Fog Climatology & Nowcasting at Dublin Airport

Joseph Oluwasanya and Julita Swiatek

Dublin City University, Collins Ave Ext, Whitehall, Dublin 9
joseph.oluwasanya2@mail.dcu.ie
julita.swiatek2@mail.dcu.ie

Abstract. In this work, we constructed classification models for the purpose of predicting fog at short lead times at Dublin Airport utilizing machine learning (ML) and deep learning (DL) algorithms. We aimed to outperform Harmonie-Arome, the physics-based model currently in use, at fog nowcasting [2]. We conducted an in-depth evaluation of these models considering fog onset and dissipation forecast skill. Additionally we analysed hourly weather observations recorded at Dublin Airport between 2011 and 2021 by Met Éireann [27]. A fog climatology of the Airport region for the 11-year period and a short-term persistence analysis of fog state using Markov Chains is presented. Through combining our learnings from the predictive models and analysis, we aimed to improve the fog nowcasting capabilities at Dublin Airport. The findings demonstrate that our models outperformed Harmonie-Arome in several metrics. We achieved an F1 score and Heidke Skill Score of 0.67, outperforming Harmonie (0.22 and 0.21, respectively), but fog state transitions remained a weak point, with all our models having near-zero Transition HSS.

Keywords: Meteorology · Machine Learning · Time Series

1 Introduction

Airports are forced to implement low-visibility procedures when visibility at an airfield drops below a particular threshold. As per these procedures, airports need to function with a reduced capacity, which means that each departing or landing aircraft will be allotted more time and distance separation from the preceding aircraft compared to regular operating conditions. In practice, this can involve doubling the standard separation distance, which results in substantial capacity constraints and delays. During fog events, Air Traffic Controllers' actions include reducing taxiway occupation, prolonging periods between takeoffs and landings, and even suspending airport operations to avoid runway incursions and other possible incidents. Missed approaches are also much more common during fog events due to insufficient runway visibility. Research has found that the rapid onset of fog incurs substantial financial losses for airports [18, 43].

Therefore, it is crucial that airports have the most up-to-date information on weather conditions so that Air Traffic Controllers can carry out their work

effectively. An accurate fog forecasting model would aid Air Traffic Controllers to determine the current airport capacity, manage the aircraft flow, and ultimately improve safety at the airport.

In this work, we built several ML models for predicting fog at short lead times at Dublin Airport with the aim of outperforming Harmonie-Arome, the model currently in use. Harmonie-Arome is a mesoscale Numerical Weather Prediction (NWP) model, which has been evaluated in many fog prediction studies [33,38], and is currently used by Met Éireann for their weather forecasts [46]. NWP models are used extensively in predicting weather patterns at different time ranges but have been shown to under-perform on hyper-localised fog prediction tasks [32, 44]. We adopted a more data-driven approach; our objective was to build ML models that predict fog more effectively.

In Section 2, we discuss related work in fog forecasting, specifically ML approaches to the problem. The datasets used for analysis and modelling are introduced in Section 3. A background on the data sources used, and explanations of the variables are also given. Section 4 describes the analysis of low-visibility events we conducted. Fog frequency and duration analysis, as well as short-term persistence analysis, are further explained. A description of our modelling setup as well as the particular models that were built is given in Section 5. Finally in Section 6 we present the results of our modelling attempts, benchmarked against the currently in-use model.

2 Related Work

Research into fog nowcasting has historically focused on the use of NWP models, which are mathematical models of the atmosphere and oceans used to create weather forecasts based on current weather conditions. NWP models require initial conditions to be set, which are obtained from weather stations and satellites. The output of an NWP model is deterministic; they produce a single forecast based on a set of initial conditions. Met Eireann, for example, uses the Harmonie-Arome model for weather forecasting [2], which includes fog prediction [46]. NWP weather forecasting is highly sensitive to small changes in initial values of parameters, which are challenging to parameterize in numerical weather models [33]. This makes it especially difficult to accurately forecast fog using such systems [31,38]. Therefore, NWP approaches are primarily used in ensembles of various runs for weather forecasting [30, 45]. In Section 6 of this study, a comparison between the fog predictions generated by our models and those from the Harmonie model in the Dublin airport region is presented.

Fog prediction using ML has been an active research area since the 1980s. Early studies, such as Koziara, Renard and Thompson, used model output statistics (MOS) and multiple linear regression to predict marine fog probability, achieving 78% accuracy [25]. In a more recent study, Bayesian decision networks were used to forecast fog at Melbourne Airport, achieving an AUC score of 0.9 [3]. In 2021, Cornejo-Bueno et al. trained neural networks using the Extreme Learning Machines (ELM) algorithm to predict horizontal visibility within

a half-hour lead time at the A-8 Motor Road in Spain, achieving good results in terms of RMSE [11]. In 2022, various ML algorithms, including Extreme Gradient Boosting (XGB), were employed to predict visibility in Seoul, achieving RMSE of 0.08km for training data and 2.08km for testing data [24]. However, their random sampling method for creating training and testing datasets may overestimate performance (See Section 5.1).

The majority of modern ML approaches to predicting low-visibility events focus on neural networks, due to their ability to effectively model nonlinear relationships between variables. Fabbian et al. and Marzban et al were among the early researchers to adopt this method [15, 26]. Their Artificial Neural Network (ANN) achieved high performance across different lead times. Dutta and Chaudhuri [14] also proposed an ANN approach, achieving the best performance when forecasting very dense visibility within a 50m horizontal distance. Kamangir et al. used a 3-D Convolutional Neural Network (CNN) to predict fog visibility categories, outperforming the High-Resolution Ensemble Forecast in the USA [22].

Despite these studies, no gold-standard approach for fog prediction using ML exists. An analysis of ML methods for fog forecasting in the Mondoñedo region in Spain was recently conducted by Castillo-Boton et al. [6]. They found that Gradient Boosting with Undersampling was the best approach for classification-based fog nowcasting.

