forecasting

MDPI

Article

Assessment of Deep Neural Network Models for Direct and Recursive Multi-Step Prediction of PM10 in Southern Spain

Javier Gómez-Gómez *, Eduardo Gutiérrez de Ravé and Francisco J. Jiménez-Hornero

GEPENA Research Group, Gregor Mendel Building (3rd Floor), Campus Rabanales, University of Cordoba, 14071 Cordoba, Spain; eduardo@uco.es (E.G.d.R.); fjhornero@uco.es (F.J.J.-H.)
* Correspondence: f12gogoj@uco.es

Abstract: Western Europe has been strongly affected in the last decades by Saharan dust incursions, causing a high PM10 concentration and red rain. In this study, dust events and the performance of seven neural network prediction models, including convolutional neural networks (CNN) and recurrent neural networks (RNN), have been analyzed in a PM10 concentration series from a monitoring station in Córdoba, southern Spain. The models were also assessed here for recursive multi-step prediction over different forecast periods in three different situations: background concentration, a strong dust event, and an extreme dust event. A very important increase in the number of dust events has been identified in the last few years. Results show that CNN models outperform the other models in terms of accuracy for direct 24 h prediction (RMSE values between 10.00 and 10.20 $\mu g/m^3$), whereas the recursive prediction is only suitable for background concentration in the short term (for 2–5-day forecasts). The assessment and improvement of prediction models might help the development of early-warning systems for these events. From the authors' perspective, the evaluation of trained models beyond the direct multi-step predictions allowed to fill a gap in this research field, which few articles have explored in depth.

Keywords: PM10; Saharan dust; neural networks; convolutional neural networks; recurrent neural networks; recursive prediction; multi-step prediction

check for updates

Academic Editor: Jun A. Zhang

Received: 20 December 2024
Revised: 18 January 2025
Accepted: 21 January 2025
Published: 26 January 2025

Citation: Gómez-Gómez, J.; Gutiérrez de Ravé, E.; Jiménez-Hornero, F.J. Assessment of Deep Neural Network Models for Direct and Recursive Multi-Step Prediction of PM10 in Southern Spain. *Forecasting* **2025**, *7*, 6. https://doi.org/10.3390/forecast7010006

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Particulate matter is of great importance because of its negative impacts on human health, such as the exacerbation of chronic respiratory and cardiovascular diseases, conjunctivitis, skin irritation, and others [1,2]. African dust incursions in Western Europe have recently attracted the attention of the scientific community for their influence on the increase of PM10 in the atmosphere in several countries [3–7] and their growing frequency and intensity in recent decades, particularly in the period of 2020–2022 [3]. March 2022 was characterized by an extremely exceptional Saharan dust outbreak, which derived from the transportation of a plume of dust that crossed the Mediterranean into Western Europe [8]. In this period, the PM10 concentrations over Spain reached unprecedented record levels and caused red rain, especially visible in the Sierra Nevada and the Pyrenees [3,6]. Cuevas-Agulló et al. found that more than half of dust events in February and March are accompanied by cut-off lows, normally located between the Canary Islands and the Iberian Peninsula [3]. However, these concerning phenomena are not clear whether they belong to exceptional cases or are due to climate change. There are discrepancies in the literature for reasons such as the detection indices, the large temporal variations, and the short records of these events [9,10].

In the last two decades, a growing number of studies based on artificial neural networks have designed and proposed different models to improve the prediction of PM10 concentration series [11] on a daily [12–16], 6-h [17], hourly [12,18–25] and even 1-min [26] basis. Artificial neural networks are a set of processing units or perceptrons disposed of in parallel layers with the ability to adaptatively recognize patterns from the input data [12,27,28]. They generate output results that increase their predictive accuracy as the network learns these patterns. The structure of a perceptron or neuron is composed of three simple elements: a set of weighted inputs, an activation function, and an output [18]. These weights determine the strength of connections between the neurons and, therefore, networks essentially learn throughout their adaptation [19]. The most popular adjustment process of weights in neural networks is the back propagation algorithm developed by Rumelhart et al. [29]. This algorithm uses the steepest gradient descent method to correct the weights by error convergence of the output for a given input, propagating the error backward throughout the network up to the input layer [27,29].

The simplest architecture is the feed-forward multilayer perceptron (MLP), an arrangement of one or more hidden layers of neurons between the input and the output layers with non-linear activation functions that do not contain any loop or feedback connection [19,28]. As explained by Díaz-Robles et al., in this model, the effective incoming signal to a neuron in a particular layer is the weighted sum of all the incoming signals, sometimes added to an extra bias, and then the activation function is applied to the result producing the output of that neuron [12,28]. Other architectures include the convolutional neural networks (CNN) designed for image processing and based on convolutional operations [30,31] and recurrent neural networks (RNN), which learn patterns in sequential data by loops or recurrent feeding of previously processed time steps and using memory cells [18,32–34].

Córdoba is a city situated in southern Spain in a region characterized by a continental Mediterranean climate with dry and hot summers and very infrequent red rain phenomena, which were traditionally associated with summer storms. In the last decade, they have become more and more common, emerging in other seasons coinciding with the increasingly frequent Saharan dust incursions, which have a prominent impact in this area. This is a consequence of the efficiency of dust particles as condensation nuclei forcing rainfall [35,36]. In this study, the hourly PM10 concentration recorded in a monitoring air quality station in Córdoba has been analyzed and used for training prediction models based on neural networks.

Few articles have explored in depth the limitations of pre-trained models for longer forecast periods beyond the initial boundary of the outputs by recursive procedures, and there are even fewer that study the PM10 series. Thus, Gilik et al. (2022) [37] showed that convolutional neural networks (CNNs) exhibit strong robustness in capturing spatial dependencies in PM10 datasets, particularly in regions affected by dust events. Their models achieved high accuracy in short-term PM10 predictions, making them effective for monitoring and forecasting in spatially variable environments. Wang et al. (2022) [38] state that RNNs, specifically bidirectional long short-term memory (BiLSTM) networks, demonstrated superior accuracy in predicting PM10 trends over extended periods compared to other models. Their ability to capture long-term temporal dependencies allowed them to outperform simpler architectures, particularly in scenarios with sequential data. However, the study also highlighted that BiLSTM models require more computational resources and longer training times, making them less suitable for real-time applications without optimization. Hu et al. [26] developed and evaluated a smart warning model for predicting PM10 concentration exceedances at construction sites, leveraging a long short-term memory (LSTM) neural network and time-series processing techniques. The

model achieves multi-step predictions by iteratively using prior predictions as inputs for subsequent predictions.

While studies on recursive multi-step predictions for PM10 are limited, related fields such as meteorology and energy demand forecasting have extensively explored the application of neural network architectures, including convolutional and recurrent neural networks. For instance, Mishra and Desai (2006) [27] highlighted that recursive and direct multi-step neural networks significantly outperform traditional ARIMA models in drought forecasting using the Standardized Precipitation Index. Shi et al. (2015) [39] demonstrated the effectiveness of convolutional models for precipitation nowcasting, highlighting their ability to capture spatiotemporal dependencies in weather data. Similarly, Cai et al. (2019) [32] applied CNNs and RNNs for day-ahead building-level energy demand forecasting, showcasing their robustness in handling sequential patterns and long-term dependencies. These studies provide valuable insights into the strengths and limitations of similar models, particularly in handling sequential data and predicting complex temporal patterns, which are relevant to PM10 forecasting.

Hence, the aim of this study is to analyze the dust events, train neural network models, and explore the limitations of recursive multi-step predictions in the monitoring station of Córdoba. This study aims to systematically evaluate the performance of established neural network architectures in forecasting PM10 levels under varying conditions, focusing on their strengths and limitations in predicting background, strong, and extreme dust events.

Two experiments are proposed to assess the forecasting of the PM10 series: training of neural network models with 24 h ahead (multi-step) prediction and their extrapolation beyond these limits. After preliminary analysis and identification of PM10 concentration peaks, the first proposed experiment, which includes the design and training of seven neural networks based on CNN and RNN architectures, is performed. Next, the second experiment consists of the extrapolation of the trained network models by recursive prediction in three different situations, including background concentration levels, one strong dust event, and one extreme dust event, which are tested over six different forecast periods, with 2 to 7-day forecasts.

The paper is organized into the following sections: Section 2 explains the data acquisition and preprocessing, the imputation of missing data, the criteria for the identification of dust events, CNN and RNN architectures, and the experimental setup. Section 3 discusses the results obtained and the performance of neural network models. Lastly, Section 4 summarizes the main conclusions drawn.

2. Materials and Methods

2.1. Data Source and Study Area

The hourly PM10 concentration data from two open-access databases were used: one includes the historical data on air quality in Spain sourced from the official website of the Spanish Ministry for the Ecological Transition and Demographic Challenge (MITECO, initials in Spanish) [40] and the other is the air quality reports database from the Environmental Portal of Andalusia (PAA, initials in Spanish) [41]. The area of study covers the city of Córdoba in the region of Andalusia, in the south of the Iberian Peninsula, and in the southwestern part of Europe.

The PM10 concentration records were obtained from stations belonging to the Air Quality Monitoring and Control Network of Andalusia. Four monitoring air quality stations, named Asomadilla (ASLLA), Avda. Al-Nasir (AVAN), Parque Joyero, and Lepanto, are in this area. While all stations were initially considered, only the ASLLA station was used in the experiments, as it better represents the background PM10 concentrations with

minimal influence from local traffic, unlike the AVAN station, which is in an urban area heavily influenced by traffic emissions (see Table 1).

Table 1. Characterization of the selected air quality stations and basic descriptive statistics of hourly PM10 concentration series.

Station	Latitude (°N)	Longitude (°W)	Type of Station	Type of Area	Mean (μg/m^3)	SD (μg/m^3)	Missing Data (%)
ASLLA	37.90	−4.78	Background	Suburban	22.10	18.86	2.98
AVAN	37.89	−4.78	Traffic	Urban	28.26	20.98	7.31

2.2. Data Preprocessing

To obtain the longest available time series, some stations were discarded. At the time of data collection, the longest available series obtained from the combination of MITECO and PAA databases covers the period from 2015 to May 2024 for stations ASLLA and AVAN. In particular, the data gaps from the MITECTO database were filled with those data obtained from the PAA database.

