Diagnosis and assessment of pre-fog in the Mainland Portuguese International Airports: statistical and Neural Network models comparison †

Pedro M. P. Guerreiro [1,*], Gonçalo Cruz [2]

[1] Portuguese Air Force Academy Research Center, 2715-021 Pêro Pinheiro, Portugal; pmguerreiro@academiafa.edu.pt

[2] Portuguese Air Force Academy Research Center, 2715-021 Pêro Pinheiro, Portugal; gccruz@academiafa.edu.pt

* Correspondence: pmguerreiro@academiafa.edu.pt; Tel.: +351-914-005-767

† Presented at the title, place, and date.

Abstract: The prediction of fog is a challenging task in operational weather forecast. Due to its dependency on small scale processes, numerical weather models struggle to deal with under scale features resulting in uncertainties on the fog forecast. Unawareness of the onset time and the duration of fog leads to disproportionate impact on open air activities, especially in aviation. Nevertheless, in a small sized country such as Portugal mainland, the fog varies greatly. The traffic of the two busiest Portuguese international airports of Porto and Lisbon is affected by the occurrence of fog in different times of the year. The fog occurrence at Porto is a predominant winter phenomenon and a summer one at Lisbon. Observational variables and their trend are local indicators of favouring conditions to the fog onset, like cooling, water vapour saturation and turbulent mixing. A dataset corresponding to 17 years of half hourly METAR from the airports of Porto and Lisbon is used to diagnose the pre fog conditioning. Two diagnostic models are proposed to assess pre fog conditions. The first model is adapted from the statistical method proposed by Menut et al. (2014), which performs a diagnosis from key variables trend, such as temperature, wind speed and relative humidity. Thresholds are defined from the METAR samples in the six-hours period prior to the formation of fog. Due to the local character of fog, the presented thresholds are the most appropriate ones for each airport. The predictability of fog is then assessed using observations. The second approach consists of neural networks such as a fully connected (FC) network and a recurrent neural network (RNN), which are especially well suited for time series. By experimenting with different types of neural networks (NN), we will try to capture the connection between the temporal evolution of measured variables in the dataset and the fog onset. These experiments will include different time windows to measure its influence in the prediction performance.

Keywords: fog prediction; statistics; diagnosis; deep learning; Multiplayer Perceptron; Recurrent Neural Networks

1. Introduction

Among the weather phenomena that affects the visibility range near the Earth's surface, the fog is the one that constrains human activities with most economic impact, and sometimes even jeopardizes human lives. Due to its great incidence on aviation, climatological studies and field campaigns have been focused on busy airports around the world [1,2,3,4,5,6,7].

World Meteorological Organization defines fog as the visibility restriction due to water droplets in the lower atmosphere that reduces the horizontal visibility to less than 1 km near to the ground [8]. The knowledge of local weather conditions that culminate in

fog plays a major role in operational forecast [1]. A study carried out by Policarpo et al. [9] has shown that a large artificial lake of 250 km^2 and its irrigated area become a local important source of moisture, that increases the availability of water vapour, favouring the formation of fog over the lake and its surroundings. Egli et al. [10], described the relation between the terrain characteristics, the predominant weather situations and the classification of fog patterns.

Later, Guerreiro et al. [7] studied the fog characteristics at the Portuguese international airports, the conditions prior to the onset, at large and local scales, and the classification of fog into advection, radiation, cloud base lowering, precipitation and evaporation types.

The connection between the prior conditions to the onset and the fog forecast was proposed by Menut et al. [11] using data from a single instrumental site (SIRTA laboratory), and a statistical methodology to estimate the probability of existing meteorological conditions favourable to radiative fog formation. Later, Róman-Cascón et al. [12] had applied the same methodology using data from two research facilities, one in Netherlands (CESAR) and the other at Spain (CIBA), to evaluate the radiation fog forecast based on the observational data and Weather and Research Forecast model (WRF) output. In the present study the same methodological basis is used to diagnosis conditions extended to all types of fog using METAR data, from the international Portuguese airports.

