

Article

Using Hybrid Deep Learning Models to Predict Dust Storm Pathways with Enhanced Accuracy

Mahdis Yarmohamadi ¹, Ali Asghar Alesheikh ^{1,*}  and Mohammad Sharif ^{2,*} 

¹ Department of Geospatial Information Systems, K. N. Toosi University of Technology, Tehran 19967-15433, Iran

² Institute of Mobility and Urban Planning, University of Duisburg-Essen, 45127 Essen, Germany

* Correspondence: alesheikh@kntu.ac.ir (A.A.A.); mohammad.sharif@uni-due.de (M.S.)

Abstract: As a potential consequence of climate change, the intensity and frequency of dust storms are increasing. A dust storm arises when strong winds blow loose dust from a dry surface, transporting soil particles from one place to another. The environmental and human health impacts of dust storms are substantial. Accordingly, studying the monitoring of this phenomenon and predicting its pathways for early decision making and warning are vital. This study employs deep learning methods to predict dust storm pathways. Specifically, hybrid CNN-LSTM and ConvLSTM models have been proposed for the 24 h-ahead prediction of dust storms in the region under study. The Modern-Era Retrospective Analysis for Research and Applications version 2 (MERRA-2) product that includes the dust particles and the meteorological information, such as surface wind speed and direction, relative humidity, surface air temperature, and skin temperature, is used to train the proposed models. These contextual features are selected utilizing the random forest feature importance method. The results indicate an improvement in the performance of both models by considering the contextual information. Moreover, a 0.2 increase in the Kappa coefficient criterion across all forecast hours indicates the CNN-LSTM model outperforms the ConvLSTM model when contextual information is considered.

Keywords: dust storm; CNN-LSTM; ConvLSTM; time series; context information; MERRA-2



Academic Editor: Antonello Pasini

Received: 31 October 2024

Revised: 23 December 2024

Accepted: 7 January 2025

Published: 12 January 2025

Citation: Yarmohamadi, M.; Alesheikh, A.A.; Sharif, M. Using Hybrid Deep Learning Models to Predict Dust Storm Pathways with Enhanced Accuracy. *Climate* **2025**, *13*, 16. <https://doi.org/10.3390/cli13010016>

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

As a result of urbanization, industrialization, and climate change, dust storms have become one of the sources of global air pollution [1]. Dust storms usually occur when powerful winds lift large amounts of sand and dust from bare and dry soils and transport them over great distances [2]. These storms are often linked to dry, hot weather patterns, low humidity, and atmospheric instability. Dust storms contribute significantly to air pollution and health risks [3]. Consequently, transport systems can be disrupted considerably as dust storms damage roads and runways and decrease visibility at the same time [4]. A dust storm can also lead to elevated concentrations of bioaerosols, which may pose health risks to individuals [5]. Moreover, the deposition of dust particles on soil and vegetation during storms can disrupt ecosystems [6]. Over long distances, dust can transport foreign materials and pollutants, affecting the climate and spreading contaminants in different regions [7]. Therefore, the far-reaching effects of dust storms underscore the need for comprehensive strategies to mitigate their impacts on both human health and the environment.

A dust storm occurs when dust particles are transported through the atmosphere. Dust storms, characterized by their variable sizes, can impact expansive regions in a short amount of time. To mitigate the harmful effects of dust storms, predictive models (e.g.,

deep learning and machine learning algorithms) have been widely used. Numerous deep learning methodologies have been employed to predict the dynamics of moving processes, such as wildfires, oil spills, and dust storms. For example, Khennou et al. [8] formulated FU-NetCast, a deep learning architecture derived from U-Net. This model explored the propagation of wildfires over a period of 24 h and obtained an Area Under the Curve (AUC) score of 80%. Wang et al. [9] implemented a Long Short-Term Memory (LSTM) in consideration of different oil spill characteristics involving edge contour features, diffusion speed, centroids, as well as the surface. The authors also developed a model targeted at tracking and forecasting oil spilling trends. From this study's results, it follows that this model demonstrates high levels of stability. In dust storm studies, forecasting the transport pathways of dusts were mainly conducted by numerical approaches, such as the NOAA Hybrid Single-Particle Lagrangian Integrated Trajectory model (HYSPLIT) [10,11]. HYSPLIT uses real-time meteorological data from sources like weather stations and satellites. By analyzing wind patterns from the meteorological data, the model predicts how these particles will travel through the atmosphere over time [12]. Tiancheng et al. [13] used an improved naïve Bayesian–Convolutional Neural Network (CNN) framework to forecast dust storms by analyzing atmospheric dynamics and terrestrial variables. In another study, Li et al. [14] predicted the 16th day of a storm using satellite images of the storm and meteorological data of the previous 15 days using the CNN and LSTM models. Yarmohamadi et al. [15] employed a CNN model to predict the subsequent 24 h of a dust storm using MERRA-2 Aerosol Optical Depth (AOD) data and contextual information. This model was able to extract relevant features from different dust storms in the dataset, although it required significant computational power for training. The ability to recognize intricate patterns and trends in large moving process datasets is one of the most substantial advantages of deep learning models. This makes them effective in developing early warning systems for events, such as natural disasters, disease outbreaks, or environmental hazards. Deep learning models can be integrated to maximize their strengths and enhance predictive performance. Hybrid deep learning models are often more robust and less susceptible to overfitting, as they integrate diverse perspectives and mitigate vulnerabilities of individual models [16]. Moreover, deep learning models can learn abstract and hierarchical representations of data, so they are adaptable to new situations or contexts [17].

Dust storms are complex phenomena, their simulation is influenced by a convergence of surface and meteorological conditions, referred to as the environmental context in this study [18,19]. Depending on the severity of the event, they may last for a few hours or days [20]. It is essential to identify the primary factors that influence dust storm events to mitigate their frequency and intensity [21]. The impacts of environmental factors have been extensively researched [22,23], and studies have demonstrated their effectiveness in influencing dust emissions. As a result of the environmental context, high-speed wind events transport large quantities of soil to other areas, resulting in dust storms [24]. Yarmohamadi et al. [15] predict dust storm transport pathways by considering the surface wind speed and direction, relative humidity, surface air temperature, and surface temperature. Khashi et al. [25] evaluated the effects of precipitation and wind speed on dust propagation in arid regions. An analysis of the social, economic, and ecological impacts of dust storms in Iran was carried out by integrating nine datasets, which included drought events, temperature, precipitation, locations of sandy soil, frequency of sand and dust storms (SDSs), human-induced soil degradation, human influence index, rain use efficiency, and losses in net primary productivity [26]. To determine the number and duration of dust storms in Northern parts of China from 1978 to 2007, Wang et al. [27] considered wind speed and wind direction. Frequent dust storms occur in the world's dust belt, which includes North Africa, the Middle East, Central Asia, and parts of China [28]. To predict and mitigate

dust storms effectively, it is necessary to have an in-depth understanding of how these environmental contexts interact.

Due to the dynamic nature of dust storms, which includes changes in size and movement, advanced tools are needed for accurate predictions and a comprehensive understanding [4]. Despite the consequences of dust storms, limited research has focused on predicting their paths, often treating these processes as point events. This study emphasizes the need for robust tools capable of modeling their complex mechanisms and investigating the environmental variables that impact them. Although the nature of dust storms differs from one region to another, it is evident that their movement is linked to the environmental context. This study aims to develop deep learning techniques to forecast dust storm transport pathways for the next 24 h by taking into account historical data and environmental contextual information. This study contributes to (i) the development of a CNN-LSTM and a ConvLSTM architecture for accurately predicting dust storm pathways, and (ii) the examination of how environmental contexts—such as surface wind speed and direction, precipitation, air temperature, surface skin temperature, and relative humidity—affect the accuracy of dust storm predictions.