Table 1 shows basic information from the selected stations ASLLA and AVAN and their records. As can be seen in the table, the finally retrieved raw data have less than 10% missing values in the study period.

To fill missing values and gaps, the linear regression imputation method based on the work of Barrero-González et al. was used [18]. This method consists of the computation of a series of linear regressions to make reiterated imputations between the dependent and independent variables in the case of multiple linear regression. The first imputation has been made by substitution of missing values with random values from normal distributions with μ and σ parameters equal to the mean and standard deviation, respectively, of each variable. Imputations of completely missing rows were addressed by adding a second random number from a uniform distribution in the interval $(-\sigma, \sigma)$ for each variable before the application of linear regressions to avoid regions with high stability, which could skew the training [18]. The authors set a maximum number of 60 regression imputations and a stopping criterion based on the root mean squared error (RMSE) between iterations lower than 0.001 for each variable. Results can be seen in Table 2.

Table 2. RMSE between the two last regression imputations and descriptive statistics of the imputed series.

Station	RMSE of 41th Imputation (μg/m^3)	μ (μg/m^3)	Median	σ (μg/m^3)	Range
ASLLA	8.59×10^{-4}	22.11	19.00	18.71	0–1116
AVAN	7.58×10^{-4}	28.41	25.00	21.19	1–1044

For the subsequent analyses, the authors focused on the ASLLA station, as it is less influenced by traffic emissions and provides a clearer representation of the background PM10 concentrations (see Table 1). Data from the AVAN station, although not directly used in the experiments, contributed to the preprocessing phase by filling gaps in the ASLLA dataset. This complementary use of AVAN data ensured the completeness and reliability of the ASLLA time series.

2.3. Dust Events

To describe dust events at the ASLLA station, some imposed criteria are defined to identify and classify them, similar to the study on aerosol optical depth of Cuevas-Agulló

et al. [3]. Table 3 provides the criteria used for this classification and the corresponding computed thresholds. Figure 1 depicts the reconstructed series of the ASLLA station, where the strong and extreme dust events are highlighted.

Table 3. PM10 concentration thresholds and criteria used to classify the intensity of dust events.

Type of Dust	Criteria	PM10 Thresholds (μg/m^3)
Background	$PM10 < \mu + 2\sigma$	$PM10 < 59$
Strong event	$\mu + 2\sigma \leq PM10 \leq \mu + 4\sigma$	$60 \leq PM10 \leq 97$
Extreme event	$PM10 > \mu + 4\sigma$	$PM10 > 97$

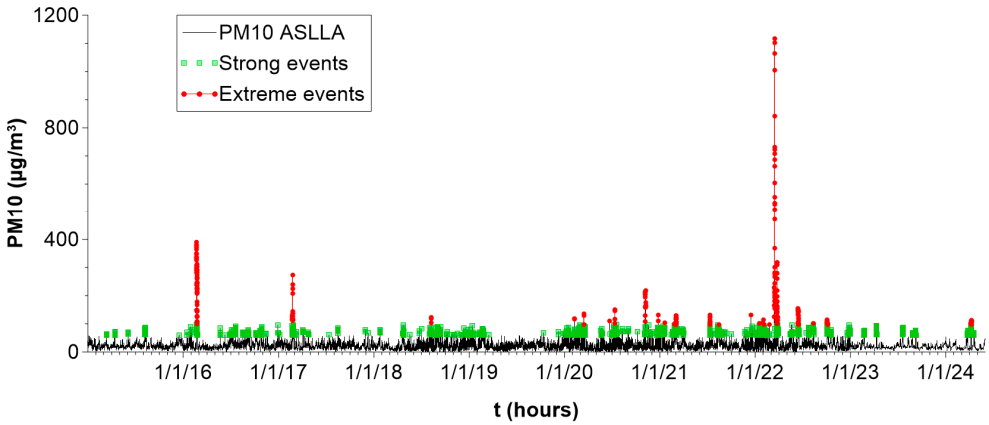


Figure 1. Hourly PM10 concentration series (black line), strong dust events (green squares), and extreme dust events (red circles and lines).

2.4. Neural Network Architectures

2.4.1. Convolutional Neural Networks (CNN)

The one-dimensional convolutional neural network (1D CNN) architecture is based on the conventional 2D CNN, which was developed for image processing and has shown a remarkable performance in the field of computer vision [30]. It contains neurons that apply convolution computation to the inputs [31,32]. 1D CNNs are a flattened version of 2D CNNs for pattern recognition in time-series data and are applied over the time dimension of the dataset.

2.4.2. Recurrent Neural Networks (RNN)

Recurrent neural networks (RNN) are more complex neural network architectures that consist of dynamic models designed to learn from sequential data by recurrently processing the inputs [18,32]. They can be trained one step at a time using the outputs to feed the inputs [18,32,33].

LSTM units are RNNs that were developed to address the gradient vanishing problem that arises in these networks [42], improving their capacity to store the previous data information and their long-range dependencies. This functionality is due to their architecture. An LSTM cell is composed of four neurons, unlike the conventional RNN, which only has one [17,32,34]. These neurons are the memory cells (c_t) and the three sigmoid or gating functions: the input gate (i_t), the forget gate (f_t) and the output gate (o_t). The hidden layer function of the LTSM cell is the following composite function [32]:

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \quad (1)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \quad (2)$$

$$f_t = \sigma\left(W_{xf}x_t + W_{hf}h_{t-1} + b_f\right) \quad (3)$$

$$c_t = f_t * c_{t-1} + i_t * tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_C) \quad (4)$$

$$h_t = o_t * tanh(c_t) \quad (5)$$

where x_t is the input vector at each time step; h_t is the hidden vector at the current time step, which has the same size as the i_t, f_t, o_t, c_t activations vectors; W_{ij} are the weight matrices connecting elements i and j; b_j are the bias terms for vector j; σ are sigmoid functions and $*$ represents the pointwise multiplication.

On the other hand, gated recurrent units (GRU) are other classes of RNNs and simpler variants of LSTMs that combine input and forget gates into update gates and also contain a reset gate [17,20,34]. This architecture was proposed by Cho et al. [43]. The hidden layer function of GRUs is defined as follows [17,20]:

$$z_t = \sigma(W_{xz}x_t + W_{hz}h_{t-1} + b_z) \quad (6)$$

$$r_t = \sigma(W_{xr}x_t + W_{hr}h_{t-1} + b_r) \quad (7)$$

$$c_t = tanh(W_{xc}x_t + W_{hc}(r_t * h_{t-1}) + b_c) \quad (8)$$

$$h_t = r_t * h_{t-1} + (1 - r_t) * c_t \quad (9)$$

where x_t, W_{ij}, b_j, and h_t have the same meaning as in Equations (1)–(5); z_t is the update gate vector; r_t is the reset gate vector and c_t is the memory cell vector (also the previous state).

2.5. Experimental Setup and Neural Network Models

After the analysis of dust events, the authors designed two experiments using different neural network models: the evaluation of 24 h ahead multi-step prediction models based on different neural network architectures and the assessment of recursive multi-step forecasts using predictions as inputs in the subsequent iteration.

Neural network models were designed, trained, and recursively used for forecasting using Python libraries tensorflow and keras [44,45] in a computer with the following specifications: Intel® Core™ i7-6700 CPU 3.40 GHz and RAM memory 16.0 GB. The total computation time reached approximately 9 h. The multiple hyperparameter optimizations contributed to the extended training time. This effort ensured the models were appropriately tuned for the dataset and task. While decision tree (DT) based methods, such as gradient boosting or random forests, could potentially reduce training time, they may not achieve comparable performance in tasks requiring temporal feature extraction. Their main limitation is that they grow and become too complex, and high variance and low bias can be observed, leading to overfitting [46].

2.5.1. Evaluation of Multi-Step Prediction Models

The evaluation of 24 h ahead multi-step prediction models with different hyperparameters was made by means of training, validation, and test phases. A subset from the PM10 concentration series with a total of $N = 61,368$ data from 2015 to 2021 was used as the dataset for this task. During the training phase, models were optimized to produce direct 24-h ahead predictions based on lagged PM10 values and seasonal features. Recursive multi-step predictions were not used during training or validation but were exclusively applied during testing to evaluate model performance over extended forecast periods.

As shown by Chaloulakou et al. [13,14,24], pollutant predictive models seem to benefit from using persistence information, i.e., lagged concentration values, as inputs. Grivas and Chaloulakou determined by a genetic algorithm optimization procedure that the 24 h lagged PM10 concentration was one of the most useful predictors [24]. Other authors found

similar results, also including PM10 concentrations lagged by 25 and 26 h [16,19]. In this study, 1 h and 2 h lagged PM10 concentrations were proposed as inputs. On the other hand, the Fourier transform of PM10 concentration in the ASLLA station revealed biannual and daily cycles as the two most important frequencies. Thus, sinusoidal and cosinusoidal variables per cycle were also aggregated as inputs, similarly to previous studies [19,24], making a total of 7 features in the dataset. Next, this dataset was split into the training (70%), validation (20%), and test (10%) sets, and the three sets were normalized to the mean and standard deviation of the training set to avoid prior information of other sets in this one. For consistency, the dimensions of inputs and outputs were 24 h × 7 features in every model for each sample.

CNN and autoregressive LSTM (AR LSTM) and GRU (AR GRU) architectures, designed by wrapping the LSTM and GRU cells in RNN layers, respectively, were implemented in Python as follows [47]:

- CNN:
  - Down-sampling layer, which extracts the last time steps from inputs with a length equal to the kernel size k of the Conv1D layer.
  - Conv1D layer with m filters and kernel size k.
  - Dense layer with 24 × 7 output units initialized with null tensors.
  - Reshape layer, which reshapes the dimensions of the output.
- AR LSTM/AR GRU:
  - LSTM/GRU cell with m units embedded in a RNN layer for the first time step [48].
  - Dense layer with 7 outputs.
  - One LSTM/GRU cell fed with the previous network states followed by the same dense layer for each subsequent time step from step 2 to 24.
  - Stacking of predictions in 24 × 7 tensors.