In 2022, Vondran et al. discovered issues with current ML approaches to classification-based fog nowcasting systems, particularly in predicting fog state transitions [39]. Accurately predicting transitions is crucial to ensure the smooth functioning of an airport and to enable the airport management to plan and enforce low-visibility procedures, which are necessary to maintain safety during adverse weather conditions. The study highlighted that ML classifiers struggle to predict fog onset and dissipation. They also emphasized the importance of considering temporal context during model training, suggesting time series split cross-validation for a more accurate estimate of model performance in a time series context. This is more reliable than the standard k-fold cross-validation, which may overestimate performance by evaluating the model's imputation skill. Taking this into consideration, we have incorporated these measures into our model evaluation process.

3 Data

We created two separate datasets; the first is used for analysis and modelling, and the second for analysing fog event duration and severity. The data was obtained from several sources, including Met Éireann, and the Global Solar Atlas [1].

Primary dataset: Met Éireann provided us with hourly observations from the Dublin Airport Weather Station for the 11-year period of 2011 to 2021, inclusive. We had 100,000 observations roughly, and 71 variables in the raw data including information about rainfall, wind, relative humidity, atmospheric pressure, temperature, clouds, sunshine, and visibility. The full data dictionary for the Met

Éireann weather dataset can be accessed in this CSV file. Additionally, direct normal irradiation (DNI) data for Dublin Airport was acquired from the Global Solar Atlas, which is a Geographical Information System that provides solar radiation map data [1]. DNI refers to the amount of solar radiation received by a surface perpendicular to the rays coming in a straight line from the direction of the sun measured in kWh/m^2. It does not take into account radiation that is scattered by the atmosphere (*e.g.*, by clouds and aerosols). During data preparation, the initial 72 variables (including DNI) were reduced to 56 variables by removing redundant variables. The full data preparation code can be found in this Jupyter notebook.

Fog event dataset: The Primary dataset captures hourly weather observations; however, it does not provide information about the duration of fog events. We created a separate dataset to measure these. For each observation where visibility was less than 1000m (which is the definition of fog as stated by the WMO [41]) the number of subsequent time steps where the visibility was below the threshold was recorded. The count of low-visibility observations was used to build the fog event dataset. Additionally, we recorded the minimum visibility and season of each event. In Section 4.2, we present our findings from this data, and how it can be used to improve fog event duration predictions at Dublin Airport.

4 Analysis

Fog is defined by the World Meteorological Organization as horizontal visibility of less than 1000m on the ground [41]. There are several types of fog, each differing in the way it forms. In Dublin, the most common types of fog are radiation and advection fog.

Radiation fog occurs on a clear night when the ground loses heat to the atmosphere through radiation. The moisture in the air that comes in contact with the cooling surface also cools, and when the temperature falls below the dew point for that air, fog forms. The fog eventually dissipates when the sun rises and there is an increase in surface temperature. This type of fog is common in autumn in northern Europe [42]. Advection fog, on the other hand, occurs when a warm, moist air mass flows over a colder surface. The air mass cools from below due to the colder surface, and if the temperature of the air mass is reduced to the dew point, fog forms. Advection fog can occur with any level of wind and often occurs in springtime in the coastal areas of northwestern Europe [40].

In the subsequent sections, we present our analysis of low-visibility events at Dublin Airport. In Section 4.1, we explain the relationships between the variables in the data and fog formation. Section 4.2 analyses fog frequency and duration using various methods (code can be found in the "Fog analysis" section of this notebook). Section 4.3 presents our short-term persistence analysis of fog state, which reveals the strong persistence behaviour, and motivates our use of the Naïve persistence operator as a benchmark for our fog nowcast models. The persistence analysis code can be found in this notebook.

4.1 Important Variables

Dry-bulb temperature is measured using a regular thermometer, while wet-bulb temperature is measured using a thermometer with its bulb covered in a wet cloth. The difference between the two temperatures provides an indication of humidity, which is also used to calculate relative humidity (RH). RH is crucial to fog onset because saturated air at 100% RH can't hold more water, causing excess water vapour to condense into water droplets when cooled. This process results in the onset of fog when it occurs at low altitudes (less than 200 meters) [36].

Temperature and relative humidity are used to calculate the dew point temperature; the temperature the air must reach to become saturated. The relationship between dew point, temperature, and relative humidity gives insight into the fog formation trends analysed in Section 4.2. Dew point and temperature have a strong positive linear correlation, revealing that higher temperatures lead to increased moisture absorption. However, the rate of water absorption is slower than the growing capacity for absorption, causing a negative correlation between relative humidity and temperature. Due to this, fog is less likely to occur in the summer. Figure 1 displays a pair plot of these variables, including correlation coefficients and p-values.

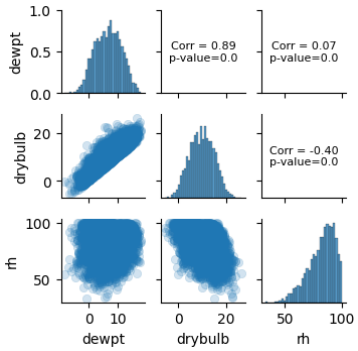

Fig. 1: Pair plot of dew point, temperature, and relative humidity.

Fog formation is dependent on the air temperature being cooled to the dew point. Using this information, we concluded that the closer the air temperature is to the dew point, the higher the likelihood of a low-visibility event. Hence, we created the variable which captures the difference between dry-bulb temperature and dew point. This variable is positively correlated with horizontal visibility ($\rho = 0.566$), showing that generally, the lower this difference is, the lower the horizontal visibility.

Cloudiness and atmospheric pressure both play a role in fog formation. High cloud coverage reduces solar radiation reaching the ground, causing the ground to cool and increasing the likelihood of fog. Fog is more likely to occur in low-pressure environments because high air pressure compresses air near the ground, causing it to warm and reducing relative humidity. In contrast, low pressure

causes air near the ground to expand and cool, increasing the likelihood of fog. The relationships between variables discussed in this section are verified in the Correlation Analysis section of this Jupyter notebook.