2. Materials and Methods

2.1. Study Region

Dust storms are frequent in Central Asia, particularly during the spring and summer months [29]. The study area for this research included several countries in Central and South Asia, including Afghanistan, Iran, Pakistan, Tajikistan, Turkmenistan, and Uzbekistan (Figure 1). These regions are characterized by arid areas and several deserts, making them vulnerable to dust storms. Dust storms frequently occur in these regions and significantly impact agriculture and transportation [30], interfere with everyday operations, and present risks to human health [31]. The reason for choosing this specific study region is mainly related to the dust-source areas, which are mostly composed of dried-up lakes, like the Aral Sea and the Hamouns in Sistan, and the deserts that are located in eastern Pakistan.

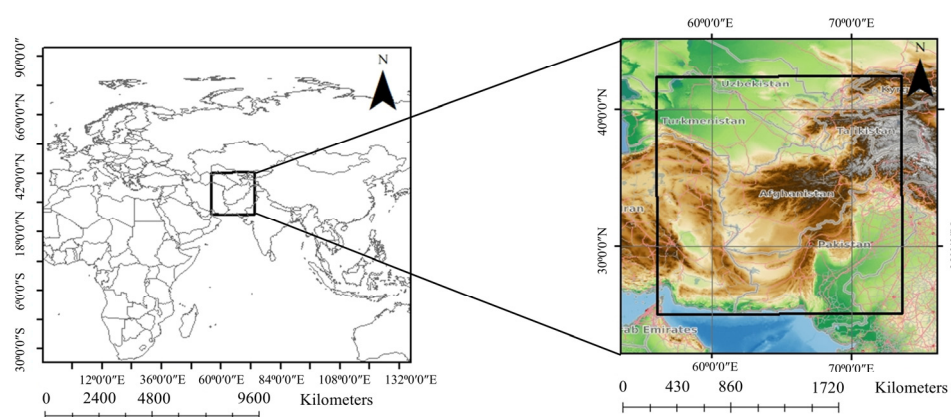


Figure 1. Study region: West, Central, and South Asia.

There are about 40 dusty days every year in Central Asia due to the gradual shrinking of the Aral Sea [32]. The Sistan Basin is identified as a major source of dust in the region, primarily due to the desiccation of the Hamoun lakes during the summer months [22]. Massive dust storms that cover the region are facilitated by favorable atmospheric dynamics, including the strong Levar flow [33]. In addition, dust storms are frequent in Pakistan because the country has an arid or semi-arid landscape covering more than 80% of its territory [34]. Winds can carry fine dust particles thousands of kilometers downwind from these sources.

2.2. Data

The second Modern-Era Retrospective Analysis for Research and Applications (MERRA-2), an atmospheric reanalysis dataset developed by NASA's Goddard Space Flight Center Global Modeling and Assimilation Office (GMAO), was selected for this study to obtain the AOD dataset (<https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/>, accessed on 5 February 2023). Both ground-based and satellite-based AOD retrievals are included in the MERRA-2 AOD dataset. It incorporates the assimilation of bias-corrected AOD from the Advanced Very-High-Resolution Radiometer (AVHRR) and Moderate-Resolution Imaging Spectroradiometer (MODIS), AOD over bright surfaces obtained from the Multi-angle Imaging SpectroRadiometer (MISR), and AOD from the AErosol RObotic NETwork (AERONET) [35]. The aerosol lifetimes are influenced by the sea-surface temperature and sea ice, daily emissions from volcanic activity and biomass burning, and high-resolution inventories of anthropogenic emission sources. MERRA-2 land surface estimates do not directly assimilate land surface observations; instead, they represent the time integration of surface meteorological conditions as modeled by the GEOS-5 land model [36]. The MERRA-2 aerosol reanalysis offers hourly aerosol parameter simulations from 1980 to the present, featuring a spatial resolution of $0.5^\circ \times 0.625^\circ$. Furthermore, various environmental context information, including surface wind speed and direction, surface air temperature, surface skin temperature, relative humidity, and 10 m and 50 m above-ground wind directions, have been used in this study, which are also provided by MERRA-2. The most effective parameters were selected using a feature selection approach. In addition to the environmental factors previously discussed, surface vegetation and soil type can significantly affect soil erosion. For instance, vegetation can prevent soil particles from lifting from the surface. While this study did not specifically examine these aspects, future research could benefit from considering the effects of these factors on dust particle transmissions.

The dataset used in this study has four dimensions: (samples, 140, 143, and features), where 140 and 143 represent the spatial dimensions that are constant across all models. The number of features was 1 when training models without contextual information and increased to 6 when contextual information was included. For evaluating the performance of the proposed models, the dataset was divided into training, testing, and validation sets. For the CNN-LSTM model, the dataset was divided into 2710 training samples, 368 test samples, and 307 validation samples. Similarly, for the ConvLSTM model, the dataset consisted of 2759 training samples, 363 test samples, and 309 validation samples. Additionally, for training the random forest model, 1000 samples with 8 features were used.

In this work, a specific system was utilized for training the models. This system, due to its features, such as an Intel Core i7 processor, 24 GB of RAM, and an NVIDIA GeForce RTX 3060 with 6 GB of dedicated memory, is capable of handling complex models efficiently. Comparing the models, it was observed that the ConvLSTM model required a considerably higher training time.

Convolutional Neural Networks (CNNs) rely heavily on ground truth data, which comprises precisely labeled datasets used as a benchmark during model training and validation. Different approaches have been utilized in various research to produce this crucial information [37,38]. In most machine learning problems, the large number of data is often labeled manually for training classifiers and modeling features, which is time-consuming and tedious [39]. To automatically execute this process by machine, we present the unsupervised and automatic labeling algorithm, which uses Otsu threshold [40]. Otsu thresholding is an automatic method to obtain an optimal threshold value for image segmentation. After this threshold has been obtained, pixels whose values are greater than the threshold are assigned to class one (dust pixels) and those with less value than the threshold are assigned to class zero (non-dust pixels). Images taken every hour in a dust

storm were evaluated and divided into two groups in a binary manner based on the Otsu threshold. Subsequently, the average threshold values obtained from all images were used to create binary-labeled images.

2.3. Methodology

To predict long-term dust storm transport pathways, two deep learning models, namely CNN-LSTM and ConvLSTM, were developed. The methods and resources used to achieve this objective are explained in this section.

2.3.1. CNN

The Convolutional Neural Network (CNN) was introduced by Lecun et al. [41]. It consists of an input layer, an output layer, and multiple hidden layers, which are categorized into convolutional, pooling, and fully connected layers. Feature extraction, parameter reduction, and training velocity are accomplished by the convolutional and pooling layers [42]. Convolutional layers use filters on input images to identify features like edges, textures, or patterns. Pooling layers then reduce the spatial dimensions of the convolved feature maps to decrease computation and control overfitting. In a fully connected layer, each neuron is linked to every neuron in the subsequent layer, allowing for classification based on the features that have been learned. The CNN begins by receiving the input image, with each pixel serving as an input neuron to the network. The network slowly extracts features from the information as it passes through the layers.

2.3.2. LSTM

Traditional deep learning architectures are not well-equipped to learn the temporal features present in time-series data. In contrast, Recurrent Neural Networks (RNNs) excel at capturing these temporal dependencies through memory cells that retain information from previous inputs [43]. LSTM [44] is a variant of RNNs. Unlike traditional RNNs, which often experience vanishing or exploding gradients when processing previous units sequentially, LSTM introduces a cell state that helps retain long-term memory. This allows the network to determine which states to maintain or discard, effectively addressing the gradient issues. Each memory cell contains three gates: an input gate, an output gate, and a forget gate. These gates regulate the flow of information into and out of the memory cell, enabling the network to selectively retain or discard information. The cell state is updated according to the forget gate and the input gate. The input gate manages the influx of new information into the memory cell, while the forget gate governs the removal of outdated information. The output gate decides whether the information stored in the memory cell can be used as output [45]. The LSTM network produces an output based on the modified cell state and the input sequence.