Given that usually it is difficult to model the behavior of weather phenomena, data-driven methods are attractive to use. NN, in particular, have two theoretical advantages over conventional statistical methods. The first advantage is that no distribution needs to be assumed; the second is that theoretically it can approximate any smooth function. Considering the advantages, neural networks have been used in atmospheric sciences for several decades [13].

One of the first successful types of NN was the multilayer perceptron (MLP). This type of network consists in an arrangement of artificial neurons, organized into layers. Neurons between layers are connected by weights and the output of each neuron is a function of the sum of the inputs, modified by a non-linear activation function. Because each neuron in a layer connects to every neuron in the following, MLP are also usually designated as fully connected (FC) or densely connected networks. One of the main drawbacks of MLP is that the number of weights increases rapidly with the increasing number of neurons and the inputs vectors with higher dimensionality. This makes the adjustment of the weights (commonly called training phase) extremely difficult for problems with many variables.

The limitations of MLP prevented its successful application to image-like data. For this application, an architecture with shared weights was proposed, that performs convolutions of kernels across the input vector. This type is designated as convolutional neural network and is especially suited to explore the spatial structure of the data. This not only allowed super-human performance in classification of common images but was used in several weather prediction tasks [14].

While both types of the presented networks are considered feedforward networks, there is also another family of networks, which is called recurrent neural networks. These are especially adequate to deal with time sequences. In this family, there are two types which are used in the context of atmospheric sciences: Long Short-Term Memory (LSTM) layers [15, 16] and Gated Recurrent Units (GRU) [17]. These units are composed of a cell and gates. The cell has an internal state that retains information over several time instants and the gates control the flow information into and out of the unit. Other approaches, like proposed by Shi et al. [18], have explored methods that have both a convolutional structure and are recurrent. This makes these networks able to deal with a time series of image-like data. In our context, we will not consider image-like data, so we will focus on FC feedforward networks and on LSTMs.

Therefore, the main goal of this study is the assessment of two extended diagnosis methods of favourable conditions to fog formation using aerodrome routine meteorological reports (METAR) data and compare both performances.

In this article, the airport's locations, the used data and the diagnosis methods are described in Section 2. In Section 3, the pre-fog conditions and the main features of fog events are presented. The two diagnosis methods are assessed using the airports METAR data and their performance is presented. In Section 4, results from both methods are discussed and compared. Finally, conclusions are presented in Section 5.

2. Materials and Methods

2.1. Materials

Seventeen years of half-hourly METAR data from the Portuguese mainland international airports were used to identify fog events. The dataset was subjected to quality control after retrieved from the METAR code form, which was generated using observational practices following the WMO Manual on Codes, Volume I, Part A [9]. From January 2002 to December 2018, the data availability is 99.68% at Porto, 99.9% at Lisbon, and 99.92% at Faro.

Fog occurrences are identified in the dataset by gathering observations of prevailing horizontal visibility less than 1000 m, associated with the report of significant weather phenomenon as fog, in the site or in the vicinity of the airport. The results presented by Guerreiro et al. [7] have shown that fog seldom occurs at Faro. In the period of 2002-2018, the daily fog occurrences were less than 1% (0.97%), while at Porto and Lisbon were 9.68% and 5.15% respectively. Therefore, Faro airport is discarded from this study, and only the 2019 data from Porto and Lisbon will be used in the methods' assessment.

2.2. Methods

2.2.1 Forecast score

At both airports, the classification of fog into precipitation, advection, cloud base lowering, radiation, and evaporation types was performed by Guerreiro et al. [7]. Each type associates an observed parameter to a primary mechanism that triggers the formation of fog. From the METAR data, the wind drives the advection and the turbulent mixing, decreasing of temperature measures the cooling rate, and dew point is used to compute relative humidity, that quantifies the amount of water vapor available for condensation.

The departure point of the diagnosis method is the statistical method proposed by Menut et al. [11] used to estimate the probability to detect observed radiation fog events. The forecast score computed using meteorological key parameters reflects the relationship between the meteorological conditions and the formation of fog. Since the method is focused on the detection of radiation fog, the pre-fog conditions are characterized by the key variables of relative humidity measured at 2 m height, 2 m temperature tendency, 10 m wind speed and infrared radiation budget. The METAR does not include infrared radiation data. Therefore, in this study, the key variables are relative humidity (RH), 3-hour temperature tendency (ΔT), and wind speed $U(t)$, regarding any type of fog.