In this study, the selected measurement error for adjustments in the training and validation phases was the mean squared error (MSE) and training stops according to the early stopping technique [24,44,45]. A maximum number of 1000 training loops or "epochs" was set. A total of seven models were configured. The rest of the hyperparameters can be found in Table 4. The Adam optimizer algorithm was used in every model [49].

Table 4. Model configuration.

Model	Patience	Kernel Size (k)	Batch Size	Filters/Units (m)	Learning Rate	Optimizer Algorithm
CNN 1	500	3	32	64	0.01	Adam
CNN 2	800	3	32	64	0.01	Adam
CNN 3	800	6	32	64	0.01	Adam
AR LSTM 1	500	-	32	64	0.01	Adam
AR LSTM 2	800	-	32	64	0.01	Adam
AR GRU 1	500	-	32	64	0.01	Adam
AR GRU 2	800	-	32	64	0.01	Adam

Four different performance metrics were used for assessing the models: the fractional bias (FB) [50], the index of agreement (IA) [51], the mean absolute error (MAE) [52], and the RMSE [53].

FB quantifies the bias in predictions by comparing the mean of predicted and observed values. It ranges from −2 to +2, with 0 indicating no bias,

$$\mathrm{FB} = \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)}{\frac{1}{2}\left(\sum_{i=1}^{N} y_i + \sum_{i=1}^{N} \hat{y}_i\right)} \tag{10}$$

where y_i and $\hat{y}_i$ are the observed and predicted values, respectively, and N refers to the total number of observed-predicted pairs. It determines whether the model systematically overpredicts (positive FB) or underpredicts (negative FB) the target variable.

IA assesses the degree of agreement between observed and predicted values. It ranges from 0 (no agreement) to 1 (perfect agreement),

$$\mathrm{IA} = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(|\hat{y}_i - \overline{y}| + |y_i - \overline{y}|)^2} \tag{11}$$

where $\overline{y}$ is the mean of the observed values. *IA* evaluates how well the predictions match observed values, considering both magnitude and distribution.

MAE represents the average absolute difference between observed and predicted values. Unlike RMSE, it treats all errors equally without penalizing larger deviations more heavily.

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \tag{12}$$

It provides a straightforward measure of prediction accuracy, useful for understanding typical prediction errors.

RMSE measures the average magnitude of the errors between observed and predicted values. It gives greater weight to larger errors due to squaring the differences, making it sensitive to outliers.

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{13}$$

This metric evaluates overall prediction accuracy, with lower RMSE values indicating better model performance.

2.5.2. Assessment of Prediction Models for Recursive Multi-Step Forecasting

The assessment of the previously trained models for recursive multi-step forecasts was designed by feeding the previous 24 h predictions as inputs in each subsequent iteration [32]. Three different situations with periods of seven days were selected from the remaining subseries which covers the years 2022–2024: one case of PM10 background levels, one strong dust event, and one extreme dust event. The last day of every 7-day period (matching the PM10 concentration peaks) was predicted by 2–7-day forecasts (2–7 recursive iterations), and the performance of every model was evaluated with the same metrics used in Section 2.5.1.

3. Results

3.1. Analysis of Dust Events

The results of a simple analysis of the identified dust events are in Table 5. There was a number of strong events significantly higher than extreme events between 2015 and May 2024. Note that strong and extreme events are not mutually exclusive since a strong event might derive into an extreme event and vice versa. Both kinds of events show a variable duration from one hour up to a little more than 2 days. Moreover, the analysis of differences between subperiods 2015–2019 and 2020–2024 shows that these events have increased in the last years in this area (see also Figure 1), with a pronounced growth rate of 1166.67% in the case of extreme events. The concerning rise in the frequency of PM10 concentration peaks

is in agreement with the results obtained by Cuevas-Agulló et al. in their study of aerosol optical depth [3]. A very unusual and particularly relevant Saharan dust event, clearly identified as an extreme event in Figure 1 on 15–25 March 2022, raised PM10 concentration levels throughout most of western Europe, reaching the United Kingdom [3,8].

Table 5. Basic information of the identified events.

Type of Dust	Number of Events (2015–2024)	Event Duration (h)	Growth Rate 2015–2019 and 2020–2024 (%)
Strong event	399	1–50	157.41
Extreme event	38	1–53	1166.67

3.2. Performance of the Trained Models

The quality of the trained models was assessed with the validation and test sets. All models successfully predicted the observed values in the test set, as shown in Figure 2, where both kinds of dust events can be found. The final number of training loops and all computed metrics can be found in Table 6. These metrics correspond to the errors measured for the PM10 concentration variable and are averaged over all samples and hours of the day (outputs).

FB, which is given in percentage, shows less average percentage deviation for CNN models in contrast to the autoregressive models in the validation set, while IA, MAE, and RMSE metrics show rather similar values with very slightly higher prediction performance in CNN models. In the test set, the results show that models CNN 2 and AR LSTM 1 have lower FB, but again, all CNN models display a slightly higher performance according to the rest of the metrics. The validation set is used for the early stopping of the training phase. On the contrary, the test set includes data that models have not seen directly or indirectly during the training process, allowing for better generalization of their performance. These outcomes show better prediction performances in the test set than in the validation set, although the FB disagrees.

Table 6. The final number of epochs or training loops and performance metrics for the validation and test sets. Bold values indicate the best results.

		Validation				Test			
Model	**Epochs**	**FB (%)**	**IA**	**MAE (μg/m^3)**	**RMSE (μg/m^3)**	**FB (%)**	**IA**	**MAE (μg/m^3)**	**RMSE (μg/m^3)**
CNN 1	572	**0.86**	**0.70**	**9.05**	**13.77**	0.81	**0.79**	**6.96**	**10.20**
CNN 2	824	**−0.74**	**0.69**	**8.94**	**13.57**	**−0.36**	**0.77**	**6.90**	**10.12**
CNN 3	837	**0.59**	**0.69**	**8.86**	**13.44**	1.18	**0.77**	**6.90**	**10.07**
AR LSTM 1	507	−5.46	0.66	9.96	14.68	**−0.41**	0.73	8.12	11.50
AR LSTM 2	823	−5.86	0.67	9.89	14.74	−4.28	0.75	8.02	11.44
AR GRU 1	510	−4.90	0.68	9.67	14.46	−0.87	0.76	7.83	10.96
AR GRU 2	809	−5.05	0.68	9.63	14.61	−6.50	0.73	7.99	11.49

These results are comparable to those found in the literature. According to the RMSE results obtained in the test set, our models outperform those obtained by Becerra-Rico et al. (2020) [20] for 24 h in advance. FB and RMSE values are similar in the best models fed with imputed data reported in the work of Kukkonen (2003) [21]. Higher IA values but also higher RMSE and MAE were found in the work of Paschalidou (2011) [19]. The same applies to the results obtained by Grivas and Chaloulakou (2006) [24] for the best models, although the MAE is very similar. Our best models (CNN) have higher MAE and RMSE than the hybrid Autoregressive Integrated Moving Average-Artificial Neural Network (ARIMA-ANN) model developed by Díaz-Robles et al. (2008) [12], but their values are pretty close.