4.2 Fog Climatology

Prior to examining fog occurrence, we inspected the visibility data since it determines the fog state. Visibility is the horizontal distance which is clearly viewable at ground level in the airport. This is measured manually by a human, and its raw distribution reveals that although it isn't entirely continuous. When visibility exceeds 20,000m, measurements appear to be rounded to the nearest 5,000m as can be seen in Figure 2. In Plot (b), the visibility values appear to be rounded for visibility exceeding 5,000 metres. As such, we decided to treat fog nowcasting strictly as a classification task.

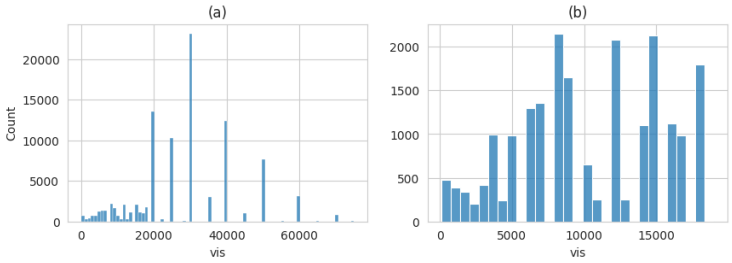

Fig. 2: (a) Full histogram of visibility (b) Only includes observations $< 20,000m$

Next, we analysed fog event frequency and duration using Table 1. We created the table using various aggregations on the fog event dataset. The P_{fog} columns show the probability of fog duration falling within the stated threshold.

Table 1: Fog duration and frequency table.

Month	**Fog Events**	**Max Duration**	P_{fog}			
			$> 16h$	$> 12h$	$< 3h$	$< 6h$
1	18	9	0.0	0.0	22.2	72.2
2	20	15	0.0	5.0	45.0	75.0
3	37	10	0.0	0.0	40.5	70.3
4	16	12	0.0	6.2	37.5	75.0
5	12	8	0.0	0.0	66.7	91.7
6	16	15	0.0	6.2	62.5	81.2
7	9	5	0.0	0.0	55.6	100.0
8	26	14	0.0	7.7	42.3	80.8
9	28	11	0.0	0.0	35.7	85.7
10	21	8	0.0	0.0	33.3	71.4
11	26	19	3.8	7.7	38.5	84.6
12	25	19	4.0	4.0	68.0	92.0

Fog events occur year-round at Dublin Airport, with fewer events in the summer and the highest number in autumn and winter. The longest fog events typically happen in November and December, with the shortest in the summer

months. When considering fog duration cutoffs, only November and December had events lasting 16 or more hours, while four additional months had events exceeding 12 hours. The summer months had the shortest fog events, with all July events lasting less than 6 hours. Finally, we examine the statistical properties of fog duration at Dublin Airport over the entire period.

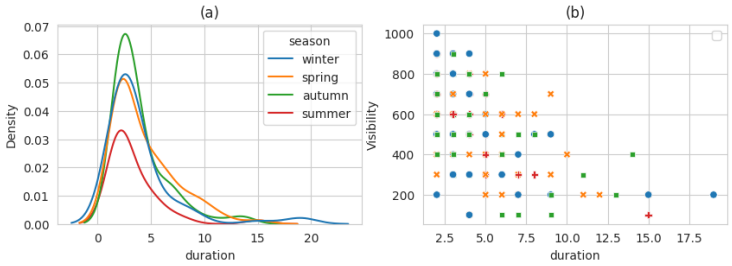

Fig. 3: (a) KDE plot of fog duration, separated by Season (b) Scatter plot of event duration vs minimum visibility in event

In Figure 3(a), we observe a right-skewed, unimodal distribution of fog duration for each season, with slight variations. Outliers on the right tail are visible for the winter season due to extreme fog events. The spring season displays the thickest tail, as reflected in its highest average fog duration (Table 2). Conversely, the summer season, with the lowest frequency of fog events, exhibits the smallest density and distribution. Moreover, fog events in summer clear up faster than in other seasons. Figure 3(b) demonstrates a negative correlation between fog duration and minimum horizontal visibility during the event. This implies that lower minimum visibility predicts longer fog duration, but it cannot be included in our models due to data leakage. Table 2 lists the average fog durations and interquartile ranges (IQRs) for each season. Spring had the longest mean duration at 4.34 hours, while summer had the shortest at 3.35 hours.

Table 2: Fog event duration Statistics by Season

	Spring	**Summer**	**Autumn**	**Winter**
Mean	4.34	3.35	3.88	3.91
IQR	4.0	2.0	2.0	2.0

The IQR represents the middle 50% of the data and indicates the range of durations that are typical for each season. The IQR for spring is the largest at 4.0 hours, indicating that there is a wider range of fog event durations in this season. In contrast, the IQR for summer is the smallest at 2.0 hours, suggesting that fog events in summer are relatively consistent in duration. Overall, the figures suggest that while there are some differences in the mean duration of fog events between seasons, the range of durations within each season varies as well. This can be used to better inform estimates of fog event duration at the airport. For

instance, if we use Bayesian methods to estimate fog duration, then these values help us construct an informative prior [3].