2.3.3. CNN-LSTM

Hybrid deep learning models are a type of neural network architecture that combines different types of deep learning models as well as other kinds of machine learning models to solve complex problems. These models use the strengths of different types of models and can be very effective in solving problems that are difficult to solve with a single model. CNN and LSTM networks are merged to process sequential data with spatial and temporal dependencies. CNNs are effective in learning spatial features from input data, such as images or text [46]. Furthermore, LSTM networks are designed to handle the challenge of long-term dependencies in sequential data [45].

In the CNN-LSTM architecture, the CNN layer consists of an input layer that receives the variables as the input, an output layer that feeds the features to the LSTM component, and a certain number of hidden layers. The output features from the CNN component are

fed into the LSTM network component, which processes them sequentially over time. The LSTM component considers spatial features extracted by CNNs and temporal dependencies in the data. This allows the network to efficiently capture both the spatial and temporal aspects of the input data [46]. The input to the network is typically a 3D tensor, where the first two dimensions represent the spatial dimensions of the input (e.g., height and width of an image) and the third dimension represents the temporal dimension (e.g., time steps for time-series data). CNN layers operate in the first two spatial dimensions, while the LSTM network layer operates in the third temporal dimension.

In this study, four convolutional layers were stacked with the number of kernels of 64, and kernel sizes of 3×3 . After each convolutional layer, max pooling was applied. The time distributed layer was used for all convolutional operations to preserve the temporal dimension. The LSTM part was integrated after the spatial features were extracted with the convolutional layers. In addition to the LSTM layer with 600 memory cells, two fully connected layers with 64 and 280,280 neurons were used to create the predictor. The final cascaded architecture of the CNN-LSTM is illustrated in Figure 2.

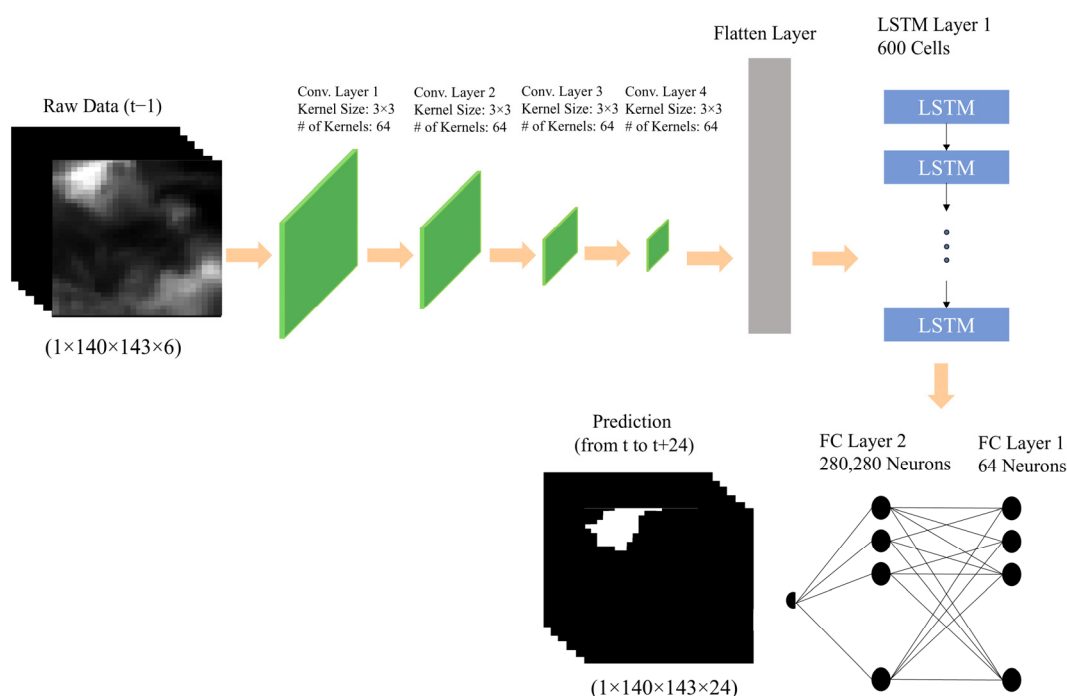


Figure 2. Architecture of proposed CNN-LSTM model.

2.3.4. ConvLSTM

ConvLSTM networks are a type of neural network architecture that combines convolutional layers and LSTM network layers. Combining these two types of layers allows ConvLSTM networks to model spatial and temporal information in the input data. ConvLSTM networks extend traditional LSTMs by replacing the fully connected layers with convolutional layers, allowing them to process spatial information efficiently.

In this study, six ConvLSTM layers were used—the first three layers had 64 filters and the following three layers had 32 filters. Finally, two fully connected layers were used for prediction by analyzing the extracted features. The final cascaded architecture of the ConvLSTM model is illustrated in Figure 3.

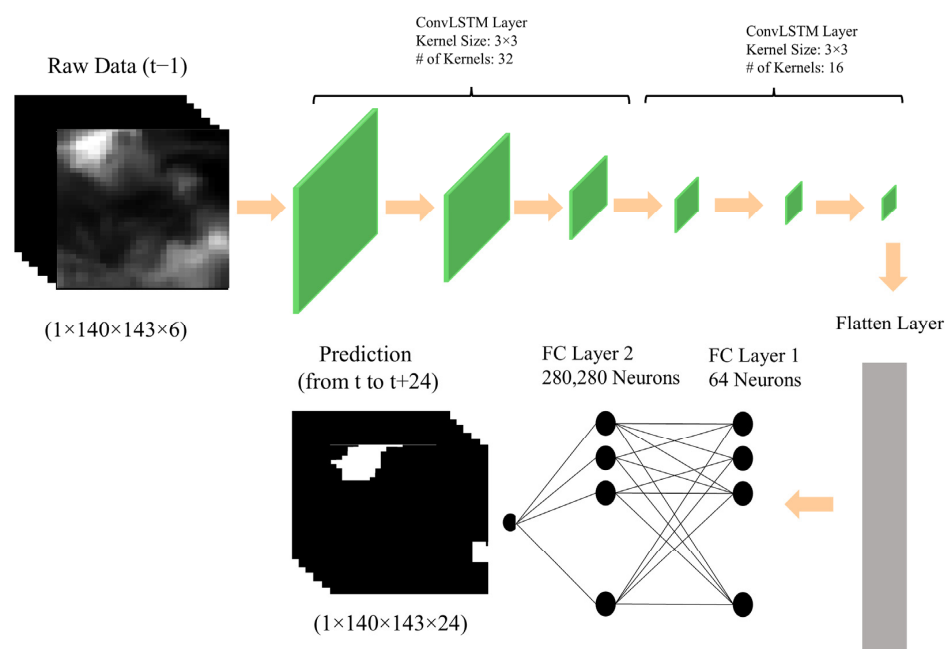


Figure 3. Architecture of proposed ConvLSTM model.

In both models, except for the last fully connected layer that used the sigmoid function, the remaining layers utilized the ReLU activation function. In the next step, the network learning rate was determined. In this study, the learning rate was adjusted from 0.1 to 0.00001, and ultimately, a value of 0.0005 was selected. The early-stopping method from the Keras library was employed to determine the optimal number of epochs. The number of epochs for CNN-LSTM was 150, while for ConvLSTM it was 80.