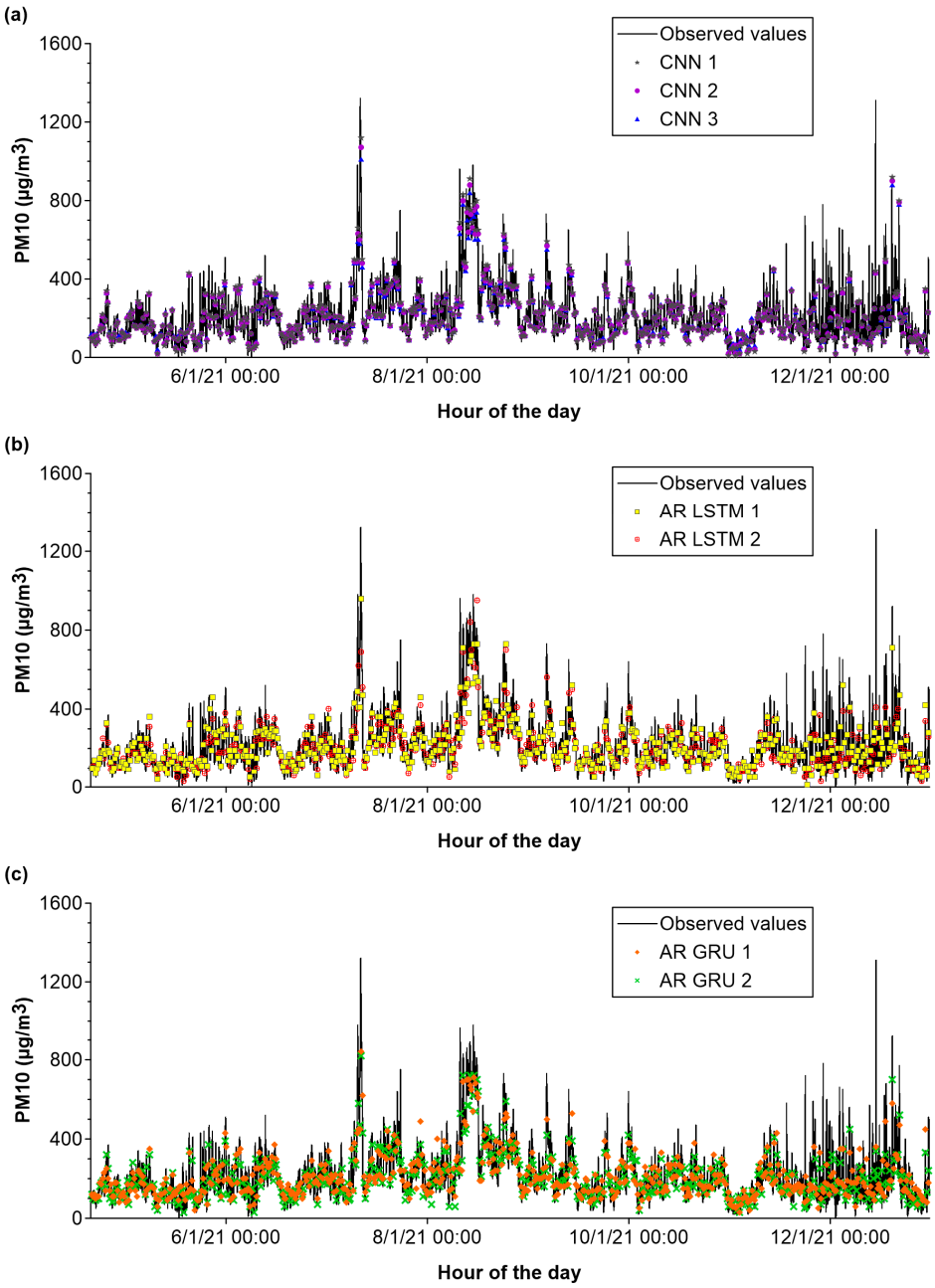


Figure 2. Observed (black line) and predicted values (symbols) of hourly PM10 concentration for (**a**) CNN, (**b**) AR LSTM, and (**c**) GRU models in the test set. For clarity purposes, symbols were depicted for every 10 values.

An analysis of RMSE per network output or per hour of the day was also made [32]. Results can be found in Figure 3. As expected, all models agree that the first predicted hour in advance is the best prediction while, as the hour of the day progresses, higher prediction errors are found. In CNN models, the ascending curve seems to stabilize around an hour of the day (approximately around the 8th hour). Other models exhibit much less clear outcomes or continuously increase up to the last hour, such as the AR LSTM 2 model. However, all of them agree on fast growth in the prediction error in the first 5–6 h. This could be due to a higher deterministic dependency of predictions on inputs in the short term and a more statistical dependency in the long term.

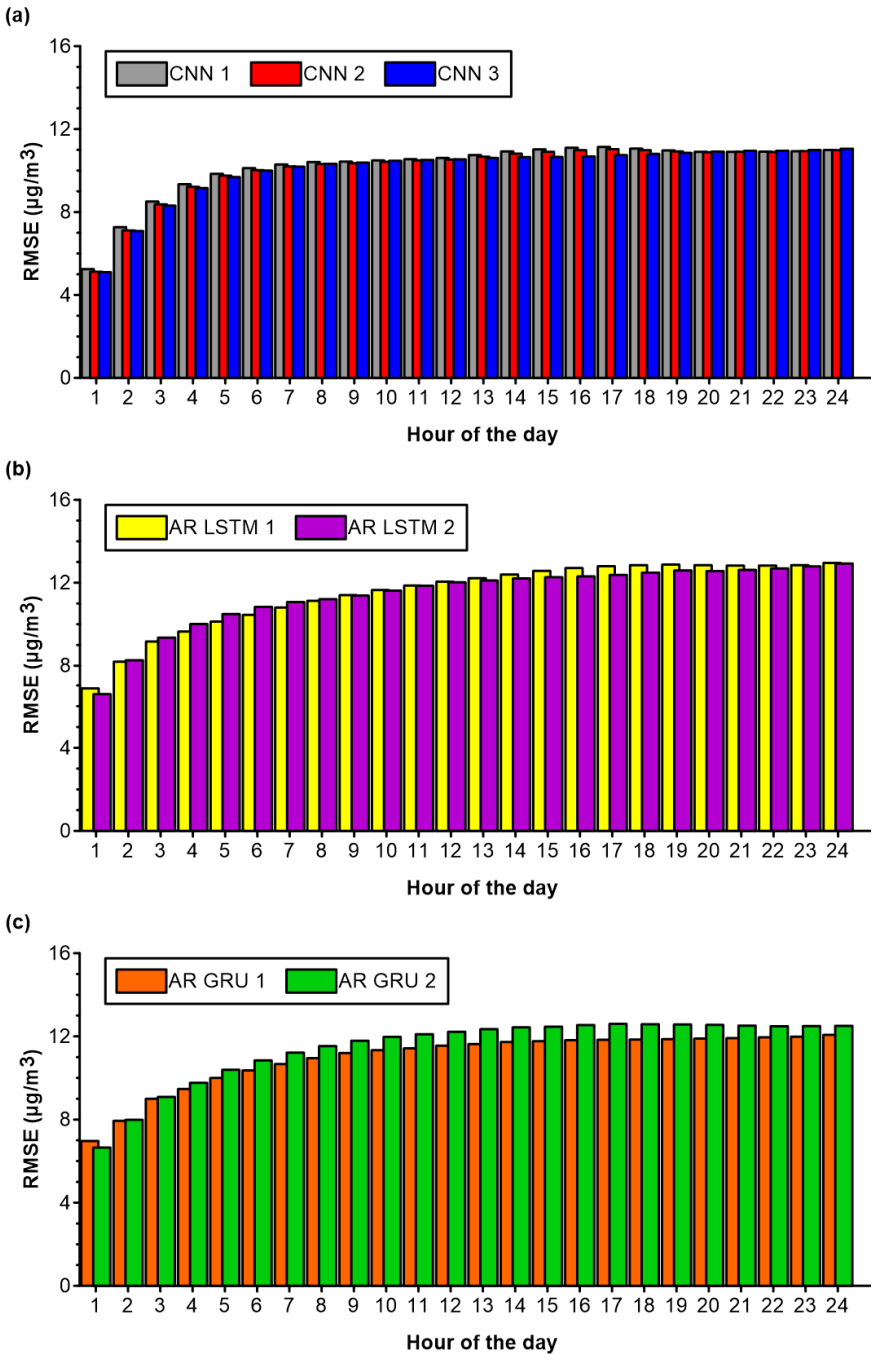


Figure 3. RMSE per hour of the day for (**a**) CNN, (**b**) AR LSTM, and (**c**) AR GRU models.

In Figure 4, predicted vs. observed values for the CNN 1 model are depicted as an example. The same phenomenon discussed above can be appreciated in this figure, where linear fits of first outputs are much closer to the bisector than the last ones. As seen in the figure, R^2 takes values in the range $[0.36, 0.86]$. Consequently, the further ahead the PM10 concentration is predicted, the less accurate the prediction of extreme values becomes.

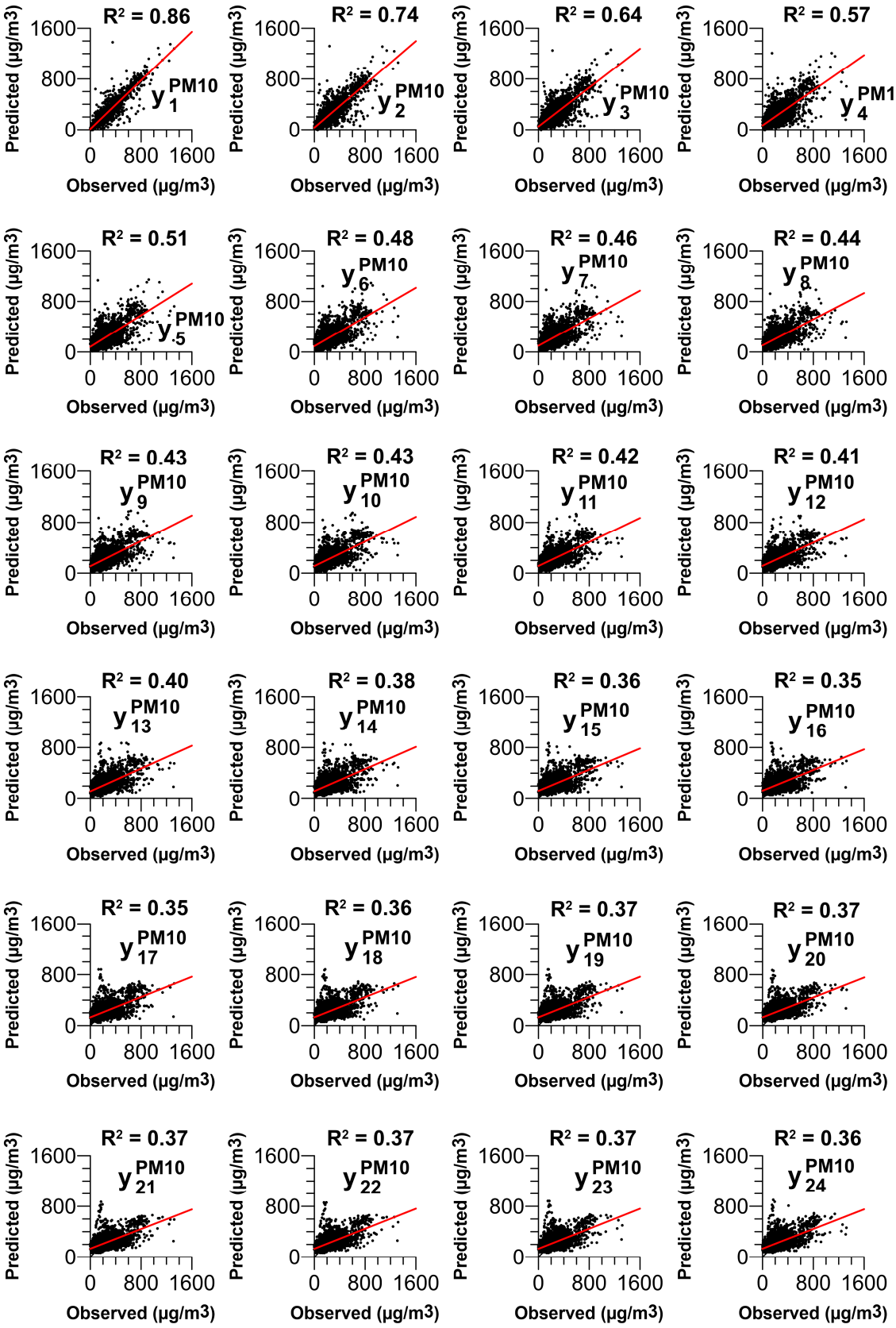


Figure 4. Predicted vs. observed values for the CNN 1 model (black dots) and linear fits (red lines).