The forecast scores α_{kv}, between 0 and 1, and they are dependent on the distance between the observed values of the key variables and the respective thresholds, following a Gaussian distribution [12]. These reference values are set by the key variables average in the fog events period. The scores are computed as

$$\alpha_{kv} = \begin{cases} 1, & kv \leq th_{kv} \\ exp\left[-\dfrac{(kv - th_{kv})^2}{2\sigma_{kv}}\right], & kv > th_{kv} \end{cases} \tag{1}$$

where kv are the key variables and th_{kv} the respective thresholds.

At a specific observed moment *t*, the forecast scores α_{kv} are computed regarding the relative humidity *RH(t)*, the wind speed *U(t)*, and the 3-hour temperature tendency *ΔT=T(t)-T(t-3h)*.

The diagnosis of the fog formation is finally computed as

$$\alpha = \alpha_{RH} \times \alpha_{\Delta T} \times \alpha_U. \tag{2}$$

For $\alpha \geq 0.9$, the forecast score method states that meteorological conditions, at time *t*, will favour the formation of fog, later in the following 6-hours period (*t+6h*), designated as pre-fog period as well [11, 12]. The forecast score is computed to each time step in the pre-fog period of the fog events detected in 2019. Since α diagnoses the formation of fog up to six hours, from each time step, it means that its performance varies from the moment of the fog onset, the worst estimation, up to six hours prior to the onset, the best estimation. Therefore, when the estimates are $\alpha \geq 0.9$, we consider that a pre-fog epsiode is present; if $\alpha \leq 0.9$, then we consider that there are no pre-fog conditions.

2.2.2 Neural Networks

Neural networks offer the theoretical advantage of not assuming any *a priori* information about the statistical distribution of the input data. Nonetheless, the performance of a given network depends on the type of data and on correctly formatting the data. One of the factors is whether the input is a single data point (*e.g.,* one image at a time) or a temporal sequence of data.

The dataset consists of sequential observations of different variables, therefore, we decided to use a network based on LSTM units, as presented in [19], to explore the time dependence. To assess if there was really any advantage, we compared the performance with a neural network with similar complexity but based on FC neurons. We denote the input data for the time instant t' as $X_{t=t'} = \left[x_{t=t'}^{feature=1} \quad x_{t=t'}^{feature=2} \quad \dots \quad x_{t=t'}^{feature=N} \right]$, where $x_{t=t'}^{feature=1}$, $x_{t=t'}^{feature=2}$ up to $x_{t=t'}^{feature=N}$ are the different features in that time instant. As presented in Figure 2 a), the input vectors for each time instant are supplied sequentially to the LSTM-based NN to create an estimate $\hat{y}_{t=t'}$. This estimate is whether there will be fog onset in the interval $t = t'$ up to $t = t' + 6 \ hours$. For the case of the FC NN, all the input data is stacked and provided to the NN in one shot, as depicted in Figure 2 b).