4.3 Persistence Analysis

We conducted a Short-Term Persistence Analysis on low-visibility events using Markov Chain Models (MCMs). To model fog state, we treated it as a discrete binary variable, where 0 corresponds to no fog ($visibility \geq 1000m$) and 1 corresponds to fog. This allowed us to observe how often states remained the same hour-by-hour and to evaluate the Naïve persistence operator as a predictor of fog, which involves simply predicting the fog state in the next hour to be the same as the current state. Markov chains were employed because they are effective for short-term prediction of meteorological time series [7, 12]. Initially, we modelled fog state as a first-order Markov Chain, which assumed that fog occurrence in the present hour was dependent only on the previous hour's fog state. The corresponding transition probabilities of the Markov Chain Model are as follows:

$$p_{i \to j} = P(X_t = j | X_{t-1} = i) \tag{1}$$

Where i and j take values 0 or 1. The transition probability matrix for a first-order Markov Chain can therefore be described as such:

$$\begin{pmatrix} p_{0\to0} & p_{0\to1} \\ p_{1\to0} & p_{1\to1} \end{pmatrix} \tag{2}$$

We estimate state persistence using the main diagonal of transition probability matrices. For a first-order Markov Chain, this is calculated as $\mathfrak{p} = \frac{1}{2}(p_{0\to0} + p_{1\to1})$. However, for higher-order Markov Chains, the memoryless property of the first-order model no longer holds, as transition probabilities depend on previous states. For instance, for a second-order Markov Chain, the state at time t depends on states at times $t-2$ and $t-1$, while for a k-order model, states at times $t-k, t-k+1, \ldots, t-1$ are considered. Tables 3a, 3b, 4a, and 4b present the results of the Short-Term Persistence Analysis. They display transition probability matrices for 1st to 4th-order Markov Chain Models, along with the corresponding state persistence estimates.

Our analysis indicates strong short-term persistence behaviour of fog state. From the first-order MCM, predictions are already equivalent to a Naïve persistence operator. The persistence becomes stronger with higher-order Markov Chain Models. When looking at the first-order Markov Chain Model, the probability of state persistence, given there was fog in the last hour is $p_{1\to1} = \mathbf{0.66}$. However, this value goes up to $p_{1111\to1} = \mathbf{0.73}$ for a fourth-order Markov Chain Model when considering the probability of state persistence, given there was fog for the last four hours prior. The persistence values $\mathfrak{p}$ show the same pattern. The short-term persistence behaviour, coupled with the rarity of fog events, suggests that the Naïve persistence model strongly predicts fog state. Section 6 compares the Naïve persistence operator to various ML and DL models.

Table 3: Transition probabilities for the first-order (3a) and second-order (3b) Markov Chain Models

	0	**1**
0	**0.997**	0.003
1	0.34	**0.66**

(a) $\mathfrak{p} = \frac{1}{2}(p_{0\to0} + p_{1\to1}) =$ **0.829**

	0	**1**
0 0	**0.998**	0.002
0 1	0.44	0.56
1 0	0.91	0.09
1 1	0.29	**0.71**

(b) $\mathfrak{p} = \frac{1}{2}(p_{00\to1} + p_{11\to1}) =$ **0.855**

Table 4: Transition probabilities for the third-order (4a) and fourth-order (4b) Markov Chain Models

	0	**1**
0 0 0	**0.998**	0.002
⋮	⋮	⋮
1 1 1	0.273	**0.727**

(a) $\mathfrak{p} = \frac{1}{2}(p_{000\to1} + p_{111\to1}) =$ **0.862**

	0	**1**
0 0 0 0	**0.998**	0.002
⋮	⋮	⋮
1 1 1 1	0.262	**0.73**

(b) $\mathfrak{p} = \frac{1}{2}(p_{0000\to1} + p_{1111\to1}) =$ **0.868**

5 Modelling

The implementation code for all of our models can be found in Jupyter notebooks, hosted on Google Colab in this folder. We implemented the ML models, as well as the feedforward neural network (ANN) using sci-kit learn. The remaining DL models were implemented in Keras and TensorFlow.

5.1 Setup

The data split used for training and testing is as follows: observations from 2011 to 2019 were used for training/validation, while 2020 and 2021 were used for out-of-time testing. We used Time series split 5-Fold cross-validation on the train/validation set to estimate model performance and make modelling decisions. This was done to robustly evaluate forecast skill, which is reflected in the time series split, in contrast to standard k-fold cross-validation, which evaluates the model's imputation skill and can lead to overestimation of performance [39].

Hyperparameters were tuned using GridSearchCV, RandomizedSearchCV, or Bayesian optimization, with GridSearchCV being the preferred method due to its exhaustive search of the defined hyperparameter space. In some cases, we used RandomizedSearchCV due to computation constraints. Bayesian optimization was only used for the ANN but was left out for subsequent models due to run time constraints. Hyperparameters were selected based on the lowest average binary cross-entropy loss. In the case of DL models, different parameters were tested

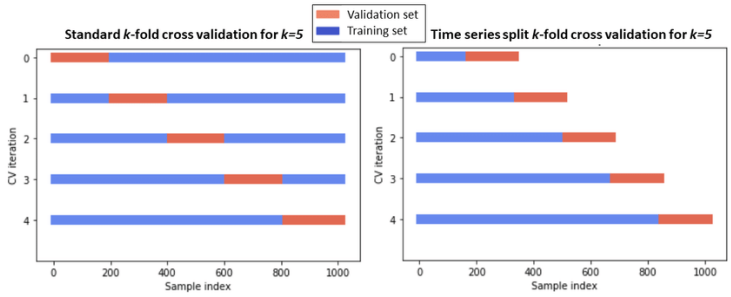

Fig. 4: Standard k-fold cross validation versus Time Series split k-fold cross validation.

manually for each of the algorithms, and their effects on training and validation loss curves were observed. Afterwards, GridSearchCV was used to find the best configuration of hyperparameters. The final hyperparameter configurations for each model can be found in Table 8 in the appendix. To streamline our data preprocessing and model evaluation workflow, we created an ML Python module built off scikit-learn, which can be accessed here. We used Neptune AI for keeping track of model metadata, including performance metrics [29].

Feature selection was carried out using a combination of domain knowledge from relevant research [11, 12], and discussions with experts in the field, including meteorologists at Met Éireann. To further validate our selection, we also employed the XGBoost gain importance metric, which assesses the relative importance of each variable based on how often it occurs in the ensemble of decision trees [9]. The scores were calculated using time series cross-validation, with a mean calculated across the 5 folds. Out of the 56 variables in the data, we trained our models using 16 variables. See Table 7 in the appendix for the list of variables.

