renew. Energy*, vol. 126, pp. 1126–1140, 2018. doi: 10.1016/j.renene.2017.06.081.

[3] M. A. Gulzar, H. Asghar, J. Hwang, and W. Hassan, "China's pathway towards solar energy utilization: Transition to a low-carbon economy," *Int. J. Environ Res. Publ. Health.*, vol. 17, no. 12, pp. 4221, 2020. doi: 10.3390/ijerph17124221.

[4] M. Salimi, M. Hosseinpour, and T. N. Borhani, "Analysis of solar energy development strategies for a successful energy transition in the UAE," *Process.*, vol. 10, no. 7, pp. 1338, 2022. doi: 10.3390/pr10071338.

[5] P. Kumari and D. Toshniwal, "Impact of lockdown measures during COVID-19 on air quality–A case study of India," *Int. J. Environ Health. Res.*, pp. 1–8, 2020. doi: 10.1080/09603123.2020.1778646.

[6] W. de Soto, S. A. Klein, and W. A. Beckman, "Improvement and validation of a model for photovoltaic array performance," *Sol. Energy*, vol. 80, no. 1, pp. 78–88, 2006. doi: 10.1016/j.solener.2005.06.010.

[7] A. Dolara, S. Leva, and G. Manzolini, "Comparison of different physical models for PV power output prediction," *Sol. Energy*, vol. 119, pp. 83–99, 2015. doi: 10.1016/j.solener.2015.06.017.

[8] X. Qing and Y. Niu, "Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM," *Energy*, vol. 148, pp. 461–468, 2018. doi: 10.1016/j.energy.2018.01.177.

[9] S. Srivastava and S. Lessmann, "A comparative study of LSTM neural networks in forecasting day-ahead global horizontal irradiance with satellite data," *Sol. Energy*, vol. 162, pp. 232–247, 2018. doi: 10.1016/j.solener.2018.01.005.

[10] M. Ding, L. Wang, and R. Bi, "An ANN-based approach for forecasting the power output of photovoltaic system," *Procedia Environ. Sci.*, vol. 11, pp. 1308–1315, 2011. doi: 10.1016/j.proenv.2011.12.196.

[11] G. Cervone, L. Clemente-Harding, S. Alessandrini, and L. Delle Monache, "Short-term photovoltaic power forecasting using artificial neural networks and an analog ensemble," *Renew. Energy*, vol. 108, no. 1, pp. 274–286, 2017. doi: 10.1016/j.renene.2017.02.052.

[12] A. Mellit, S. Sağlam, and S. A. Kalogirou, "Artificial neural network-based model for estimating the produced power of a photovoltaic module," *Renew. Energy*, vol. 60, pp. 71–78, 2013. doi: 10.1016/j.renene.2013.04.011.

[13] M. Khodayar and J. Wang, "Spatio-temporal graph deep neural network for short-term wind speed forecasting," *IEEE Trans. Sustain. Energy*, vol. 10, no. 2, pp. 670–681, 2018. doi: 10.1109/TSTE.2018.2844102.

[14] A. Muhammad, J. M. Lee, S. W. Hong, S. J. Lee, and E. H. Lee, "Deep learning application in power system with a case study on solar irradiation forecasting," in *2019 Int. Conf. Artif. Intell. Inf. Commun. (ICAIIC)*, 2019, pp. 275–279.

[15] S. Mishra and P. Palanisamy, "An integrated multi-time-scale modeling for solar irradiance forecasting using deep learning," 2019. arXiv preprint arXiv:1905, 02616.

[16] T. P. Chu, J. H. Jhou, and Y. G. Leu, "Image-based solar irradiance forecasting using recurrent neural networks," in *2020 Int. Conf. Sys. Sci. and Eng.(ICSSE)*, 2020, pp. 1–4.

[17] G. Guariso, G. Nunnari, and M. Sangiorgio, "Multi-step solar irradiance forecasting and domain adaptation of deep neural networks," *Energies*, vol. 13, no. 15, pp. 3987, 2020. doi: 10.3390/en13153987.

[18] A. Mukherjee, A. Ain, and P. Dasgupta, "Solar irradiance prediction from historical trends using deep neural networks," in *2018 IEEE Int. Conf. on Smart Energy Grid Eng.*, 2018, pp. 356–361.

[19] T. A. Farrag and E. E. Elattar, "Optimized deep stacked long short-term memory network for long-term load forecasting," *IEEE Access*, vol. 9, pp. 68511–68522, 2021. doi: 10.1109/ACCESS.2021.3077275.

[20] M.Ş. Özçoban, M. E. Isenkul, S. Sevgen, S. Acarer, and M. Tüfekci, "Modelling the effects of nanomaterial addition on the permeability of the compacted clay soil using machine learning-based flow resistance analysis," *Applied Sciences*, vol. 12, no. 1, pp. 186, 2021. doi: 10.3390/app12010186.