3.3. Performance of Recursive Multi-Step Forecasting

The 7-day periods selected to evaluate the performance of recursive multi-step forecasting in different situations are the following (see Figure 1):

- Background PM10 levels: from 13 February 2024-01:00 to 20 February 2024-00:00;
- Strong dust event: from 14 July 2023-01:00 to 21 July 2023-00:00;
- Extreme dust event: from 19 March 2022-01:00 to 26 March 2022-00:00.

Figure 5 shows RMSE values for these three different situations in distinct forecast periods (from 2 to 7-day forecast periods). All models have similar RMSE values. For extreme events, these errors are progressively lower as the forecast period increases, but in general, these values are very high, which makes these models inappropriate for recursive prediction. The performance of models for the strong event similarly improves in long-term forecasts. RMSE is below 40 $\mu g/m^3$ for every model except for AR GRU 2. Thus, these models are not suitable for recursively predicting strong events. Lastly, RMSE in the background concentration is below 20 $\mu g/m^3$ in every model and is lower in the short term. Therefore, all models provide good estimations for recursive prediction of background PM10 concentration according to this metric.

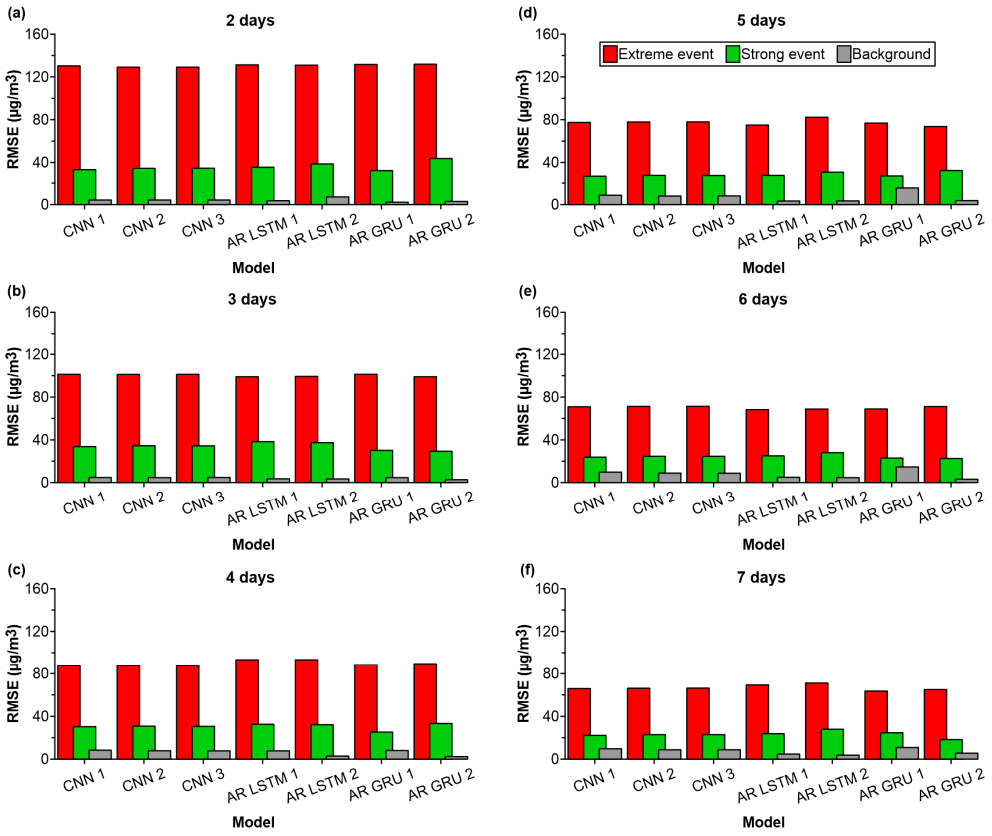


Figure 5. RMSE in the recursive multi-step approach for background (grey bars), strong event (green bars), and extreme event (red bars) and (**a**) 2, (**b**) 3, (**c**) 4, (**d**) 5, (**e**) 6, and (**f**) 7-day forecasts.

The rest of the metrics were also analyzed in Appendix A. According to the results shown in Figures A1–A3, there is a general improvement of metrics as the forecast period increases from 2 to 7 days, which is less evident for IA. They also agree that models are not suitable for the recursive prediction of extreme events regardless of the forecast period considered. For the strong event, FB, IA, and MAE results disagree and show contradictory performances to each other and with the RMSE. As a consequence, the suitability of models to predict strong events by this approach cannot be confirmed. For background PM10 concentration, FB and MAE metrics agree on the suitability of models for recursive prediction and the best performance in the short term, whereas only IA results disagree.

4. Conclusions

Strong ($\mu + 2\sigma \leq PM10 \leq \mu + 4\sigma$) and extreme ($PM10 > \mu + 4\sigma$) dust events were classified and identified in the hourly PM10 concentration series, obtaining an increase of 157.41% and 1166.67% of these events, respectively, between periods 2015–2019 and 2020–May 2024. This growth rate does not necessarily imply a trend, as explained in the literature. Nevertheless, it makes the use of prediction models that attempt to forecast the PM10 concentration more relevant. The duration of these events was found to be between 1 and 53 h (between 1 h and 3 days). The four most prominent extreme events found in the subperiod 2020–May 2024 are:

- 11–12 July 2020, which peaked at 150 $\mu g/m^3$;
- 6 November 2020, with a maximum of 218 $\mu g/m^3$;
- 15–27 March 2022 (the most exceptional and highest recorded extreme dust event), which peaked at 1116 $\mu g/m^3$;
- 13–15 June 2022 with a maximum of 155 $\mu g/m^3$.

The most exceptional extreme dust event (15–27 March 2022) is clearly related to the most relevant Saharan dust incursion which occurred in the last decades all over the Iberian Peninsula. This one caused orange-colored skies and historical levels of PM10 concentration all over the Iberian Peninsula.

Seven neural network models based on CNN, LSTM, and GRU architectures and different hyperparameters were tested with four metrics (FB, IA, MAE, and RMSE). The best performances were found for CNN models according to all metrics, except for FB, which displays lower absolute values for CNN 2 and AR LSTM 1 model. In conclusion, CNN models perform better in predicting these PM10 concentration series than the more complex models based on cells embedded in RNN layers. Furthermore, they exhibit accurate 24 h ahead predictions, which demonstrate that CNN models are effective for short-term predictions of PM10 concentrations, particularly in scenarios with stable background levels. This makes them suitable for operational use in early warning systems for Southern Spain and similar regions.

Prediction error increases faster in the first predicted 5–6 h, and then it stabilizes in CNN models around the 8th hour. A hypothesis for the cause of this phenomenon could be the higher deterministic dependence of predictions on inputs in the short term and a more statistical dependence in the long term, probably due to the activation and interplay of more neurons in the network.

The experiment designed to assess the suitability of the extrapolation of trained models for recursive multi-step (24 h ahead) predictions allowed us to distinguish between three situations in which models could perform differently: background levels, strong dust events, and extreme dust events. The authors conclude that extreme and strong dust events cannot be accurately predicted with the recursive approach, as shown by the disparity between the different metrics and low performances in every model. Nevertheless, background PM10 concentration might be recursively predicted in the short term (by 2–5-day forecasts), producing small prediction errors.

It is important to note that the RMSE threshold of 40 $\mu g/m^3$, referenced in the study, is an empirical observation derived from the performance of the evaluated models under specific conditions. This value is not a universally accepted benchmark but serves as an indicator of the limitations of the models when predicting strong dust events using a recursive multi-step approach. Comparisons with prior studies, such as those by Grivas and Chaloulakou [24] and Kukkonen [21], indicate that RMSE values of a similar magnitude are common in PM10 forecasting, depending on the data and methodologies used.

The evaluation provides critical insights into the current capabilities and limitations of neural network models proposed here for PM10 forecasting, particularly in handling strong

and extreme dust events. These findings serve as a valuable reference for future efforts to enhance forecasting accuracy for high-intensity pollution episodes. While traditional statistical and deterministic models have been widely used in Spain for PM10 forecasting, neural network models offer distinct advantages, including the ability to capture non-linear patterns and sequential dependencies.

The current study relied on the best available data sources, focusing on PM10-specific predictors and seasonality and point predictions. The results demonstrated that CNNs outperformed RNN-based models for short-term predictions, offering a computationally efficient approach for localized applications. However, the models faced significant challenges in predicting strong and extreme events, as evidenced by higher RMSE values and limitations in capturing rare extreme peaks. To address these limitations, future research should explore probabilistic forecasting methods, such as Bayesian neural networks or Monte Carlo dropout, to provide prediction intervals that quantify uncertainty and enhance reliability for critical data. Additionally, incorporating meteorological variables (e.g., wind speed, direction) could improve the models' ability to capture complex patterns associated with strong and extreme events.