In the following step, the cost function was optimized. For this purpose, the model was trained using the Dice, Tversky, and Lovasz SoftMax functions, with the Dice function being selected as the optimal cost function. In the final step, the network optimization method was chosen. The model was trained using common optimization algorithms: SGD, RMSProp, Adam, Nadam, and Adamax; ultimately, the Adam algorithm was selected.

2.4. Evaluation Metrics

The performances of the proposed models were evaluated using the recall, precision, overall accuracy, Kappa coefficient, and F1 score. Figure 4 represents a confusion matrix used for deriving these metrics. A true positive (TP) refers to a positive outcome predicted by the model accurately, while a true negative (TN) signifies the accurate prediction of a negative outcome by the model. False positives (FPs) are positive outcomes incorrectly predicted by the model, and false negatives (FNs) are negative outcomes incorrectly predicted by the model.

Overall, accuracy refers to the percentage of correctly predicted image pixels. An overall accuracy score of zero indicates that the classifier consistently predicts incorrectly, while a score of one signifies that the classifier consistently predicts the correct label. The mathematical formula of overall accuracy is presented in Equation (1).

$$\text{Overall Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \quad (1)$$

The Kappa coefficient measures the degree of agreement between the predicted and observed data. Kappa coefficient values equal to one indicate 100% agreement, while zero

values indicate no agreement between the two criteria. The Kappa coefficient is calculated using Equation (2).

$$k = \frac{\rho_o - \rho_e}{1 - \rho_e}, \quad (2)$$

where ρ_o is the proportion of cases where the raters agreed and ρ_e is the proportion of agreement that would be expected if raters were making random guesses.

Precision, also known as a positive prediction, indicates the number of pixels correctly predicted for each category, and is calculated by using Equation (3). Recall, also known as sensitivity, indicates the number of actual pixels detected in each category, and is calculated using Equation (4). The F1 score measures the balance between precision and recall (Equation (5)). The maximum value of the F1 score is equal to one, which indicates full precision and recall. F1 scores of one indicate that the model is perfect in terms of precision and recall, correctly identifying all positives without producing any false positives or false negatives.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (4)$$

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

		Actual values	
		Positive	Negative
Predicted values	Positive	TP	FP
	Negative	FN	TN

Figure 4. Confusion matrix.

2.5. Random Forest Feature Importance

For prediction purposes, random forests (RFs) combine multiple decision trees. Each tree in a forest makes a prediction and the final prediction is made by aggregating their predictions [47]. A decision tree is a hierarchical structure in which internal nodes represent a test on an attribute, branches represents the test outcomes, and leaf nodes represents a classification based on attribute values. Decision trees are constructed using a divide-and-conquer algorithm, and data are iteratively split based on attribute values until the majority of records are classified under a particular class label. The process selects the best attribute to be split.

RF provides a measure of feature importance that indicates the relative importance of each feature in the prediction process. Unlike most other classifiers, RF performs feature selection directly. By using random forest feature importance (RFFI), we can identify the most influential features or variables within a dataset that can be used to predict a particular outcome [48]. RF models are trained on the dataset and their accuracy is evaluated by evaluating each feature's contribution. Feature importance in the RF model refers to a numerical value that describes the contribution of a feature in an accurate prediction. The importance of a feature is calculated as the decrease in node impurity averaged over all trees in the forest. When a feature is used to split the data, node impurity is reduced. The lower the impurity level, the more important the feature. To assess the significance of individual features, the Mean Decrease in Impurity (MDI) measures the impurity reduction (or purity increase) in a decision tree as a particular feature is utilized for segmentation. The more significant the decrease in the average impurity, the more critical the features

become. A conclusive importance score for each attribute is measured by calculating the average impurity reduction score across all decision trees in the model [49].

The dataset used in the random forest method contained 1000 samples with sizes of 140, 143, and 8, where 8 represents the raw data from dust storms and contextual information. First, a sliding kernel of size 3×3 was applied to each input image. It extracted all the values of neighboring pixels in the kernel and flattened them into a feature vector. The central pixel's value in the label image was stored as the target label. A random forest model was then instantiated with 250 decision trees, with the maximum tree depth set to 7 to avoid over-fitting. Finally, feature importance was calculated with the help of the `feature_importances_` tool in the `keras` library from the trained random forest model, to identify which features contribute most to the classification decision.

3. Results

3.1. RFFI Model Output

The RFFI method is used to select relevant contextual information. First, all the features, including surface wind speed and direction, relative humidity, air temperature, surface skin temperature, as well as wind direction at 10 m and 50 m above the ground surface, are fed into the RF model, and the importance score of each feature is calculated. These features had feature importance values of 0.002, 0.001, 0.0006, 0.0003, 0.0002, 0.0001, and 0.00009. The wind directions at 10 m and 50 m from the ground have the least effect on the forecast and are excluded from the prediction process. These importance scores represent the level of their contribution to the RF classification. To mitigate errors caused by randomness, feature importance calculations are repeated three times, and the importance variables are selected with the average importance scores. These features have a significant impact on the occurrence and movement of dust storms; therefore, using them in the forecasting process can help improve forecast accuracy. It should be noted that, in this study, the wind direction is generated using the U and V components of the wind.

3.2. CNN-LSTM Model Output

To investigate the effect of environmental context information on predicting dust storm pathways, the CNN-LSTM model underwent training both with and without the inclusion of contextual data (surface wind speed and direction, relative humidity, surface air temperature, and surface skin temperature). The model predicted the pathway of the storm in the next 24 h. Precision, recall, overall accuracy, Kappa coefficient, and F1 score are computed using the confusion matrix (Figure 4). Table 1 shows the forecast results for the next 6, 12, 18, and 24 h.

Table 1. Evaluation of the CNN-LSTM model.

Prediction Time	Without Contextual Information		With Contextual Information				
	Recall	Precision	Recall	Precision	Overall Accuracy	F1 Score	Kappa Coefficient
t + 6	0.7122	0.8	0.7136	0.8160	0.9825	0.7609	0.612
t + 12	0.7475	0.8552	0.7559	0.8608	0.9852	0.8051	0.734
t + 18	0.7313	0.842	0.7234	0.8688	0.9856	0.792	0.766
t + 24	0.7256	0.8386	0.7242	0.8689	0.9853	0.7939	0.769

Figure 5 shows the predicted and actual shapes of a dust storm originating from Turkmenistan at t, t + 6, t + 12, t + 18, and t + 24. These results were obtained by using the CNN-LSTM method and considering the selected environmental context information.