(**a**)   (**b**)
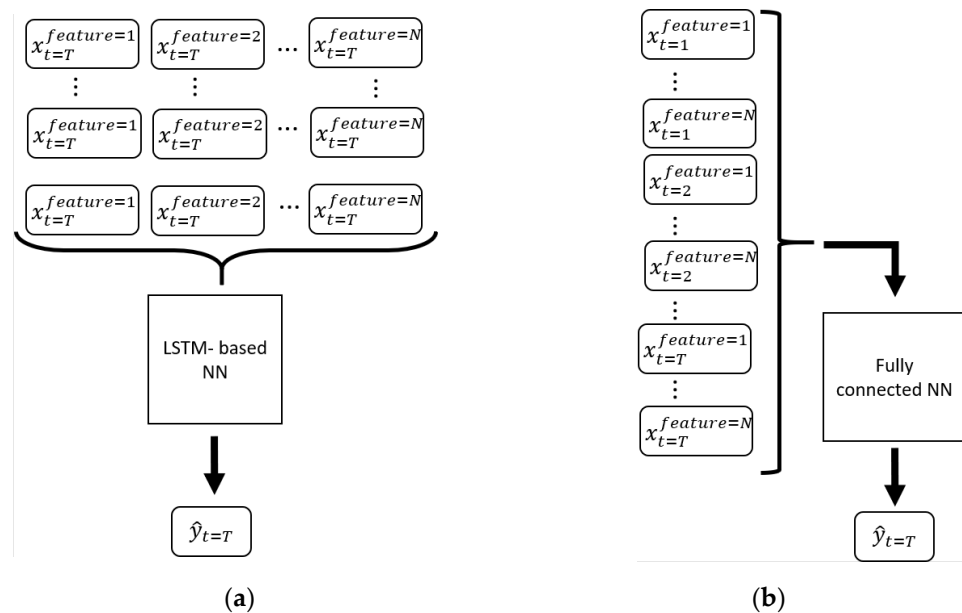
Figure 2. Representation of how the input data is fed to distinct types of networks: (**a**) In the
case of the LSTM-based NN, at each time instant t, the features from that correspondent instant
are provided. The estimate ŷ in the time instant t, is then generated not only based on that
input but is also considering the previous instants; (**b**) For the case of FC networks, the output
is calculated based on all features corresponding to a finite number of time instants. Unlike the
LSTM-based NN, in the FC case, there is no information about which time instant a given value
was observed.

Another factor is the range of values assumed by the input, since NN are usually
designed to deal with normalized data. Additionally, the visibility features (one of the
most important in our problem) has most of the occurrences with very high values and
occasionally decreases one order of magnitude, which correspond to fog. In general, the
normalization approach that was followed consists in a subtraction of the average value
and dividing it by the standard deviation. Both these values correspond to the average
and standard deviation calculated in the training data. This yields to negative and positive
values around 0.

At last, despite not explicitly selecting key variables as in the previous section, some
caution must be devoted to curate the data before providing it to the network. For in-
stance, the dataset had 25 parameters, but we combined them into 13. These combinations
consist in transforming some features into new ones which are more meaningful to the
fog occurrence. For instance, we combined the month and day of the month into only one
variable with a periodic behavior by computing.

$$x_{time\ of\ the\ year} = \frac{2\pi}{12}\cos\left(month + \frac{day\ of\ the\ month}{number\ of\ days\ in\ the\ month}\right) . \tag{3}$$

The main idea of this operation is that, for example in the last days of December will have
a similar value to the first days of January, thus is expected to have a similar contribution
to the prediction of weather phenomena. We did a similar operation with the hours and
minutes, creating a signal with a period corresponding to 24 hours.

$$x_{time\ of\ the\ day} = \cos\left((60 \times hour + minutes) \times \frac{2\pi}{1440}\right) \tag{4}$$

We also transformed the wind intensity and direction into two components, one aligned
with North-South, and another aligned with East-West. At last, we did not consider the
data regarding cloud layers above the lowest layer.

After the data curation, a given network configuration must be designed. The main
building block for the recurrent network is the LSTM cell. This cell contains a cell state c,
receives an input z and computes an output ŷ. The LSTM cell also contains gates that
allow or block the flow of information into and out of the cell and the update of the cell
state. The amount of information allowed by input gate i, output gate i and forget gate
f at time instant t is defined as

$$f_t = \sigma_g(W_{xf}x_t + W_{hf}\ h_{t-1} + b_f) \tag{5}$$

$$i_t = \sigma_g(W_{xi}x_t + W_{hi}\ h_{t-1} + b_i) \tag{6}$$

and

$$o_t = \sigma_g(W_{xo}x_t + W_{ho}\ h_{t-1} + b_o) \tag{7}$$

where W represent weight matrices that connect two quantities (*e.g.*, W_{zo} connects the
input of the cell with the output gate) and b correspond to bias.