Fog events are rare and make up only 0.8% of hourly observations in our 11-year dataset. This presents a challenge for ML models that tend to favor the majority class to minimize classification error [16]. To address this issue, we applied various data sampling techniques, such as random undersampling of the majority class and oversampling of the minority class using SMOTE [8]. The models didn't improve when we used such methods, so they aren't included in Section 6 but their implementation and results can be found here.

5.2 Algorithms

Below are brief descriptions of the algorithms used to build fog nowcasting models.

Random Forest: Ensemble of decision trees. While creating each tree, bagging and feature randomization are employed to produce an uncorrelated forest of trees whose prediction by the majority is more accurate compared to each

tree [5].
XGBoost: Ensemble of decision trees constructed using gradient boosting. XGBoost iteratively improves the model by adding decision trees to predict error of the current ensemble. This process continues until a stopping criterion is reached, such as a maximum number of trees or a minimum improvement in the loss function [10].
Bagged Ensemble: Ensembled XGBoost models using bagging (bootstrap aggregating). Models are trained on bootstrap samples and averaged to aggregate their results [4].
Deep Artificial Neural Network (ANN): A fully connected multi-layer feed-forward neural network with multiple hidden layers. Each layer applies a nonlinear activation function to a linear combination of inputs and is trained through backpropagation, an optimization algorithm that adjusts the weights and biases of the neurons to minimize a specified loss function. [19].
Recurrent Neural Network (RNN): A type of ANN where connections between nodes create a cycle, allowing output from a hidden layer to be the input to the same layer. It is useful for time series forecasting as information from prior inputs influences the current input and output, thus making the RNN's output dependent on previous elements of a sequence. RNNs frequently suffer from the vanishing and exploding gradient problems, wherein the gradient (the slope of the loss function along the error curve) becomes too small or too big, respectively, leading to the weight parameters becoming insignificant (vanishing gradient) or too large (exploding gradient) [20].
Long Short-Term Memory Network (LSTM): A modification of RNN which addresses the problem of the vanishing gradient through a memory mechanism. The hidden layers of LSTMs have cells which contain input, forget, and output gates. These gates control the flow of information which is needed to predict the output in the network. This gives LSTMs the ability to learn long-term dependencies from sequential input [17].
Gated Recurrent Unit (GRU): A simplified version of LSTM, also aimed at solving the problem of long-term dependencies. It uses only two gates; update and reset, allowing the model to keep relevant information and pass it on to the next time step. It is easier to train than LSTM as it has fewer trainable parameters [13].

5.3 Evaluation

F1 and Heidke Skill Score (HSS) were used to evaluate models; F1 scores models based on Precision and Recall [37], while HSS is a categorical forecast metric comparing model skill to random forecasting based on category frequency. F1 takes values in the range $[0, 1]$, and is a robust metric for imbalanced binary classification problems. HSS is a commonly used metric for evaluating categorical forecasting systems that ranges from -1 to 1, where 0 -1 implies the model outperforms random, and -1 to 0 implies it performs worse than random. Near zero values imply forecasts are roughly equivalent to random. Equations for these

metrics are given below:

$$\text{F1} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad \text{HSS} = \frac{TP + TN - \text{Hits by chance}}{n - \text{Hits by chance}} \tag{3}$$

Several intermediate values are used to obtain the above metrics. These are Precision $= \frac{TP}{TP+FP}$, Recall $= \frac{TP}{TP+FN}$, and Hits by chance $= \frac{1}{n}((TP + FP) \times (TP + TN) + (TN + FP) \times (TN + FN))$ where TP, FP, TN, FN are True Positive, False Positive, True Negative, and False Negative, respectively and n is the total number of validation examples. We evaluate the models' ability to predict fog onset and dissipation by filtering the validation set to only the fog state transition time points and using the same metrics as described earlier. This approach helps us determine how well the models perform in predicting fog state change. Figure 5 illustrates how the transition time points are selected.

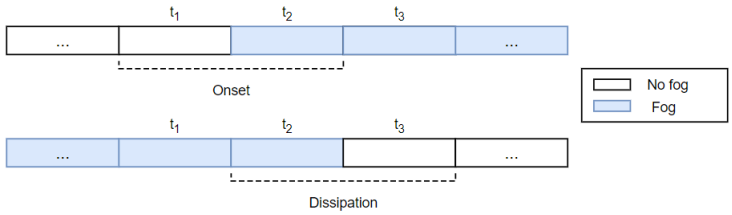

Fig. 5: Fog onset and dissipation time points

It is important to note that the LSTM and RNN DL models exhibited stochastic behaviour each time they were retrained due to the use of a GPU, which introduces randomness even when seeded. Due to this, we ran each of the models multiple times and computed sample statistics of the results.

6 Results

In Table 5, we present the results of our models. Overall, our best model is the Bagged XGBoost Ensemble, which achieves the highest F1 and HSS scores. The ML approaches outperformed the DL models at predicting hourly fog states, but we note that GRU achieves the highest Transition F1 score. As previously suggested in Section 4.3, the persistence operator is a strong predictor of future fog states. The results in Table 6 show that persistence outperformed most of our models at predicting hourly fog states. In terms of fog onset and dissipation forecasts, Harmonie-Arome scores the highest. Although, note that its Transition HSS score is zero and that all of our models have near-zero Transition HSS scores, implying that none of the models effectively predict fog onset and dissipation. The transition metrics are left blank for the persistence model as it is unable to predict fog transitions since it forecasts fog state in the next hour to be the same as the current fog state.