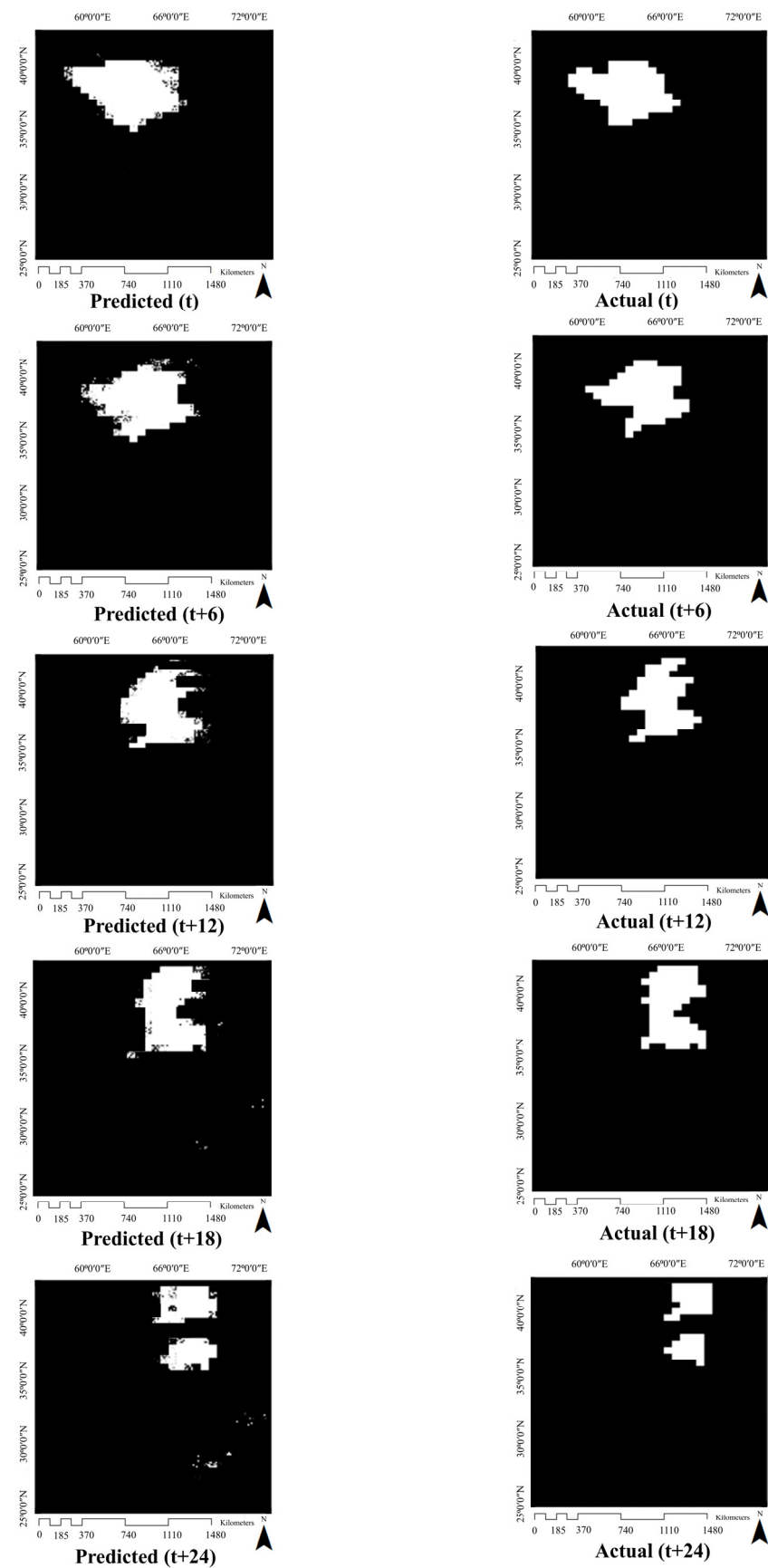


Figure 5. Predicted and actual results for a sample dust storm using the CNN-LSTM method at t , $t + 6$, $t + 12$, $t + 18$, and $t + 24$.

3.3. ConvLSTM Model Output

Similar to the CNN-LSTM model, to evaluate the performance of the ConvLSTM model and the effect of contextual information, all the steps performed in the CNN-LSTM model are implemented. The confusion matrix, precision, recall, overall accuracy, F1 score, and Kappa coefficient values are used to examine the classification accuracy. Table 2 shows the forecast results for the next 6, 12, 18, and 24 h.

Table 2. Evaluation of the ConvLSTM model.

Prediction Time	Without Contextual Information		With Contextual Information				
	Recall	Precision	Recall	Precision	Overall Accuracy	F1 Score	Kappa Coefficient
t + 6	0.684	0.826	0.759	0.8271	0.9682	0.7919	0.5379
t + 12	0.7	0.831	0.752	0.8246	0.9679	0.7871	0.5433
t + 18	0.685	0.851	0.7479	0.8266	0.968	0.7852	0.5435
t + 24	0.641	0.853	0.739	0.8169	0.9676	0.7756	0.5465

The predicted and actual shapes of a sample dust storm originating from Turkmenistan at t, t + 6, t + 12, t + 18, and t + 24 are shown in Figure 6. These results were obtained by using the ConvLSTM method and by including contextual information.

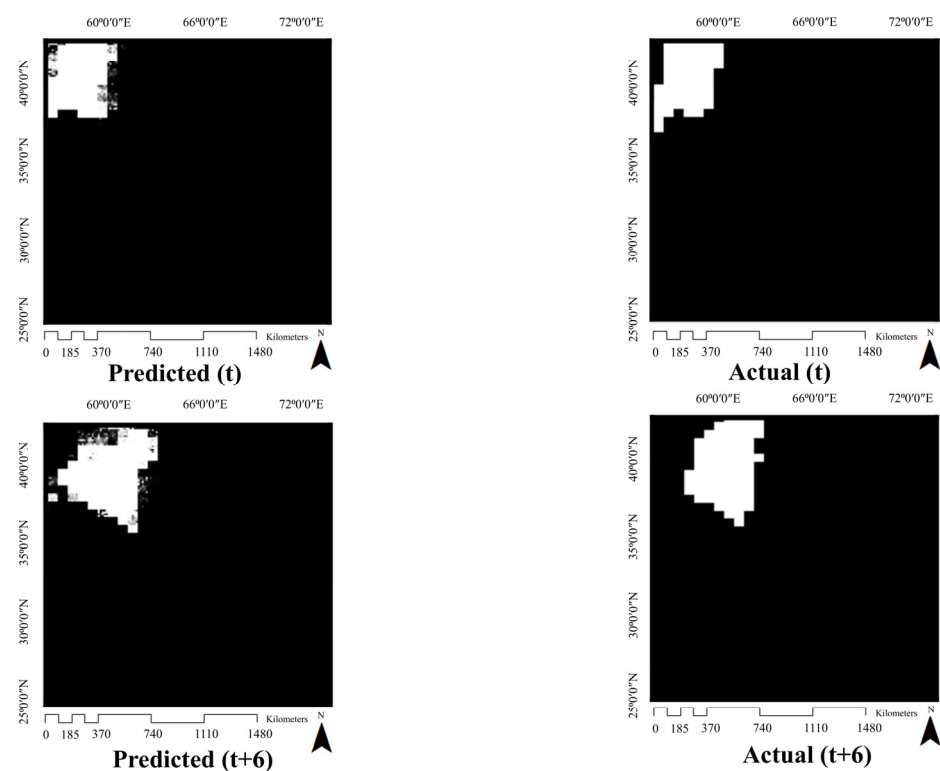


Figure 6. Cont.