The findings and contributions of the models introduced here are summarized below:

1. Tailored analysis for Southern Spain:
    - This study specifically focuses on the Saharan dust incursions affecting PM10 concentrations in Córdoba, Spain, a region where deep neural network models for forecasting this phenomenon have been underexplored;
    - By analyzing hourly PM10 concentrations and their dependency on dust events, the study provides insights unique to this geographic context, offering practical benefits for regional early warning systems.
2. Performance of CNNs for short-term predictions:
    - The proposed CNN models demonstrated notable accuracy in short-term predictions (24-h forecasts) with RMSE values between 10.00 and 10.20 $\mu g/m^3$;
    - This highlights the efficiency of CNNs for operational use in predicting PM10 concentrations, particularly for stable background levels and strong dust events.
3. Simplified recursive multi-step prediction for background and strong events:
    - The recursive prediction approach tested in this study provided reasonable accuracy for background PM10 concentrations and strong dust events over short forecast periods (2–5 days), with RMSE values remaining under 20 $\mu g/m^3$ for background levels;
    - While other models, such as CNN-LSTM hybrids (e.g., Gilik et al., 2022 [37]) emphasize long-term forecasting, our approach simplifies implementation without sacrificing accuracy for short-term operational needs.
4. Challenges in extreme event prediction and practical insights:
    - Our study confirms that all tested models, including CNNs, LSTMs, and GRUs, face significant limitations in predicting extreme events using a recursive approach due to high variability and large RMSE values (>40 $\mu g/m^3$);
    - This aligns with findings in studies like Wang et al. (2021) [38], and Hu et al. (2023) [26], but our work explicitly quantifies these limitations in the context of Saharan dust incursions, filling a gap in the literature.
5. Local data integration without meteorological variables:
    - Unlike previous studies that rely heavily on meteorological inputs or satellite data, our models achieve competitive performance using PM10-specific predic-

tors (e.g., lagged values and seasonality). This simplifies application in resource-limited contexts while maintaining accuracy for localized predictions.

In comparison with other existing CNNs and RNNs, which incorporate hybrid architectures and advanced features, the benefits provided by the presented models are:

1. Efficiency and simplicity:
   - Compared to hybrid models [37], our CNNs are computationally efficient, requiring less training time and resources. This makes them suitable for real-time applications in air quality monitoring systems.
2. Localized relevance:
   - While broader models like CNN-BiLSTMs [38] aim for general applicability across regions, our study directly addresses the unique characteristics of Saharan dust incursions in Southern Spain, offering more precise regional forecasting.
3. Insights into strong and extreme dust events:

By separately analyzing background, strong, and extreme dust events, this study provides practical guidelines for the applicability and limitations of neural network models under different scenarios. This level of granularity has been explored less in prior studies.

4. Practical for data-limited scenarios:

Unlike models relying on extensive meteorological or satellite data [26], our models perform effectively with limited data inputs, making them more accessible for local agencies with constrained resources. They provide actionable predictions with minimal data dependencies, making them easier to integrate into early warning systems.

These unique benefits demonstrate how the study models excel in localized short-term forecasting, offering a practical alternative to broader, more complex models while addressing the operational needs of resource-constrained regions.

The significant growth in the frequency and intensity of Saharan dust events in southern Spain makes the development of early-warning systems a fact of extreme importance since the high PM10 concentration has numerous harmful health effects, such as the exacerbation of chronic respiratory and cardiovascular diseases. In this context, the assessment and improvement of neural network prediction models in a multi-step approach should be considered as a relevant objective by policymakers due to the benefits in the improvement of the services rendered by the public institutions that are responsible for the monitorization of PM10 concentration levels and other particulate matter. From the authors' perspective, evaluating the trained models for 24-h ahead predictions and extending them using the recursive multi-step prediction approach addressed a significant gap in this research field that had not been thoroughly explored.

Author Contributions: Conceptualization, J.G.-G. and F.J.J.-H.; methodology, J.G.-G. and F.J.J.-H.; software, J.G.-G.; validation, J.G.-G., F.J.J.-H. and E.G.d.R.; formal analysis, J.G.-G. and F.J.J.-H.; investigation, J.G.-G., F.J.J.-H. and E.G.d.R.; resources, F.J.J.-H. and E.G.d.R.; data curation, J.G.-G.; writing—original draft preparation, J.G.-G., F.J.J.-H. and E.G.d.R.; writing—review and editing, F.J.J.-H. and E.G.d.R.; visualization, J.G.-G. and F.J.J.-H.; supervision, F.J.J.-H. and E.G.d.R.; project administration, F.J.J.-H. and E.G.d.R.; funding acquisition, F.J.J.-H. and E.G.d.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Research Program of the University of Córdoba (2024), Spain.

Data Availability Statement: The original data presented in the study are openly available at https://www.miteco.gob.es/es.html (accessed on 16 October 2024) and https://www.juntadeandalucia.es/medioambiente/portal/home (accessed on 16 October 2024).

Acknowledgments: The authors gratefully acknowledge the support of the funding source.

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

Appendix A. Evaluation of Performance of Recursive Multi-Step Forecasting with Other Metrics

FB, IA, and MAE metrics were also analyzed to assess the performance of trained models in recursive multi-step forecasting.

FB can be found in Figure A1. The largest deviations from the mean are found for the extreme dust event in the 2-day forecast, which progressively reduces as the forecast period increases. As shown by their performance, all models are inappropriate for the recursive prediction of extreme events and exhibit high variability among the different forecast periods. The selected strong event is also better estimated for larger forecast periods. The AR GRU 2 model outperforms the rest of them with 1% FB in the 7-day forecast. However, in general, these models are not suitable for the recursive prediction of strong events according to this metric, except for 6 and 7-day forecasts. The absolute FB of models for background concentration is lower than those values obtained for the strong event for 2, 3, and 4-day forecasts. For the 5-day forecast, the performance is mostly lower except for AR GRU 1. Therefore, every model gives better estimations for background levels in the short term, which makes them suitable for recursive prediction in this case.

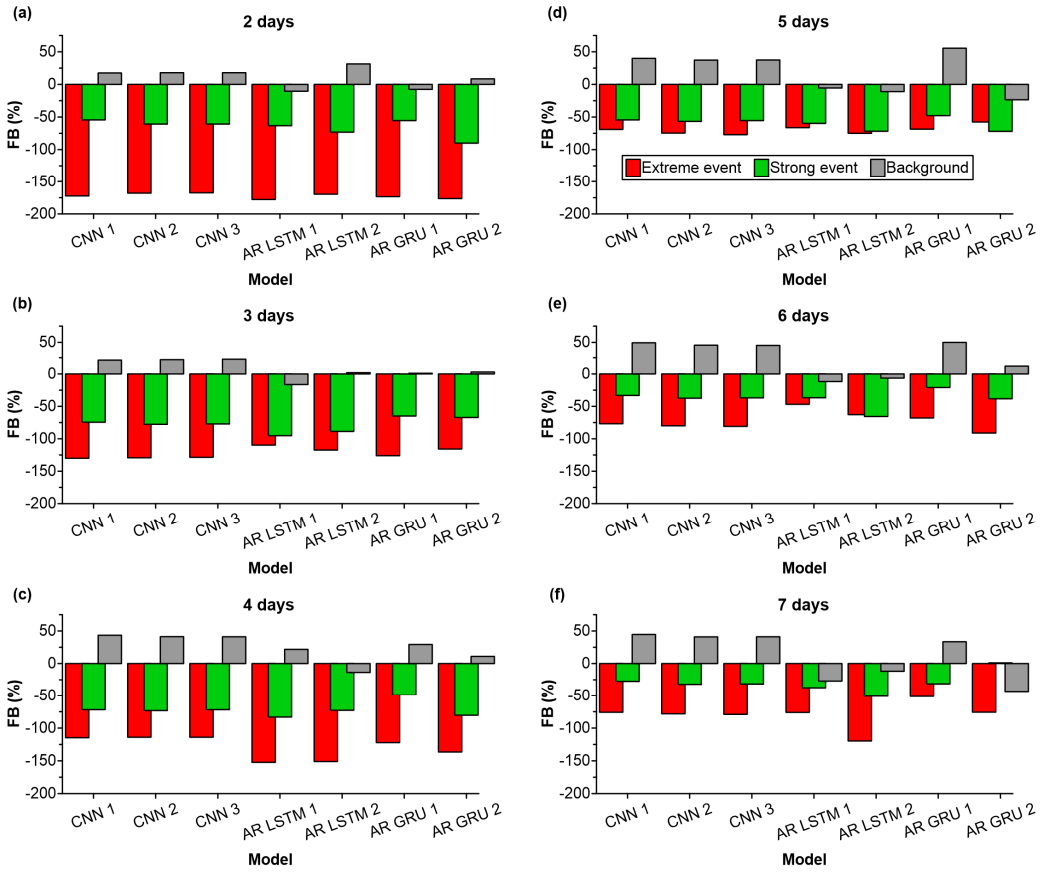


Figure A1. FB in percentage in the recursive multi-step approach for background (grey bars), strong event (green bars), and extreme event (red bars) and (**a**) 2, (**b**) 3, (**c**) 4, (**d**) 5, (**e**) 6, and (**f**) 7-day forecasts.

IA results are depicted in Figure A2. The IA for the extreme event is below 0.5 for every model, regardless of the forecast period. Moreover, models show lower IA values as the forecast period increases. According to this result, the recursive forecast approach is

again inappropriate for estimating the extreme event. In the case of the strong event, the highest IA can be found in 4 and 5-day forecasts, except for AR GRU 2. Nevertheless, all values are between 0.3 and 0.6, making these models not suitable for this kind of prediction in the case of strong events. Finally, a high disparity in the values of IA for the background levels can be found in every model and forecast period. Despite this variability, the largest value was found for the AR GRU 2 model (0.7) in the 3-day forecast period. Thus, according to this metric, models are not capable of predicting background concentration using this approach either.