Table 5: Results of model runs. Best values in bold. HSS stands for Heidke Skill Score

Model	F1 Score	HSS	Transition F1 Score	Transition HSS
Random Forest	0.6	0.59	0.27	-0.02
ANN	0.56	0.55	0.3	-0.05
Recurrent Neural Network	0.33	0.33	0.25	**0.003**
Gated Recurrent Unit	0.5	0.5	**0.34**	0.0
Long Short-Term Memory Network	0.39	0.39	0.29	-0.002
XGBoost	**0.67**	0.66	0.32	-0.1
Bagged XGBoost Ensemble	**0.67**	**0.67**	0.3	-0.1

Table 6: Best model compared to Benchmarks

Model	F1 Score	HSS	Transition F1 Score	Transition HSS
Bagged XGBoost Ensemble	**0.67**	**0.67**	0.3	-0.1
Harmonie-Arome	0.22	0.21	**0.37**	**0.0**
Persistence	0.64	0.65	-	-

7 Conclusion & Future Work

In our analysis, we investigated monthly fog event frequency and duration at Dublin Airport for the 11-year period from 2011 to 2021. Based on our empirical analysis, we identified the temporal patterns of low-visibility events with respect to their occurrence and duration throughout the year at Dublin Airport. We analysed the relationships among various weather variables such as air temperature, dew point temperature, and air pressure that influence fog formation in Section 4.3. This gave us a basis for selecting variables for modelling. The short-term persistence analysis shows the strong persistence behaviour of the fog state at the airport, implying that the Naïve persistence operator is a strong predictor of the fog state. The persistence performance was difficult to surpass when considering its ability to predict hourly fog states, as is seen in Section 6.

Previous research into classification-based machine learning models for fog nowcasting does not take into account fog onset and dissipation forecast skill [6, 28]. Our evaluation framework improves on this by separately scoring fog state transitions. We also leverage time series split cross-validation for making modelling decisions. Our models outperformed Harmonie-Arome at predicting hourly fog states, with Harmonie scoring 0.22 in F1 and 0.21 in HSS, and our best model scoring 0.67 in both metrics. When the ability to predict fog onset and dissipation was taken into account, we found that, in general, the models were unable to predict these state transitions, as they all scored near-zero transition HSS. This limits our categorical forecasting models in the operational context at the airport, as correctly predicting fog onset is the most valuable property of such a system. Air Traffic Controllers could use this information to plan for low-visibility procedures in advance. The results align with recent research [39], and show that current classification-based ML models struggle with predicting fog onset and dissipation. Below, we outline our ideas for further research to address this problem.

We attempted to address the class imbalance issue in our data by implementing random undersampling and SMOTE techniques, but there are various other methods available that we did not utilize. One such technique involves adjusting the weight of each example in the training process, which can prioritize the minority class while still preserving the time-series context of the data. In our study, we focused on predicting fog with a 1-hour lead time. Exploring longer lead times, such as 6 or 12 hours, could provide more insight into the models' ability to forecast state transitions. For example, we could evaluate the frequency of fog onset within 12 hours when the model predicts fog onset 12 hours from now. Finally, given the non-continuous distribution of visibility values in our data, we decided to treat this as a classification task exclusively. A potentially more effective approach could be to model horizontal visibility directly as a regression task. Runway Visual Range (RVR) data, measured by sensors at airports, can provide more precise and objective data for predicting visibility directly [12].

In conclusion, our models outperformed Harmonie-Arome at predicting hourly fog states, although they were unable to predict fog onset and dissipation, which limits the applicability of such models in an operational context. We suggest including the evaluation of fog onset and dissipation when building classification-based machine learning models for fog nowcasting.

Acknowledgements

Thanks to Prof. Tomás Ward for providing guidance throughout the project. His feedback, encouragement and support were invaluable in completing this work. Thanks to Met Eireann for providing the data for this project. Thanks to Dr Kevin Devine of Met Eireann for his advice throughout the project [23]. Thanks to Prof Sancho Salcedo-Sanz, Pérez Aracil Jorge, and Peláez Rodríguez César of La Universidad de Alcalá [21,34,35] for sharing their knowledge on this subject with us, and for providing implementation code for various DL Models.

8 Appendix

Table 7: Features selected

Variable	**Units**
Visibility	meters
Drybulb temp - Dewpoint temp	°C
Dewpoint temp	°C
Drybulb temp	°C
Direct Normal Irradiation	kWh/m^2
Relative Humidity	%
Cloud ceiling height	100 ft
Low cloud height	meters
Amount of cloud	Okta
Height of Second Significant Cloud layer	meters
Pressure Characteristic	hPa
Air Pressure	hPa
Vapor pressure	hPa
Wind speed	knots
Wind direction	degrees
Rainfall duration last hr	minutes

Table 8: Hyperparameters used for each ML model (and ANN). Where applicable, the random state was set to 42.

Model	**Hyperparameter**	**Value**
Random Forest	n_estimators	800
	max_depth	80
	max_features	'sqrt'
	min_samples_split	12
	min_samples_leaf	4
	bootstrap	False
XGBoost	n_estimators	415
	learning_rate	0.007
	max_depth	2
	subsample	0.825
	colsample_bytree	0.904
	gamma	0.225
Bagged XGBoost	n_estimators	100
ANN	max_iter	500
	batch_size	100
	activation	'logistic'
	hidden_layer_sizes	(762, 762)
	early_stopping	True
	n_iter_no_change	100
	solver	'adam'
	alpha	1e-06

Table 9: Hyperparameters used for each DNN. Where applicable, the random state was set to 42. Early stopping was used for each model.