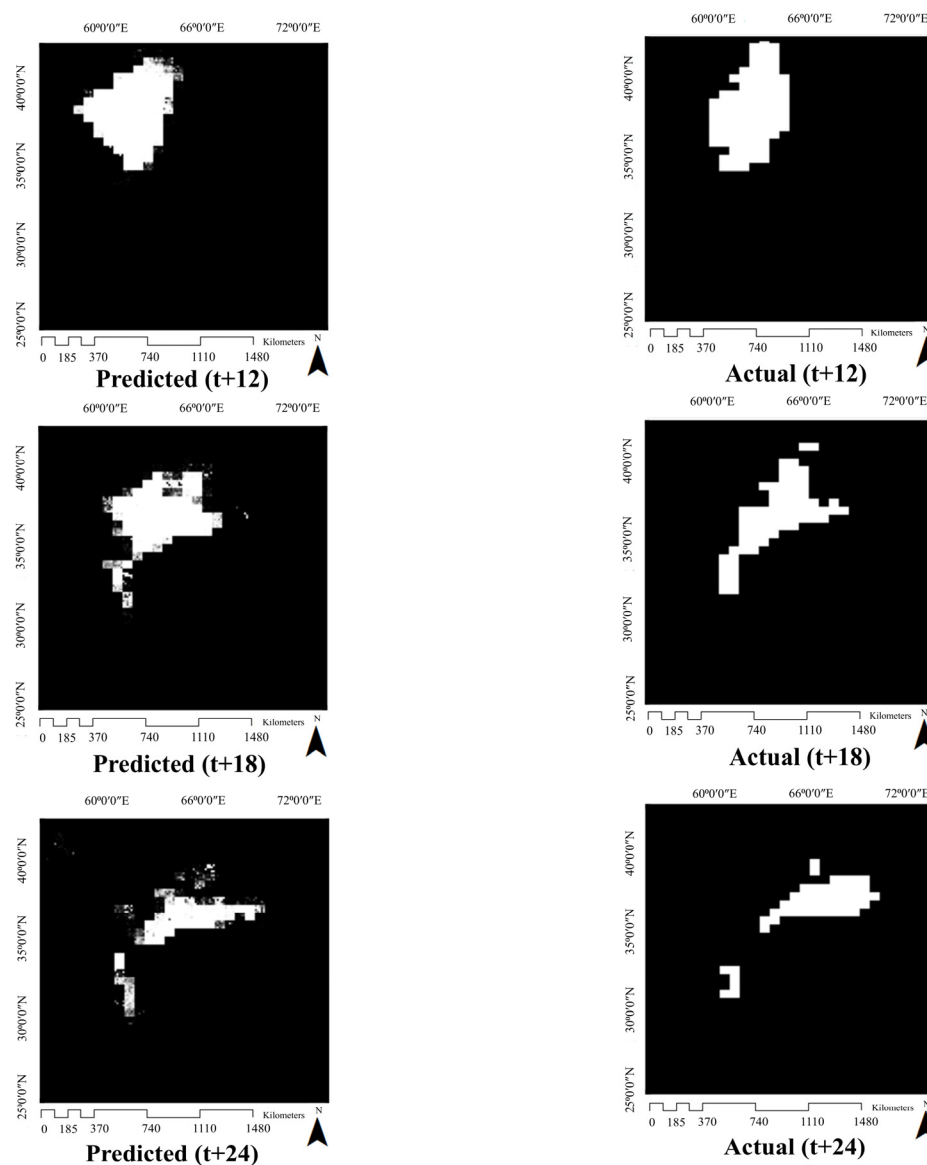


Figure 6. Predicted and actual results for a sample dust storm using ConvLSTM at t , $t + 6$, $t + 12$, $t + 18$, and $t + 24$.

4. Discussion

4.1. The Effect of Contextual Information on Prediction Accuracy

In the CNN-LSTM model, by incorporating contextual information, the model obtained recall values of 0.7136, 0.7559, 0.7234, and 0.7242 for $t + 6$, $t + 12$, $t + 18$, and $t + 24$ -time steps, respectively. In contrast, before incorporating contextual information, the model obtained recall values of 0.7122, 0.7475, 0.7313, and 0.7256 for the same time steps. The results show that adding contextual information does not significantly impact the recall values. However, the precision improved from 0.8, 0.8552, 0.842, and 0.8386 to 0.8160, 0.8608, 0.8688, and 0.8689 for $t + 6$, $t + 12$, $t + 18$, and $t + 24$ time steps, respectively, after incorporating contextual information (Table 1). The increased precision indicates that the CNN-LSTM model is better at minimizing the false-positive rate. Incorporating contextual information enables models to refine their predictions and reduces cases where non-dust storm events are misclassified as dust storms. This improvement in accuracy has practical implications, especially in scenarios where the consequences of false alarms are costly or destructive. By effectively distinguishing between true-positive and false-positive predictions, the model contributes more reliable predictions and decision making. How-

ever, the sensitivity remains relatively constant. This indicates that the CNN-LSTM models consistently capture the same proportion of true-positive samples as before integrating the contextual information.

By incorporating contextual information into the ConvLSTM model, the model obtained precision values of 0.8271, 0.8246, 0.8266, and 0.8169 for $t + 6$, $t + 12$, $t + 18$, and $t + 24$ time steps, respectively. In contrast, before incorporating contextual information, the model obtained precision values of 0.826, 0.831, 0.851, and 0.853 for the same time steps. The results show that the addition of contextual information does not significantly impact the precision values. However, the recall improved from 0.684, 0.7, 0.685, and 0.641 to 0.759, 0.752, 0.7479, and 0.739 for $t + 6$, $t + 12$, $t + 18$, and $t + 24$ time steps, respectively, after incorporating contextual information (Table 2). This indicates that the ConvLSTM model can now identify dust storms better and predict them more accurately. Integrating contextual information into the ConvLSTM model increases spatial awareness and enables the model to better understand and exploit the inherent spatial relationships of time series.

4.2. Comparison the the Prediction Accuracy of CNN-LSTM and ConvLSTM Models

Both CNN-LSTM and ConvLSTM models developed in this research have had success in predicting the movement path of dust storms. The reason for this is the artificial neural network's capability in solving non-linear problems. Adding the LSTM component to the CNN improved the Kappa coefficient from 0.5379, 0.5433, 0.5435, and 0.5465 for the next 6, 12, 18, and 24 h in the ConvLSTM model to 0.612, 0.734, 0.766, and 0.769 for the same time steps in the CNN-LSTM model, respectively. However, the F1 score remained almost constant, with values of 0.7609, 0.8051, 0.792, and 0.7939 for the next 6, 12, 18, and 24 h in the CNN-LSTM model, and values of 0.7919, 0.7871, 0.7852, and 0.7756 in the ConvLSTM model, respectively. Incorporating LSTM layers into CNN models enables them to consider temporal dependencies in addition to spatial features.

Compared to the literature, Tiancheng et al. [13] developed an improved CNN model to predict dust storms using ground and meteorological data. Their model achieved an overall accuracy of 0.883. Notably, our proposed models consistently surpass a 0.95 accuracy for a 24 h forecast. Therefore, context information enhances their ability to learn evolving patterns over time, helping to reconcile the predicted and actual results. This is particularly useful when dealing with time-series data where temporal dependencies are critical. Yarmohammadi et al. [49] developed a CNN network that predicts pathway of a dust storm in the subsequent 24 h. Because their model did not consider temporal dependencies, as the prediction time increased, the Kappa coefficient values decreased from 0.7360 at $t + 6$ to 0.6625 at $t + 24$, and the F1 score value decreases from 0.7497 at $t + 6$ to 0.6769 at $t + 24$. In contrast, the value of the Kappa coefficient for our CNN-LSTM model improved from 0.612 at $t + 6$ to 0.769 at $t + 24$, and the value of the F1 score improved from 0.7609 at $t + 6$ to 0.7939 at $t + 24$. An increase in the Kappa coefficient in CNN-LSTM models indicates their greater ability to recognize patterns and underlying data relationships, which leads to more favorable and stable results in predictions. In addition, the CNN-LSTM model has a better Kappa coefficient than the ConvLSTM model. The Kappa coefficient values for the CNN-LSTM model are 0.612, 0.734, 0.766, and 0.769 for the next 6, 12, 18, and 24 h, and 0.5379, 0.5433, 0.5435, and 0.5465 for the same time steps for the ConvLSTM model, respectively (Tables 1 and 2), which shows the positive effect of using CNN and LSTM networks simultaneously. The CNN-LSTM model combines the strengths of CNN and LSTM layers, allowing it to effectively capture spatial and temporal dependencies in time-series data. This performance improvement can be attributed to the ability of CNN-LSTM to model local patterns and long-term dependencies in time series, leading to more accurate predictions. However, the ConvLSTM model, which integrates

the convolution operation in LSTM layers, shows a lower level of agreement than another model. The ConvLSTM architecture may have had difficulty capturing the time-series data's complex dynamics and patterns.