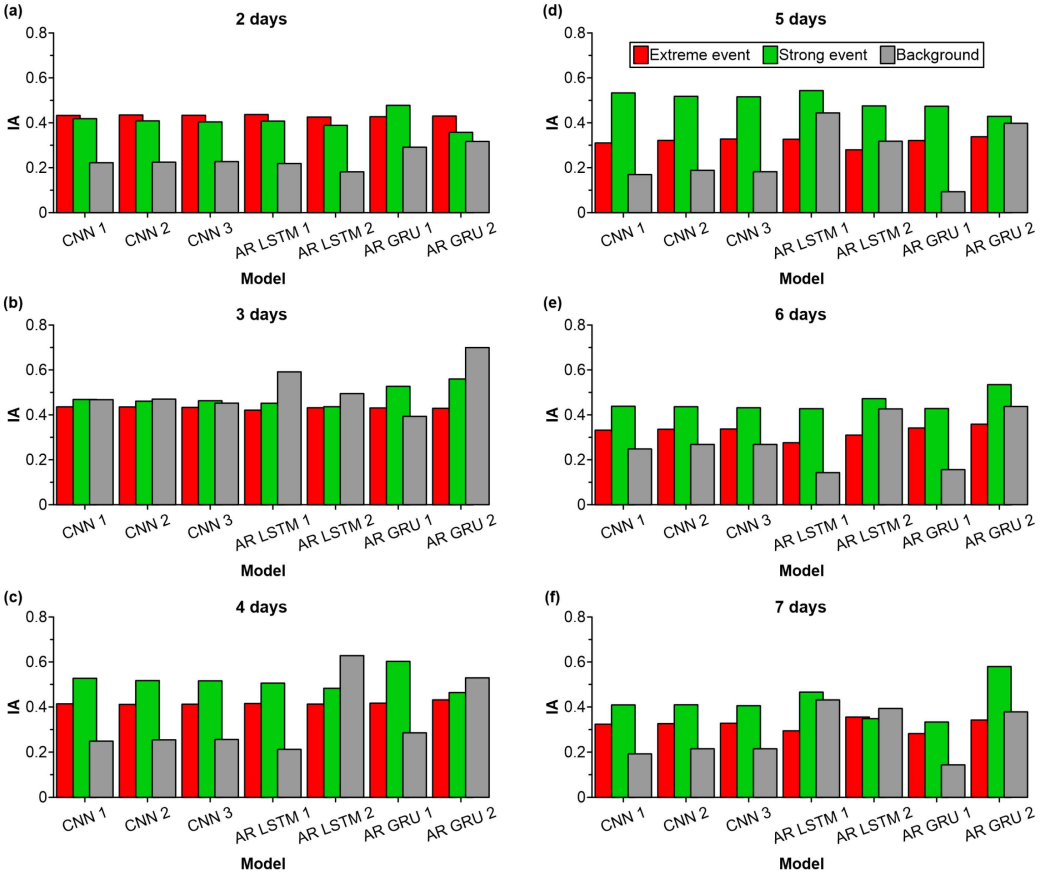


Figure A2. IA in the recursive multi-step approach for background (grey bars), strong event (green bars), and extreme event (red bars) and (**a**) 2, (**b**) 3, (**c**) 4, (**d**) 5, (**e**) 6, and (**f**) 7-day forecasts.

MAE is shown in Figure A3. All models have very high MAE values for extreme event prediction, which again makes this approach unsuitable for recursive prediction. Lower values are found for the strong event. However, they are still relatively high for 2 and 3-day forecasts when models are less suitable. Finally, background levels are well estimated by all models, which show high precision in the 2-day forecast according to this metric.

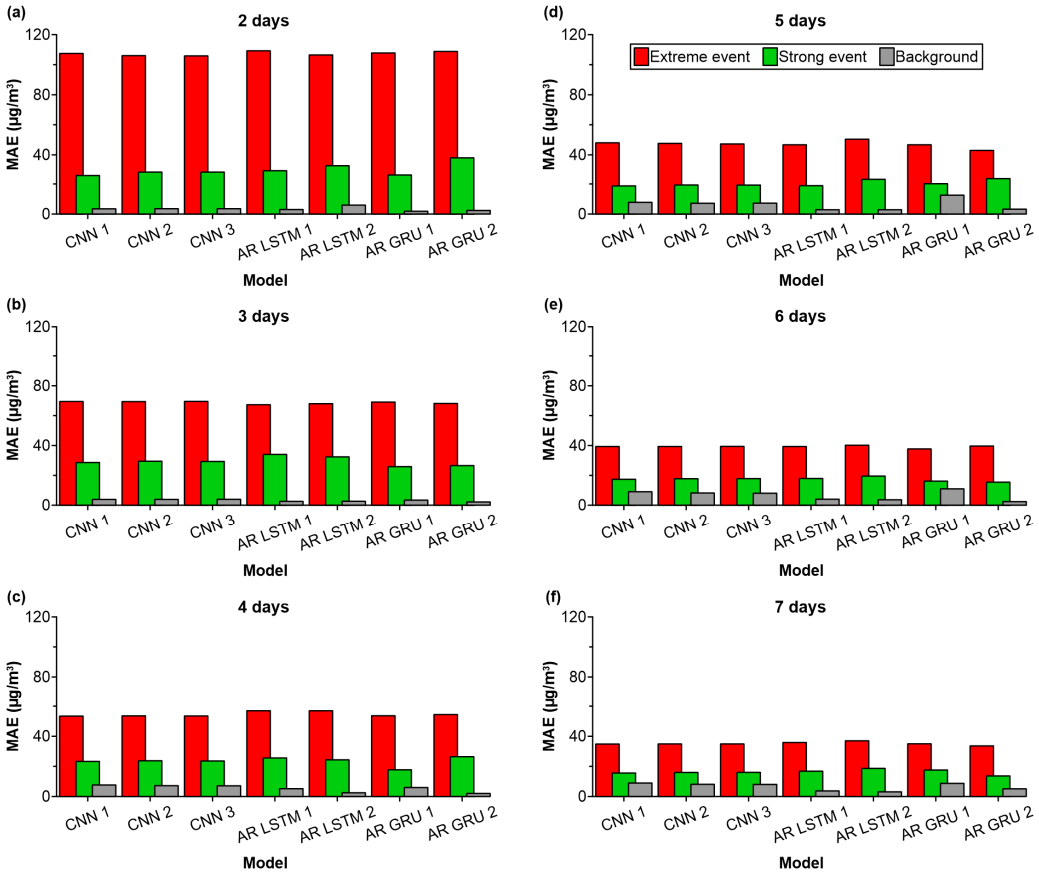


Figure A3. MAE in the recursive multi-step approach for background (grey bars), strong event (green bars), and extreme event (red bars) and (**a**) 2, (**b**) 3, (**c**) 4, (**d**) 5, (**e**) 6, and (**f**) 7-day forecasts.

References

1. Goudie, A.S. Desert Dust and Human Health Disorders. *Environ. Int.* **2014**, *63*, 101–113. [CrossRef]
2. Kim, K.-H.; Kabir, E.; Kabir, S. A Review on the Human Health Impact of Airborne Particulate Matter. *Environ. Int.* **2015**, *74*, 136–143. [CrossRef] [PubMed]
3. Cuevas-Agulló, E.; Barriopedro, D.; García, R.D.; Alonso-Pérez, S.; González-Alemán, J.J.; Werner, E.; Suárez, D.; Bustos, J.J.; García-Castrillo, G.; García, O.; et al. Sharp Increase in Saharan Dust Intrusions over the Western Euro-Mediterranean in February–March 2020–2022 and Associated Atmospheric Circulation. *Atmos. Chem. Phys.* **2024**, *24*, 4083–4104. [CrossRef]
4. Huerta-Viso, A.; Crespo, J.; Galindo, N.; Yubero, E.; Nicolás, J.F. Saharan Dust Events over the Valencian Community (Eastern Iberian Peninsula): Synoptic Circulation Patterns and Contribution to PM10 Levels. *Aerosol Air Qual. Res.* **2020**, *20*, 2519–2528. [CrossRef]
5. Oduber, F.; Calvo, A.I.; Blanco-Alegre, C.; Castro, A.; Nunes, T.; Alves, C.; Sorribas, M.; Fernández-González, D.; Vega-Maray, A.M.; Valencia-Barrera, R.M.; et al. Unusual Winter Saharan Dust Intrusions at Northwest Spain: Air Quality, Radiative and Health Impacts. *Sci. Total Environ.* **2019**, *669*, 213–228. [CrossRef] [PubMed]
6. Rodríguez, S.; López-Darias, J. Emerging Extreme Saharan-Dust Events Expand Northward over the Atlantic and Europe Prompting Record-Breaking PM_{10} and $PM_{2.5}$ Episodes. *Atmos. Chem. Phys.* **2024**, *24*, 12031–12053. [CrossRef]
7. Ansmann, A.; Bösenberg, J.; Chaikovsky, A.; Comerón, A.; Eckhardt, S.; Eixmann, R.; Freudenthaler, V.; Ginoux, P.; Komguem, L.; Linné, H.; et al. Long-Range Transport of Saharan Dust to Northern Europe: The 11–16 October 2001 Outbreak Observed with EARLINET. *J. Geophys. Res.-Atmos.* **2003**, *108*, 4783. [CrossRef]
8. World Meteorological Organization (WMO). *WMO Airborne Dust Bulletin*; WMO Airborne Dust Bulletin, 7; WMO: Geneva, Switzerland, 2023. Available online: https://wmo.int/publication-series/wmo-airborne-dust-bulletin-no-7-september-2023 (accessed on 21 January 2025).
9. Woollings, T.; Barriopedro, D.; Methven, J.; Son, S.-W.; Martius, O.; Harvey, B.; Sillmann, J.; Lupo, A.R.; Seneviratne, S. Blocking and Its Response to Climate Change. *Curr. Clim. Chang. Rep.* **2018**, *4*, 287–300. [CrossRef] [PubMed]
10. Sousa, P.M.; Barriopedro, D.; García-Herrera, R.; Woollings, T.; Trigo, R.M. A New Combined Detection Algorithm for Blocking and Subtropical Ridges. *J. Clim.* **2021**, *34*, 7735–7758. [CrossRef]