[21] D. Chandola, H. Gupta, V. A. Tikkiwal, and M. K. Bohra, "Multi-step ahead forecasting of global solar radiation for arid zones using deep learning," *Procedia Comp. Sci.*, vol. 167, pp. 626–635, 2020. doi: 10.1016/j.procs.2020.03.329.

[22] X. Zhao, H. Wei, H. Wang, T. Zhu, and K. Zhang, "3D-CNN-based feature extraction of ground-based cloud images for direct normal irradiance prediction," *Sol. Energy*, vol. 181, pp. 510–518, 2019. doi: 10.1016/j.solener.2019.01.096.

[23] F. Wang *et al.*, "Deep learning based irradiance mapping model for solar PV power forecasting using sky image," in *2019 IEEE Industry App. Society Annual Meeting*, 2019, pp. 1–9.

[24] O. El Alani, M. Abraim, H. Ghennioui, A. Ghennioui, I. Ikenbi and F. E. Dahr, "Short term solar irradiance forecasting using sky images based on a hybrid CNN-MLP model," *Energy. Reports.*, vol. 7, pp. 888–900, 2021. doi: 10.1016/j.egyr.2021.07.053.

[25] Y. LeCun, K. Kavukcuoglu, and C. Farabet, "Convolutional networks and applications in vision," in *Proc. 2010 IEEE Int. Symp. Circuits Syst.*, 2010, pp. 253–256.

[26] S. Song, Z. Yang, H. Goh, Q. Huang, and G. Li, "A novel sky image-based solar irradiance nowcasting model with convolutional block attention mechanism," *Ener. Reports.*, vol. 8, pp. 125–132, 2022. doi: 10.1016/j.egyr.2022.02.166.

[27] J. Che, Y. Yang, L. Li, X. Bai, S. Zhang and C. Deng, "Maximum relevance minimum common redundancy feature selection for nonlinear data," *Inform. Sci.*, vol. 409, pp. 68–86, 2017. doi: 10.1016/j.ins.2017.05.013.

[28] P. Verma and H. Om, "MCRMR: Maximum coverage and relevancy with minimal redundancy based multi-document summarization," *Expert. Syst. Appl.*, vol. 120, pp. 43–56, 2019. doi: 10.1016/j.eswa.2018.11.022.

[29] M. Cıbuk, U. Budak, Y. Guo, M. C. Ince, and A. Sengur, "Efficient deep features selections and classification for flower species recognition," *Measurement.*, vol. 137, pp. 7–13, 2019. doi: 10.1016/j.measurement.2019.01.041.

[30] A. R. Webb, "Clustering," in *Statistical Pattern Recognition*, Hoboken NJ: John Wiley & Sons, 2003, pp. 361–407.

[31] Y. Akbulut, A. Sengur, Y. Guo, and F. Smarandache, "NS-k-NN: Neutrosophic set-based k-nearest neighbors classifier," *Symmetry*, vol. 9, no. 9, pp. 179, 2017. doi: 10.3390/sym9090179.

[32] T. Hastie, R. Tibshirani, J. H. Friedman, and J. H. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. vol. 2. New York: Springer, pp. 1–758, 2009.

[33] J. I. E. Hoffman, "Linear regression," in J. I. E. Hoffman (Ed.), *Basic Biostatistics for Medical and Biomedical Practitioners*, 2nd ed. Academ. Press, 2019, pp. 445–489.

[34] M. Husein and I. Y. Chung, "Day-ahead solar irradiance forecasting for microgrids using a long short-term memory recurrent neural network: A deep learning approach," *Energies*, vol. 12, no. 10, pp. 1856, 2019. doi: 10.3390/en12101856.

[35] Y. Yu, J. Cao, and J. Zhu, "An LSTM short-term solar irradiance forecasting under complicated weather conditions," *IEEE Access*, vol. 7, pp. 145651–145666, 2019. doi: 10.1109/ACCESS.2019.2946057.

[36] F. Wang, Y. Yu, Z. Zhang, J. Li, Z. Zhen and K. Li, "Wavelet decomposition and convolutional LSTM networks based improved deep learning model for solar irradiance forecasting," *Appl. Sci.*, vol. 8, no. 8, pp. 1286, 2018. doi: 10.3390/app8081286.

[37] H. Y. Cheng, C. C. Yu, and C. L. Lin, "Day-ahead to week-ahead solar irradiance prediction using convolutional long short-term memory networks," *Renew. Energy*, vol. 179, pp. 2300–2308, 2021. doi: 10.1016/j.renene.2021.08.038.