The calculated metrics demonstrate the accuracy and reliability of the model predictions. This level of accuracy was achieved by using only one previous time step. In another study, the satellite images of dust storm and meteorological data from the previous 15 days were employed to forecast the dust storm's occurrence on the 16th day [14]. In addition, our research extends the predictive horizon beyond a single day, focusing on the trajectory of dust storms over 24 h. The ability of these models to predict dust storms can be attributed to their ability to extract multi-scale features and dependencies. Although these models are strong, they have some limitations, such as requiring high computing power for training. In addition, the spatial resolution of MERRA-2 may reduce the accuracy of predictions. MERRA-2 is based on numerical weather prediction models that may not fully represent the real atmosphere. Despite these limitations, MERRA-2 has a temporal resolution of one hour, allowing hourly dust storm predictions. Furthermore, MERRA-2 has been available since 1980, providing long-term data analysis.

5. Conclusions

This study proposed two hybrid deep learning models, namely CNN-LSTM and ConvLSTM, for the prediction of dust storm transport pathways incorporating environmental context information. Both models were trained on raw MERRA-2 AOD data together with surface wind speed, temperature, humidity, and labeled storm images. The performances of both models were evaluated in two conditions: considering spatial information and considering both spatial and contextual information. This contextual information that has the greatest impact on the dust storm pathway prediction process was selected using the RFFI method. The results show that the hybrid models can effectively capture temporal and spatial dependencies. The added contextual information improved recall in the ConvLSTM model and precision in the CNN-LSTM model, thus establishing the importance of such data for correct predictions. Furthermore, the addition of contextual information increased the accuracy of the predictions and generally improved the performance of the models. The hybrid models, which combined convolutional and Long Short-Term Memory networks, could process information with higher complexity and longer temporal dependencies, hence improving the prediction capability. Detecting dust storms involves analyzing AOD datasets to identify periods of significant increases in aerosol concentration, indicative of dust events. The recommendation for further studies is to develop a more robust method for detecting dust particles in satellite images to avoid the false-positive and false-negative values generated from threshold definitions.

Author Contributions: Conceptualization, M.S. and A.A.A.; methodology, M.Y. and M.S.; software, M.Y.; validation, M.Y.; formal analysis, M.Y. and M.S.; data curation, M.Y.; writing—original draft preparation, M.Y. and M.S.; writing—review and editing, M.S. and A.A.A.; supervision, A.A.A. and M.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available on request from the corresponding authors.

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

References

- Manisalidis, I.; Stavropoulou, E.; Stavropoulos, A.; Bezirtzoglou, E. Environmental and health impacts of air pollution: A review. *Front. Public Health* **2020**, *8*, 14. [\[CrossRef\]](#)
- Wang, W.; Samat, A.; Abuduwaili, J.; Ge, Y.; De Maeyer, P.; Van de Voorde, T. Temporal characterization of sand and dust storm activity and its climatic and terrestrial drivers in the Aral Sea region. *Atmos. Res.* **2022**, *275*, 106242. [\[CrossRef\]](#)
- Han, X.; Liu, L.; Shi, P.; Zhang, G.; Liu, Y.; Zhang, X.; Yang, Y.; Qu, Z.; Wang, G.; Hu, Z.; et al. Dust Storm and Haze Convergence Process and its Effect on Air Pollution in Northern China. *SSRN Electron. J.* **2022**. [\[CrossRef\]](#)
- Miri, A.; Middleton, N. Long-term impacts of dust storms on transport systems in south-eastern Iran. *Nat. Hazards* **2022**, *114*, 291–312. [\[CrossRef\]](#)
- Miri, A.; Shirmohammadi, E.; Sorooshian, A. Impacts of dust storms on indoor and outdoor bioaerosol concentration in the Sistan region of Iran. *J. Build. Eng.* **2023**, *76*, 107302. [\[CrossRef\]](#)
- Zheng, C.; Li, Z.-q.; Zhou, P.; Zhang, X.; Zhou, X.; Ma, S. Physicochemical Impacts of Dust Storms on Aerosol and Glacier Meltwater on the Northern Margin of the Taklimakan Desert. *Front. Earth Sci.* **2021**, *8*, 527663. [\[CrossRef\]](#)
- Liang, L.; Han, Z.; Li, J.; Xia, X.; Sun, Y.; Liao, H.; Liu, R.; Liang, M.; Gao, Y.; Zhang, R. Emission, transport, deposition, chemical and radiative impacts of mineral dust during severe dust storm periods in March 2021 over East Asia. *Sci. Total Environ.* **2022**, *852*, 158459. [\[CrossRef\]](#) [\[PubMed\]](#)
- Khennou, F.; Ghaoui, J.; Akhloufi, M.A. Forest fire spread prediction using deep learning. In *Geospatial Informatics XI*; SPIE: Bellingham, WA, USA, 2021.
- Wang, R.; Zhu, Z.; Zhu, W.; Fu, X.; Xing, S. A dynamic marine oil spill prediction model based on deep learning. *J. Coast. Res.* **2021**, *37*, 716–725. [\[CrossRef\]](#)
- Draxler, R.R.; Hess, G. An overview of the HYSPLIT_4 modelling system for trajectories. *Aust. Meteorol. Mag.* **1998**, *47*, 295–308.
- Khalidy, R.; Salmabadi, H.; Saeedi, M. Numerical simulation of a severe dust storm over Ahvaz using the HYSPLIT model. *Int. J. Environ. Res.* **2019**, *13*, 161–174. [\[CrossRef\]](#)
- Stein, A.F.; Draxler, R.R.; Rolph, G.D.; Stunder, B.J.; Cohen, M.D.; Ngan, F. NOAA's HYSPLIT atmospheric transport and dispersion modeling system. *Bull. Am. Meteorol. Soc.* **2015**, *96*, 2059–2077. [\[CrossRef\]](#)
- Li, T.; Ren, Q.-d.-e.-j.; Qiu, Y. Application of improved naive bayesian-CNN classification algorithm in sandstorm prediction in inner mongolia. *Adv. Meteorol.* **2019**, *2019*, 5176576.
- Li, N.; Zhang, W. Research on sand-dust storm forecasting based on deep neural network with stacking ensemble learning. *IEEE Access* **2022**, *10*, 111855–111863.
- Yarmohamadi, M.; Alesheikh, A.A.; Sharif, M.; Vahidi, H. Predicting dust-storm transport pathways using a convolutional neural network and geographic context for impact adaptation and mitigation in urban areas. *Remote Sens.* **2023**, *15*, 2468. [\[CrossRef\]](#)
- Sadaria, D.P.; Khachariya, D.H.; Hirpara, D.J. Exploring the Diverse Applications of Deep Learning Across Multiple Domains. *rrrj* **2023**, *2*, 183–200. [\[CrossRef\]](#)
- Farmanifard, S.; Alesheikh, A.A.; Sharif, M. A context-aware hybrid deep learning model for the prediction of tropical cyclone trajectories. *Expert Syst. Appl.* **2023**, *231*, 120701. [\[CrossRef\]](#)
- Sharif, M.; Alesheikh, A.A. Context-aware movement analytics: Implications, taxonomy, and design framework. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2018**, *8*, e1233. [\[CrossRef\]](#)
- Sharif, M.; Alesheikh, A.A.; Tashayo, B. CaFIRST: A context-aware hybrid fuzzy inference system for the similarity measure of multivariate trajectories. *J. Intell. Fuzzy Syst.* **2019**, *36*, 5383–5395. [\[CrossRef\]](#)
- Sarafian, R.; Nissenbaum, D.; Raveh-Rubin, S.; Agrawal, V.; Rudich, Y. Deep multi-task learning for early warnings of dust events implemented for the Middle East. *Npj Clim. Atmos. Sci.* **2023**, *6*, 23. [\[CrossRef\]](#)
- Csavina, J.; Field, J.; Taylor, M.P.; Gao, S.; Landázuri, A.; Betterton, E.A.; Sáez, A.E. A review on the importance of metals and metalloids in atmospheric dust and aerosol from mining operations. *Sci. Total Environ.* **2012**, *433*, 58–73. [\[CrossRef\]](#) [\[PubMed\]](#)
- Khusfi, Z.E.; Khosroshahi, M.; Roustaei, F.; Mirakbari, M. Spatial and seasonal variations of sand-dust events and their relation to atmospheric conditions and vegetation cover in semi-arid regions of central Iran. *Geoderma* **2020**, *365*, 114225. [\[CrossRef\]](#)
- Rashki, A.; Kaskaoutis, D.; Goudie, A.S.; Kahn, R. Dryness of ephemeral lakes and consequences for dust activity: The case of the Hamoun drainage basin, southeastern Iran. *Sci. Total Environ.* **2013**, *463*, 552–564. [\[CrossRef\]](#)
- Soleimani, Z.; Teymouri, P.; Bolorani, A.D.; Mesdaghinia, A.; Middleton, N.; Griffin, D.W. An overview of bioaerosol load and health impacts associated with dust storms: A focus on the Middle East. *Atmos. Environ.* **2020**, *223*, 117187. [\[CrossRef\]](#)
- Khashi, K.; Azhdary Moghaddam, M.; Hashemi Monfared, S. Effects of wetlands and wind velocity on dust propagation in arid areas: A critical case study in south-east of Iran. *Int. J. Environ. Sci. Technol.* **2022**, *19*, 12169–12180. [\[CrossRef\]](#)
- Cao, H.; Liu, J.; Wang, G.; Yang, G.; Luo, L. Identification of sand and dust storm source areas in Iran. *J. Arid Land* **2015**, *7*, 567–578. [\[CrossRef\]](#)
- Wang, R.; Liu, B.; Li, H.; Zou, X.; Wang, J.; Liu, W.; Cheng, H.; Kang, L.; Zhang, C. Variation of strong dust storm events in Northern China during 1978–2007. *Atmos. Res.* **2017**, *183*, 166–172. [\[CrossRef\]](#)