11. Cabaneros, S.M.; Calautit, J.K.; Hughes, B.R. A Review of Artificial Neural Network Models for Ambient Air Pollution Prediction. *Environ. Model. Softw.* **2019**, *119*, 285–304. [CrossRef]
12. Díaz-Robles, L.A.; Ortega, J.C.; Fu, J.S.; Reed, G.D.; Chow, J.C.; Watson, J.G.; Moncada-Herrera, J.A. A Hybrid ARIMA and Artificial Neural Networks Model to Forecast Particulate Matter in Urban Areas: The Case of Temuco, Chile. *Atmos. Environ.* **2008**, *42*, 8331–8340. [CrossRef]
13. Chaloulakou, A.; Grivas, G.; Spyrellis, N. Neural Network and Multiple Regression Models for PM_{10} Prediction in Athens: A Comparative Assessment. *J. Air Waste Manag. Assoc.* **2003**, *53*, 1183–1190. [CrossRef]
14. McKendry, I.G. Evaluation of Artificial Neural Networks for Fine Particulate Pollution (PM_{10} and $PM_{2.5}$) Forecasting. *J. Air Waste Manag. Assoc.* **2002**, *52*, 1096–1101. [CrossRef] [PubMed]
15. Ferrari, L.; Guariso, G. Convolutional Forecasting of Particulate Matter: Toward a Data-Driven Generalized Model. *Atmosphere* **2024**, *15*, 398. [CrossRef]
16. Hong, W.Y. Meteorological Variability and Predictive Forecasting of Atmospheric Particulate Pollution. *Sci. Rep.* **2024**, *14*, 14. [CrossRef]
17. Athira, V.; Geetha, P.; Vinayakumar, R.; Soman, K.P. DeepAirNet: Applying Recurrent Networks for Air Quality Prediction. *Procedia Comput. Sci.* **2018**, *132*, 1394–1403. [CrossRef]
18. Barrero-González, D.; Ramírez-Montañez, J.A.; Aceves-Fernández, M.A.; Ramos-Arreguín, J.M. Capability of an Elman Recurrent Neural Network for Predicting the Non-Linear Behavior of Airborne Pollutants. *Earth Sci. Inform.* **2022**, *15*, 125–135. [CrossRef]
19. Paschalidou, A.K.; Karakitsios, S.; Kleanthous, S.; Kassomenos, P.A. Forecasting Hourly PM10 Concentration in Cyprus through Artificial Neural Networks and Multiple Regression Models: Implications to Local Environmental Management. *Environ. Sci. Pollut. Res.* **2011**, *18*, 316–327. [CrossRef]
20. Becerra-Rico, J.; Aceves-Fernández, M.A.; Esquivel-Escalante, K.; Pedraza-Ortega, J.C. Airborne Particle Pollution Predictive Model Using Gated Recurrent Unit (GRU) Deep Neural Networks. *Earth Sci. Inform.* **2020**, *13*, 821–834. [CrossRef]
21. Kukkonen, J. Extensive Evaluation of Neural Network Models for the Prediction of NO_2 and PM10 Concentrations, Compared with a Deterministic Modelling System and Measurements in Central Helsinki. *Atmos. Environ.* **2003**, *37*, 4539–4550. [CrossRef]
22. Gogineni, R.; Umakanth, N.; Ramakrishna, Y.; Veeraswamy, P. Assessment of Daily Concentrations of Air Pollutants over Four Stations of Andhra Pradesh, India, during 2019–2022. *Acta Geophys.* **2023**, *72*, 1227–1249. [CrossRef]
23. Ramírez Montañez, J.A.; Aceves Fernández, M.A.; Arriaga, S.T.; Ramos Arreguin, J.M.; Salini Calderón, G.A. Evaluation of a Recurrent Neural Network LSTM for the Detection of Exceedances of Particles PM10. In Proceedings of the 2019 16th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE), Mexico City, Mexico, 11 September 2019; pp. 1–6. [CrossRef]
24. Grivas, G.; Chaloulakou, A. Artificial Neural Network Models for Prediction of PM10 Hourly Concentrations, in the Greater Area of Athens, Greece. *Atmos. Environ.* **2006**, *40*, 1216–1229. [CrossRef]
25. Bakar, M.A.A.; Ariff, N.M.; Man, P.H.; Nadzir, M.S.M. Forecasting the Air Quality Time Series Data Using Univariate and Multivariate Convolutional Neural Network Models. *AIP Conf. Proc.* **2024**, *2895*, 040011. [CrossRef]
26. Hu, D.; Li, M.; Guo, J. The Early Warning Model of Dust Concentration in Smart Construction Sites Based on Long Short Term Memory Network. In Proceedings of the 2023 5th International Conference on Intelligent Control, Measurement and Signal Processing (ICMSP), Chengdu, China, 19–21 May 2023; pp. 1132–1137. [CrossRef]
27. Mishra, A.K.; Desai, V.R. Drought Forecasting Using Feed-Forward Recursive Neural Network. *Ecol. Model.* **2006**, *198*, 127–138. [CrossRef]
28. Chaloulakou, A.; Saisana, M.; Spyrellis, N. Comparative Assessment of Neural Networks and Regression Models for Forecasting Summertime Ozone in Athens. *Sci. Total Environ.* **2003**, *313*, 1–13. [CrossRef] [PubMed]
29. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning Representations by Back-Propagating Errors. *Nature* **1986**, *323*, 533–536. [CrossRef]
30. Azizjon, M.; Jumabek, A.; Kim, W. 1D CNN Based Network Intrusion Detection with Normalization on Imbalanced Data. In Proceedings of the 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), Fukuoka, Japan, 19–21 February 2020; pp. 218–224. [CrossRef]
31. Kiranyaz, S.; Ince, T.; Abdeljaber, O.; Avci, O.; Gabbouj, M. 1-D Convolutional Neural Networks for Signal Processing Applications. In Proceedings of the ICASSP 2019—2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, UK, 12–17 May 2019; pp. 8360–8364. [CrossRef]
32. Cai, M.; Pipattanasomporn, M.; Rahman, S. Day-Ahead Building-Level Load Forecasts Using Deep Learning vs. Traditional Time-Series Techniques. *Appl. Energy* **2019**, *236*, 1078–1088. [CrossRef]
33. Graves, A. Generating Sequences with Recurrent Neural Networks. *arXiv* **2014**, arXiv:1308.0850v5. [CrossRef]
34. Greff, K.; Srivastava, R.K.; Koutnik, J.; Steunebrink, B.R.; Schmidhuber, J. LSTM: A Search Space Odyssey. *IEEE Trans. Neural Netw. Learn. Syst.* **2017**, *28*, 2222–2232. [CrossRef]

35. Querol, X.; Pey, J.; Pandolfi, M.; Alastuey, A.; Cusack, M.; Pérez, N.; Moreno, T.; Viana, M.; Mihalopoulos, N.; Kallos, G.; et al. African Dust Contributions to Mean Ambient PM10 Mass-Levels across the Mediterranean Basin. *Atmos. Environ.* **2009**, *43*, 4266–4277. [CrossRef]
36. Levin, Z.; Ganor, E. The Effects of Desert Particles on Cloud and Rain Formation in the Eastern Mediterranean. In *The Impact of Desert Dust Across the Mediterranean*; Guerzoni, S., Chester, R., Eds.; Springer: Dordrecht, The Netherlands, 1996; pp. 77–86, ISBN 978-94-017-3354-0.
37. Gilik, A.; Ogrenci, A.S.; Ozmen, A. Air Quality Prediction using CNN + LSTM-based Hybrid Deep Learning Architecture. *Environ. Sci. Pollut. Res.* **2022**, *29*, 11920–11938. [CrossRef]
38. Wang, J.; Li, J.; Wang, X.; Wang, T.; Sun, Q. An Air Quality Prediction Model Based on CNN-BiNLSTM-Attention. *Environ. Dev. Sustain.* **2022**. [CrossRef]
39. Shi, X.; Chen, Z.; Wang, H.; Yeung, D.Y.; Wong, W.K.; Woo, W.C. Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting. *arXiv* **2015**, arXiv:1506.04214v2. [CrossRef]
40. Ministerio Para La Transición Ecológica y El Reto Demográfico. Available online: https://www.miteco.gob.es/es.html (accessed on 16 October 2024).
41. Junta de Andalucía, Consejería de Sostenibilidad y Medio Ambiente Portal Ambiental de Andalucía. Available online: https://www.juntadeandalucia.es/medioambiente/portal/home (accessed on 16 October 2024).
42. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef] [PubMed]
43. Cho, K.; van Merrienboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning Phrase Representations Using RNN Encoder-Decoder for Statistical Machine Translation. *arXiv* **2014**, arXiv:1406.1078v3. [CrossRef]
44. TensorFlow. Available online: https://www.tensorflow.org/ (accessed on 21 October 2024).
45. Abadi, M.; Agarwal, A.; Barham, P.; Brevdo, E.; Chen, Z.; Citro, C.; Corrado, G.S.; Davis, A.; Dean, J.; Devin, M.; et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. *arXiv* **2015**, arXiv:1603.04467v2. [CrossRef]
46. Kaparthi, S.; Bumblauskas, D. Designing predictive maintenance systems using decision tree-based machine learning techniques. *Int. J. Qual. Reliab. Manag.* **2020**, *37*, 659–686. [CrossRef]
47. Time Series Forecasting. TensorFlow Core. Available online: https://www.tensorflow.org/tutorials/structured_data/time_series#advanced_autoregressive_model (accessed on 21 October 2024).
48. Working with RNNs. TensorFlow Core. Available online: https://www.tensorflow.org/guide/keras/working_with_rnns (accessed on 21 October 2024).
49. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv* **2017**, arXiv:1412.6980v9. [CrossRef]
50. Chang, J.C.; Hanna, S.R. Air Quality Model Performance Evaluation. *Meteorol. Atmos. Phys.* **2004**, *87*, 167–196. [CrossRef]
51. Willmott, C.J. On the Validation of Models. *Phys. Geogr.* **1981**, *2*, 184–194. [CrossRef]
52. Willmott, C.J.; Matsuura, K. Advantages of the Mean Absolute Error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim. Res.* **2005**, *30*, 79–82. [CrossRef]
53. Chai, T.; Draxler, R.R. Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature. *Geosci. Model Dev.* **2014**, *7*, 1247–1250. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.