With this building block one can build arbitrarily complex network, with several cells
in parallel and multiple consecutive layers. However, to assess the impact of the network
design choices, we started with a simple network with only two LSTM layers. The first
layer contains several cells in parallel and receives the input X_t. The second layer only has

one cell, receives the output of the first one and calculates the network's final output. We
also compared the resulted obtained with a FC layer, also with only two layers and the
same number of neurons as the number of cells in the LSTM-based network. As mentioned previously, the FC network deals with a limited number of time instants. Additionally, during training, the LSTM-based network is presented with multiple samples,
each one comprising observed features from several time instants. We have designated
this amount as time horizon. To understand, the impact of the time horizon on the performance, we have also tested several values.

The task considered in this work is to predict the first fog observation (fog onset) in
the next six hours, therefore, for the training of the neural network, we have framed this
as a classification problem. With this framework, at training time, we not only provide the
observations X from METARs but also the *Y*, which is a binary value. We have used a
common loss function for this type of problem, the binary cross-entropy.

2.2.3 Performance indicators

Using classical forecast score parameters, such as hit rate, false alarm, miss and correct reject, the model performance is evaluated from the following contingence table (Table 1), according to the model output and the fog observations. To simplify the interpretation of the results, we normalize the previously mentioned score parameters. Thus, we
will present **hit rate** which corresponds to the frequency of correct prediction of fog over
the total number of fog occurrences, **false alarm rate** corresponds to the frequency of incorrect fog prediction over the total number negative samples (no observed fog). We also
present **miss rate** and **true negative rate**, which are the complementary values of hit rate
and false alarm rate, respectively.

Table 1. Forecast performance indicators obtained from the combination of the model output and
the fog observations.

	Observed	Not observed
Fog forecasted	HIT RATE (HR)	FALSE ALARM RATE (FAR)
Absence of fog	MISS RATE (MR)	TRUE NEGATIVE RATE (TNR)

3. Results

Results obtained for both methods were obtained by using the METAR data corresponding to the period from 2002 to 2018 as training samples. Additionally, both methods
were also applied in two different locations independently: Porto and Lisbon.

After the training stage, both the forecast score method and NN-based method were
tested in METAR data gathered during 2019. Lisbon's dataset contains 17519 samples
while there are two less samples in Porto's dataset. From the fog observations, 56 episodes
were identified in Porto and 25 in Lisbon. Since our goal is to identify the pre-fog conditions, we want to classify as 1, each of the six hours before a fog episode. All other estimates outside the pre-fog period (fog episodes observations and no fog observations)
should be classified as 0. This yields 663 ground truth samples classified as 1 and the remainder classified as 0.

3.1. Forecast score

In the forecast score method, the diagnosis of fog starts after the third hour from the
first observation due to the 3-hour temperature tendency. Therefore, the estimation of pre-fog, as well as the absence of fog, is performed over 17511 observations at Porto and 17513
at Lisbon.

The following contingency table (Table 2) reflects the performance of the forecast
score method through the indicators from Table 1.

Table 2. Performance indicators of the forecast score's method.

	Porto		Lisbon	
	Observed	**Not observed**	**Observed**	**Not observed**
Fog forecasted	17.19%	6.34%	5.57%	0.84%
Absence of fog	82.8%	93.66%	94.43%	99.16%

The method's performance at Porto presents a HR of 17.19%, and 5,57% at Lisbon. At both airports, the FAR is considerable low, being residual at Lisbon (0.84%), due to many true negative diagnoses. The MR indicator reveals the method's weakness of forecasting fog, with 82.8% at Porto and 94.43% at Lisbon.

3.2. Neural Network-based method

The evaluation of the NN-based method had several objectives. The first was to assess the benefits of using an LSTM over a FC NN. The second was to measure the impact of using input sequences of variable length. Thirdly, we wanted to determine the sensitivity of the NN to the different input variables (observations in the METAR corresponding to temperature, dew point and visibility). Lastly, we wanted to compare with the Forecast Score method.

In Table 3, we present the cross-entropy loss as well as other that was obtained for each configuration for both Porto and Lisbon airports. Each of the network configurations was fed with sequences of input data of variable length (6, 12, 24, 48 and 72 time instants).

Table 3. Performance indicators obtained for each network configuration in Porto and Lisbon, when considering the test dataset (METAR observations from 2019).