28. Samman, A.E.; Butt, M.J. Aerosol Types and Their Climatology over the Dust Belt Region. *Atmosphere* **2023**, *14*, 1610. [\[CrossRef\]](#)
29. Shen, H.; Abuduwaili, J.; Samat, A.; Ma, L. A review on the research of modern aeolian dust in Central Asia. *Arab. J. Geosci.* **2016**, *9*, 625. [\[CrossRef\]](#)
30. Francis, D.; Fonseca, R.M.; Nelli, N.R.; Bozkurt, D.N.; Cuesta, J.; Bosc, E. On the Middle East's severe dust storms in spring 2022: Triggers and impacts. *Atmos. Environ.* **2022**, *296*, 119539. [\[CrossRef\]](#)
31. Eleftheriou, A.; Mouzourides, P.; Biskos, G.; Yiallourous, P.; Kumar, P.; Neophytou, M.K.-A. The challenge of adopting mitigation and adaptation measures for the impacts of sand and dust storms in Eastern Mediterranean Region: A critical review. *Mitig. Adapt. Strateg. Glob. Change* **2023**, *28*, 33. [\[CrossRef\]](#)
32. Indoitu, R.; Orlovsky, L.; Orlovsky, N. Dust storms in Central Asia: Spatial and temporal variations. *J. Arid Environ.* **2012**, *85*, 62–70. [\[CrossRef\]](#)
33. Kaskaoutis, D.G.; Francis, D.; Rashki, A.; Chaboureau, J.-P.; Dumka, U.C. Atmospheric dynamics from synoptic to local scale during an intense frontal dust storm over the Sistan Basin in winter 2019. *Geosciences* **2019**, *9*, 453. [\[CrossRef\]](#)
34. Hussain, A.; Mir, H.; Afzal, M. Analysis of dust storms frequency over Pakistan during 1961–2000. *Pak. J. Meteorol.* **2005**, *2*, 49–68.
35. Randles, C.; Da Silva, A.; Buchard, V.; Colarco, P.; Darmenov, A.; Govindaraju, R.; Smirnov, A.; Holben, B.; Ferrare, R.; Hair, J. The MERRA-2 aerosol reanalysis, 1980 onward. Part I: System description and data assimilation evaluation. *J. Clim.* **2017**, *30*, 6823–6850. [\[CrossRef\]](#)
36. Koster, R.D.; Bosilovich, M.G.; Akella, S.; Lawrence, C.; Cullather, R.; Draper, C.; Gelaro, R.; Kovach, R.; Liu, Q.; Molod, A. *Technical Report Series on Global Modeling and Data Assimilation*; NASA: Washington, DC, USA, 2015; Volume 43.
37. Yoon, J.; Yoo, C.; Ha, E. On the use of threshold for the ground validation of satellite rain rate. *Adv. Meteorol.* **2015**, *2015*, 254378. [\[CrossRef\]](#)
38. Boulila, W.; Ghandorh, H.; Khan, M.A.; Ahmed, F.; Ahmad, J. A novel CNN-LSTM-based approach to predict urban expansion. *Ecol. Inform.* **2021**, *64*, 101325. [\[CrossRef\]](#)
39. Fu, Y.; Huang, T.S. Unsupervised locally embedded clustering for automatic high-dimensional data labeling. In Proceedings of the 2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07, Honolulu, HI, USA, 15–20 April 2007; IEEE: New York, NY, USA, 2007.
40. Otsu, N. A threshold selection method from gray-level histograms. *Automatica* **1975**, *11*, 23–27. [\[CrossRef\]](#)
41. LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**, *86*, 2278–2324. [\[CrossRef\]](#)
42. Xie, G.; Shangguan, A.; Fei, R.; Ji, W.; Ma, W.; Hei, X. Motion trajectory prediction based on a CNN-LSTM sequential model. *Sci. China Inf. Sci.* **2020**, *63*, 212207. [\[CrossRef\]](#)
43. Schmidt, R.M. Recurrent neural networks (rnns): A gentle introduction and overview. *arXiv* **2019**, arXiv:1912.05911.
44. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Yu, Y.; Si, X.; Hu, C.; Zhang, J. A review of recurrent neural networks: LSTM cells and network architectures. *Neural Comput.* **2019**, *31*, 1235–1270. [\[CrossRef\]](#)
46. Yin, M.; Liao, S.; Liu, X.-Y.; Wang, X.; Yuan, B. Compressing recurrent neural networks using hierarchical tucker tensor decomposition. *arXiv* **2020**, arXiv:2005.04366.
47. Farnaaz, N.; Jabbar, M. Random forest modeling for network intrusion detection system. *Procedia Comput. Sci.* **2016**, *89*, 213–217. [\[CrossRef\]](#)
48. Louppe, G.; Wehenkel, L.; Suter, A.; Geurts, P. Understanding variable importances in forests of randomized trees. *Adv. Neural Inf. Process. Syst.* **2013**, *26*, 431–439.
49. Scornet, E. Trees, forests, and impurity-based variable importance in regression. In *Annales de l'Institut Henri Poincaré (B) Probabilités et Statistiques*; Institut Henri Poincaré: Paris, France, 2023.

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.