Model	**Hyperparameter**	**Value**
RNN	Structure	128 neuron simple RNN layer $->$ 64 unit LSTM layer $->$ 2x32 neuron dense layers $->$ output neuron (sigmoid)
	batch_size	100
	epochs	50
	learning_rate	0.001
	decay	0.00003
	activation	'relu'
	loss	'binary_crossentropy'
	sequence length	5
LSTM	Structure	256 then 128 unit LSTM layers $->$ 32 then 16 neuron dense layers $->$ output neuron (sigmoid)
	batch size	100
	epochs	30
	learning_rate	0.001
	activation	'relu'
	patience	30
	loss	'binary_crossentropy'
	sequence length	5
GRU	Structure	256 then 128 unit GRU layers $->$ 2x32 neuron dense layers $->$ output neuron (sigmoid)
	batch size	100
	epochs	50
	learning_rate	0.001
	activation	'relu'
	patience	30
	loss	'binary_crossentropy'
	sequence length	5

References

1. Global Solar Atlas: Global Solar Atlas, https://globalsolaratlas.info/map?c=53.66946,-6.177063,10&s=53.352191,-6.2677&m=site
2. Bengtsson, L., Andrae, U., Aspelien, T., Batrak, Y., Calvo, J., Rooy, W.d., Gleeson, E., Hansen-Sass, B., Homleid, M., Hortal, M., Ivarsson, K.I., Lenderink, G., Niemelä, S., Nielsen, K.P., Onvlee, J., Rontu, L., Samuelsson, P., Muñoz, D.S., Subias, A., Tijm, S., Toll, V., Yang, X., Køltzow, M.: The HARMONIE–AROME Model Configuration in the ALADIN–HIRLAM NWP System. Monthly Weather Review **145**(5), 1919–1935 (May 2017). https://doi.org/10.1175/MWR-D-16-0417.1
3. Boneh, T., Weymouth, G.T., Newham, P., Potts, R., Bally, J., Nicholson, A.E., Korb, K.B.: Fog Forecasting for Melbourne Airport Using a Bayesian Decision Network. Weather and Forecasting **30**(5), 1218–1233 (Oct 2015). https://doi.org/10.1175/WAF-D-15-0005.1
4. Breiman, L.: Bagging predictors. Machine learning **24**, 123–140 (1996)
5. Breiman, L.: Random Forests. Machine Learning **45**(1), 5–32 (Oct 2001). https://doi.org/10.1023/A:1010933404324
6. Castillo-Botón, C., Casillas-Pérez, D., Casanova-Mateo, C., Ghimire, S., Cerro-Prada, E., Gutierrez, P.A., Deo, R.C., Salcedo-Sanz, S.: Machine learning regression and classification methods for fog events prediction. Atmospheric Research **272**, 106157 (Jul 2022). https://doi.org/10.1016/j.atmosres.2022.106157
7. Cazacioc, L., Cipu, E.C.: Evaluation of the transition probabilities for daily precipitation time series using a Markov chain model. Proceedings of 3rd International Colloquium–Mathematics in Engineering and Numerical Physics pp. 82–92 (2005), https://www.academia.edu/48212308/Evaluation_of_the_transition_probabilities_for_daily_precipitation_time_series_using_a_Markov_chain_model
8. Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P.: SMOTE: Synthetic minority over-sampling technique. Journal of Artificial Intelligence Research **16**, 321–357 (jun 2002). https://doi.org/10.1613/jair.953
9. Chen, T., Guestrin, C.: XGBoost. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM (aug 2016). https://doi.org/10.1145/2939672.2939785
10. Chen, T., Guestrin, C.: Xgboost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. p. 785–794. KDD '16, Association for Computing Machinery, New York, NY, USA (2016). https://doi.org/10.1145/2939672.2939785
11. Cornejo-Bueno, S., Casillas-Pérez, D., Cornejo-Bueno, L., Chidean, M.I., Caamaño, A.J., Cerro-Prada, E., Casanova-Mateo, C., Salcedo-Sanz, S.: Statistical analysis and machine learning prediction of fog-caused low-visibility events at a-8 motor-road in spain. Atmosphere **12**(6), 679 (May 2021). https://doi.org/10.3390/atmos12060679
12. Cornejo-Bueno, S., Casillas-Pérez, D., Cornejo-Bueno, L., Chidean, M.I., Caamaño, A.J., Sanz-Justo, J., Casanova-Mateo, C., Salcedo-Sanz, S.: Persistence Analysis and Prediction of Low-Visibility Events at Valladolid Airport, Spain. Symmetry **12**(6), 1045 (Jun 2020). https://doi.org/10.3390/sym12061045
13. Dey, R., Salem, F.M.: Gate-variants of gated recurrent unit (gru) neural networks. In: 2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS). pp. 1597–1600 (2017). https://doi.org/10.1109/MWSCAS.2017.8053243