Network type	Sequence length	Cross-entropy loss	HR (%)	FAR (%)	MR (%)	TNR (%)
LSTM-based NN - Porto	6	0.02857	47.5	4.3	52.5	95.7
	12	0.02800	47.7	3.8	52.3	96.2
	24	**0.02800**	**49.3**	**3.7**	**50.7**	**96.3**
	48	0.02829	52.0	4.1	48.0	95.9
	72	0.02853	49.3	3.7	50.7	96.3
FC NN – Porto	6	0.02835	44.6	3.5	55.4	96.5
	12	**0.02851**	**47.7**	**3.8**	**52.3**	**96.2**
	24	0.02867	45.2	3.7	54.8	96.3
	48	0.02914	39.6	3.1	60.4	96.9
	72	0.02895	38.8	3.2	61.2	96.8
LSTM-based NN – Lisbon	6	0.01364	50.0	3.1	50.0	96.9
	12	0.01443	48.9	3.5	51.1	96.5
	24	**0.01395**	**51.8**	**3.2**	**48.2**	**96.8**
	48	0.01411	46.3	3.3	53.7	96.6
	72	0.01385	51.5	3.5	48.5	96.5
FC NN - Lisbon	6	0.01363	47.8	3.3	52.2	96.7
	12	0.01387	43.8	2.9	56.2	97.1
	24	**0.01358**	**47.4**	**3.1**	**52.6**	**96.9**
	48	0.01376	51.1	3.0	48.9	97.0
	72	0.01419	44.9	3.2	55.1	96.8

As shown in Table 3, LSTM based network achieved higher HR than the fully connect network, for both locations. Another interesting aspect is that the performance depends on the number of past METAR observations (sequence length) which are provided to the NN. The results show that the sequence length that produces best results is an intermediate value. This value was 24-time instants for the LSTM-based NN.

4. Discussion

The low HR of 17,19% at Porto and 5,57% at Lisbon are followed by a MR of 82,8% and 94.43%, respectively. Despite of low FAR (6,34% and 0,84%), the method is strongly penalized by high MR values These results are far from the diagnosis effectiveness of 87% obtained by Menut et al. [11] and Román-Cascón et al. [12] in radiation fog forecasting. The diagnosis method applied to the airport's observational data is constrained by the meteorological variables provided by METAR, where the relative humidity, the wind speed and the 3-hour temperature tendency are the key variables.

The results obtained with the NN-based methods showed that the recurrent network (based on LSTM) achieved a better result than the FC network. This indicates that the recurrent network is more suited to problems with sequential data, where there is a temporal dependence between the data. While the FC network also receives the same amount of data, its performance suffers as more time instants are considered. When 72-time instants are used, the number of weights in the FC network increases significantly and becomes harder to train the NN. For the case of the LSTM, the performance decreases slightly. We believe that this was caused mainly by the difficulty of propagating the error over such a long sequence (during the training stage) and not because of an increase in the number of weights.

One aspect highlighted by the results of the NNs is that there is an optimal value for the sequence length. This value was 24-time instants for the LSTM-based NN. This value corresponds to approximately 12 hours of observations, which is a time interval relevant to the formation of fog.

When considering the HR, FAR, MR and TNR obtained by the neural networks, the results are encouraging, especially if we consider that the FAR is relatively low. Even though the HR is not very high (and consequently the MR is higher than what would be desirable), we must consider that for each fog episode there are hourly estimates for a period of six hours before the episode.

5. Conclusions

The forecast score's method evaluation shows that the performed diagnosis of the pre-fog period, barely capture the pre-fog conditions that favour the formation of fog. The METAR's key variables are unable to explicit the primary mechanisms of the fog onset.

The NN-based methods showed an interesting performance, outperforming the forecast score's method. These results are especially encouraging because very simple networks where considered, with only two layers (FC and LSTM). The results that were obtained also showed that the recurrent networks are preferable, over fully connected, to deal with this type of problem. On the other hand, NN's results also have some limitations. Much more combinations of layers, architectures and parameters should be considered in the future to fully characterize the potential of NN to perform prediction based on METAR data. This characterization should also include an ablation study to determine the sensitivity of the network to different parameters like temperature or humidity.

References

1. Belo-Pereira, M.; Santos, J.A. A persistent wintertime fog episode at Lisbon airport (Portugal): Performance of ECMWF and AROME models. *Meteorol. Appl*. **2016**, 233, 353–370.
2. Izett, J.G.; vanWiel, B.J.H.; Baas, P.; van Hooft, J.A.; Schulte, R.B. Dutch fog: On the observed spatio-temporal variability of fog in the Netherlands. *Q. J. R. Meteorol. Soc*. **2019**, 145, 2817–2834.
3. Stolaki, S.N.; Kazadzis, S.A.; Foris, D.V.; Karacostas, T.S. Fog characteristics at the airport of Thessaloniki, Greece. Nat. Hazards Earth Syst. Sci. **2009**, 9, 1541–1549.
4. Tardif, R.; Rasmussen, R.M. Event-based climatology and typology of fog in the New York City region. *J. Appl. Meteorol. Climatol*. **2007**, 46, 1141–1168.
5. Houssos, E.E.; Lolis, C.J.; Gkikas, A.; Hatzianastassiou, N.; Bartzokas, A. On the atmospheric circulation characteristics associated with fog in Ioannina, north-western Greece. *Int. J. Climatol*. **2012**, 32, 1847–1862.
6. Teixeira, J.; Miranda, P.M.A. Fog prediction at Lisbon airport using a one-dimensional boundary layer model. *Meteorol. Appl*. **2001**, 8, 497–505.
7. Guerreiro, P.M.P; Soares, P.M.M.; Cardoso, R.M.; Ramos A.M. An Analysis of Fog in the Mainland Portuguese International Airports. *Atmosphere* **2020**, Volume 11, Issue 11, 1239.
8. Terradellas, E.; Cano, D. Implementation of single-column model for fog and low cloud forecasting at centralSpanish airports. *Pure Appl. Geophys*. **2007**, 164, 1327–1345.
9. World Meteorological Oranization. WMO Guide to Meteorological Instruments and Methods of Observation; WMO: Geneva, Switzerland, **2017**; Available online: www.wmo.int (accessed on 3 March 2020).
10. Policarpo, C.; Salgado, R.; Costa, M.J. Numerical simulations of fog events in southern Portugal. *Adv. Meteorol*. **2017**.
11. Menut, L.; Mailler, S.; Dupont, J.C., Haeffelin, M., Elias, T. Predictability of the Meteorological Conditions Favourable to Radiative Fog Formation During the 2011 ParisFog Campaign. *Boundary-Layer Meteorol*. **2011**. 150, 277-297.
12. Róman-Cascón, C.; Steeneveld, G. J.; Yagüe, C.; Sastre, M., Arrillaga, J. A.; Maqueda, G. Forecasting radiation fog at climatologically contrating sites: evaluation of statistical methods and WRF. *Q. J. R. Meteorol. Soc*. **2016**. 142, 1048-1063.
13. Gardner, M. W.; Dorling, S. R. Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. *Atmospheric environment*. **1998**. *32*(14-15), 2627-2636.
14. Rodrigues, E. R.; Oliveira, I.; Cunha, R.; Netto, M. DeepDownscale: a deep learning strategy for high-resolution weather forecast. In *2018 IEEE 14th International Conference on e-Science (e-Science)*. IEEE. **2018**; 415-422.
15. Zhao, J.; Deng, F.; Cai, Y.; Chen, J. Long short-term memory-Fully connected (LSTM-FC) neural network for PM2. 5 concentration prediction. *Chemosphere*. **2019**. 220, 486-492.
16. Miao, Q.; Pan, B.; Wang, H.; Hsu, K.; Sorooshian, S. Improving monsoon precipitation prediction using combined convolutional and long short term memory neural network. *Water*. **2019**. 11(5), 977.
17. Wojtkiewicz, J.; Hosseini, M.; Gottumukkala, R.; Chambers, T. L. Hour-ahead solar irradiance forecasting using multivariate gated recurrent units. *Energies*. **2019**. 12(21), 4055.
18. 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. *Advances in neural information processing systems*. **2015**. 28.
19. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural computation*. **1997**. 9 (8):1735–1780.