14. Dutta, D., Chaudhuri, S.: Nowcasting visibility during wintertime fog over the airport of a metropolis of India: decision tree algorithm and artificial neural network approach. Natural Hazards **75**(2), 1349–1368 (Jan 2015). https://doi.org/10.1007/s11069-014-1388-9
15. Fabbian, D., Dear, R.d., Lellyett, S.: Application of Artificial Neural Network Forecasts to Predict Fog at Canberra International Airport. Weather and Forecasting **22**(2), 372–381 (Apr 2007). https://doi.org/10.1175/WAF980.1
16. He, H., Garcia, E.A.: Learning from Imbalanced Data. IEEE Transactions on Knowledge and Data Engineering **21**(9), 1263–1284 (Sep 2009). https://doi.org/10.1109/TKDE.2008.239
17. Hochreiter, S., Schmidhuber, J.: Long Short-Term Memory. Neural Computation **9**(8), 1735–1780 (Nov 1997). https://doi.org/10.1162/neco.1997.9.8.1735
18. Huang, H., Zhang, G.: Case Studies of Low-Visibility Forecasting in Falling Snow With WRF Model. Journal of Geophysical Research: Atmospheres **122**(23), 12,862–12,874 (2017). https://doi.org/10.1002/2017JD026459
19. Jain, A., Mao, J., Mohiuddin, K.: Artificial neural networks: a tutorial. Computer **29**(3), 31–44 (1996). https://doi.org/10.1109/2.485891
20. Jain, L.C., Medsker, L.R.: Recurrent Neural Networks: Design and Applications. CRC Press, Inc., USA, 1st edn. (1999)
21. Jorge PÉREZ-ARACIL | Professor (Assistant) | Universidad Politécnica de Madrid, Madrid | UPM | Research profile, https://www.researchgate.net/profile/Jorge-Perez-Aracil
22. Kamangir, H., Collins, W., Tissot, P., King, S.A., Dinh, H.T.H., Durham, N., Rizzo, J.: FogNet: A multiscale 3D CNN with double-branch dense block and attention mechanism for fog prediction. Machine Learning with Applications **5**, 100038 (Sep 2021). https://doi.org/10.1016/j.mlwa.2021.100038
23. Kevin Devine, https://kmdevine.com/
24. Kim, B.Y., Cha, J.W., Chang, K.H., Lee, C.: Estimation of the Visibility in Seoul, South Korea, Based on Particulate Matter and Weather Data, Using Machine-learning Algorithm. Aerosol and Air Quality Research **22**(10), 220125 (2022). https://doi.org/10.4209/aaqr.220125
25. Koziara, M.C., Renard, R.J., Thompson, W.J.: Estimating Marine Fog Probability Using a Model Output Statistics Scheme. Monthly Weather Review **111**(12), 2333–2340 (Dec 1983). https://doi.org/10.1175/1520-0493(1983)111<2333:EMFPUA>2.0.CO;2
26. Marzban, C., Leyton, S., Colman, B.: Ceiling and Visibility Forecasts via Neural Networks. Weather and Forecasting **22**(3), 466–479 (Jun 2007). https://doi.org/10.1175/WAF994.1
27. Met Éireann, https://www.met.ie
28. Miao, K.c., Han, T.t., Yao, Y.q., Lu, H., Chen, P., Wang, B., Zhang, J.: Application of LSTM for short term fog forecasting based on meteorological elements. Neurocomputing **408**, 285–291 (Sep 2020). https://doi.org/10.1016/j.neucom.2019.12.129
29. neptune.ai | ML metadata store, https://neptune.ai/
30. Roberts, B., Gallo, B.T., Jirak, I.L., Clark, A.J.: The High Resolution Ensemble Forecast (HREF) system: Applications and Performance for Forecasting Convective Storms (Dec 2019). https://doi.org/10.1002/essoar.10501462.1
31. Román-Cascón, C., Steeneveld, G.J., Yagüe, C., Sastre, M., Arrillaga, J.A., Maqueda, G.: Forecasting radiation fog at climatologically contrasting sites: evaluation of statistical methods and WRF. Quarterly Journal of the Royal Meteorological Society **142**(695), 1048–1063 (2016). https://doi.org/10.1002/qj.2708

32. Román-Cascón, C., Yagüe, C., Sastre, M., Maqueda, G., Salamanca, F., Viana, S.: Observations and WRF simulations of fog events at the Spanish Northern Plateau. In: Advances in Science and Research. vol. 8, pp. 11–18. Copernicus GmbH (Feb 2012). https://doi.org/10.5194/asr-8-11-2012
33. Román-Cascón, C., Yagüe, C., Steeneveld, G.J., Morales, G., Arrillaga, J.A., Sastre, M., Maqueda, G.: Radiation and cloud-base lowering fog events: Observational analysis and evaluation of WRF and HARMONIE. Atmospheric Research **229**, 190–207 (Nov 2019). https://doi.org/10.1016/j.atmosres.2019.06.018
34. Sancho Salcedo-Sanz Home Page, http://agamenon.tsc.uah.es/Personales/sancho/
35. Signal, C.P.R.o.A.|.U.D.o., Student, C.T.o.E.: César PELÁEZ-RODRÍGUEZ | PhD Student | Master of Engineering | University of Alcalá, Alcalá de Henares | UAH | Department of Signal and Communications Theory | Research profile, https://www.researchgate.net/profile/Cesar-Pelaez-Rodriguez
36. American Meteorological Society: Fog - Glossary of Meteorology, https://glossary.ametsoc.org/wiki/Fog
37. Sokolova, M., Japkowicz, N., Szpakowicz, S.: Beyond accuracy, f-score and roc: A family of discriminant measures for performance evaluation. vol. Vol. 4304, pp. 1015–1021 (01 2006). https://doi.org/10.1007/11941439_114
38. Steeneveld, G.J., Ronda, R.J., Holtslag, A.A.M.: The Challenge of Forecasting the Onset and Development of Radiation Fog Using Mesoscale Atmospheric Models. Boundary-Layer Meteorology **154**(2), 265–289 (Feb 2015). https://doi.org/10.1007/s10546-014-9973-8
39. Vorndran, M., Schütz, A., Bendix, J., Thies, B.: Current Training and Validation Weaknesses in Classification-Based Radiation Fog Nowcast Using Machine Learning Algorithms. Artificial Intelligence for the Earth Systems **1**(2) (Apr 2022). https://doi.org/10.1175/AIES-D-21-0006.1
40. WMO: Advection fog, https://cloudatlas.wmo.int/advection-fog.html
41. WMO: Fog compared with Mist, https://cloudatlas.wmo.int/fog-compared-with-mist.html
42. WMO: Radiation fog, https://cloudatlas.wmo.int/radiation-fog.html
43. Wu, J., Fu, C., Zhang, L., Tang, J.: Trends of visibility on sunny days in China in the recent 50 years. Atmospheric Environment **55**, 339–346 (Aug 2012). https://doi.org/10.1016/j.atmosenv.2012.03.037
44. Zhou, B., Du, J., Gultepe, I., Dimego, G.: Forecast of Low Visibility and Fog from NCEP: Current Status and Efforts. Pure and Applied Geophysics **169**(5), 895–909 (May 2012). https://doi.org/10.1007/s00024-011-0327-x
45. Met Éireann: Launch of Ireland's high resolution ensemble-based forecasting system - Met Éireann - The Irish Meteorological Service, https://www.met.ie/irelands-first-high-resolution-ensemble-based-forecasting-system
46. Met Éireann: NWP in Met Éireann (Jul 2011), https://www.met.ie/science/numerical-weather-prediction/operational-nwp-in-met-eireann