Atmospheric Research 214 (2018) 64–73

Contents lists available at ScienceDirect

Atmospheric Research

journal homepage: www.elsevier.com/locate/atmosres

Prediction of low-visibility events due to fog using ordinal classification

D. Guijo-Rubio[a,*], P.A. Gutiérrez[a], C. Casanova-Mateo[b,c], J. Sanz-Justo[b], S. Salcedo-Sanz[d], C. Hervás-Martínez[a]

[a] Department of Computer Science and Numerical Analysis, Universidad de Córdoba, Córdoba, Spain
[b] LATUV: Remote Sensing Laboratory, Universidad de Valladolid, Valladolid, Spain
[c] Department of Civil Engineering: Construction, Infrastructure and Transport, Universidad Politécnica de Madrid, Madrid, Spain
[d] Department of Signal Processing and Communications, Universidad de Alcalá, Alcalá de Henares, Madrid, Spain

ARTICLE INFO

Keyword:
Airports
Fog events prediction
Time series
Forecasting
Ordinal classification
Time series preprocessing.

ABSTRACT

The prediction of low-visibility events is very important in many human activities, and crucial in transportation facilities such as airports, where they can cause severe impact in flight scheduling and safety. The design of accurate predictors for low-visibility events can be approached by modelling future visibility conditions based on past values of different input variables, recorded at the airport. The use of autoregressive time series forecasters involves adjusting the order of the model (number of past series values or size of the sliding window), which usually depends on the dynamical nature of the time series. Moreover, the same window size is normally used for all the data, thought it would be reasonable to use different sliding windows. In this paper, we propose a hybrid prediction model for daily low-visibility events, which combines fixed-size and dynamic windows, and adapts its size according to the dynamics of the time series. Moreover, visibility is labelled using three ordered categories (FOG, MIST and CLEAR), and the prediction is then carried out by means of ordinal classifiers, in order to take advantage of the ordinal nature of low-visibility events. We evaluate the model using a dataset from Valladolid airport (Spain), where radiation fog is very common in autumn and winter months. The considered data set includes five different meteorological input variables (wind speed and direction, temperature, relative humidity and QNH – pressure adjusted at mean sea level) and the Runway Visual Range (RVR), which is used to characterize the low-visibility events at the airport. The results show that the proposed hybrid window model with ordinal classification leads to very robust performance prediction in daily time-horizon, improving the results obtained by the persistence model and alternative prediction schemes tested.

1. Introduction

Fog is a meteorological phenomenon consisting of the suspension of very small, usually microscopic water droplets in the air, generally reducing the horizontal visibility at the Earth's surface to < 1 km (WMO 2011). Basically, fog is a cloud at ground level, that has been studied extensively from different points of view (Román Cascón 2015; Román-Cascón et al. 2012). Different fog types can be classified according to the two main physical processes which produce saturation of the air: cooling and the addition of water vapour. In the first group we have the radiation fog, that can occur in long nights and clear skies as a result of the thermal radiation cooling, the advection fog, that can arise when warm and moist air is moving over a colder surface, and the up-slope or orographic fog, formed when the air is forced up a slope undergoing an adiabatic cooling process. The second group includes in turn two main types: the steam fog, produced by rapid evaporation from an underlying warm water surface, and the frontal fog, caused by rain falling into cold air and moistening it. In all the cases, when the horizontal visibility is at least 1 km, but not > 5 km, the phenomenon is called *mist* (WMO 2011).

As it is well known, adverse weather phenomena can strongly affect air traffic management and flight operations (da Rocha et al. 2015; Dey 2018). Thunderstorms, icing, turbulence or wind shear conditions can greatly disrupt air traffic flows, leading to flight delays, diversions or cancellations. However, fog is perhaps the most important local visibility-reducing phenomenon that affects airport operations, since it can strongly reduce runway capacity (Bergot et al. 2007).

The number of runways fully available is a key element in the global airspace capacity, since they largely determine the number of departing or arriving flights within an airspace area. Therefore, any local weather

* Corresponding author at: Department of Computer Science and Numerical Analysis, University of Córdoba, Rabanales Campus, Albert Einstein Building 3rd Floor, 14071 Córdoba, Spain.
E-mail address: dguijo@uco.es (D. Guijo-Rubio).

https://doi.org/10.1016/j.atmosres.2018.07.017
Received 12 April 2018; Received in revised form 13 July 2018; Accepted 16 July 2018
Available online 20 July 2018

phenomena that can significantly reduce runway operations will have a deep impact since it can potentially cause imbalances between airspace capacity and actual traffic load. When foggy conditions are present, airport managers activate specific low-visibility procedures to sustain safely operations in reduced visibility. Operating under reduced-visibility conditions poses potential problems for the managing of the Air Traffic System since there can be an increase of missed approaches rates, larger time-intervals between landings or take-offs, increase in air traffic controllers and pilots workload, higher probability of runway incursion and, in extreme situations, suspension of runway operations. Accordingly, if Air Traffic Flow and Capacity Managers had the possibility to access early and highly reliable meteorological information about the expected visibility conditions at airports, they would probably make better decisions about the measures that can be activated in the next hours to cope with expected temporary and substantial reductions in airspace and airport capacity. In this sense, 24 h ahead accurate forecast of low-visibility conditions at airports could be of great significance for improving the safety and cost-efficiency of aeronautical operations.

Some studies have dealt with the need of developing reliable systems to detect fog formation in real time. For example, in Ahmed et al. (2015) a bi-spectral brightness temperature difference (BTD) technique applied to satellite images with a specific BTD threshold for a northern India region has been successfully used to detect nighttime fog formation. This work has been recently improved in Dey (2018) where the theoretical aspect of the new BTD threshold is discussed.

On the other hand, previous works have discussed different models for low-visibility events prediction at different time-horizons. In Colabone et al. (2015), a multilayer perceptron for hourly fog forecasting is considered, to assist in planning flight activities of the Academia da Força Aérea (AFA), using meteorological data. In Dutta and Chaudhuri (2015) a decision tree algorithm and an artificial neural network (ANN) approach are applied to a problem of very short-term (3 h) fog prediction in India. In Bremnes and Michaelides (2007), a neural network approach is applied to a problem of fog prediction in Norway, with a prediction time-horizon up to 6 h. In Cornejo-Bueno et al. (2017) different machine learning regression techniques have been applied to fog prediction in Spain.

In Marzban et al. (2007), three sources of information are combined in order to forecast ceiling and visibility at 39 airports of the Northwest of the United States by using ANNs, with a time-horizon up to 12 h. In Roquelaure et al. (2009), a Local Ensemble Prediction System (Roquelaure and Bergot 2008) (LEPS) is used for fog prediction at Paris-Charles de Gaulle Airport, with a maximum prediction time horizon of 12 h. In Chmielecki and Raftery (2011), a Bayesian model is applied to a problem of fog forecasting in different stations of USA North West Pacific, with a prediction time-horizon of 12 h. Finally, a model based on ANN is presented in Fabbian et al. (2007), where up to 18 h fog events prediction is tackled at Canberra International Airport (Australia). Note that these works mainly tackle the problem of fog prediction as a regression approach. The proposed methodology is completely novel in the sense that the fog prediction is now associated with class intervals, leading to an ordinal classification problem.

In this paper, we tackle the problem of low-visibility events prediction using three ordinal categories, based on multi-variant temporal series, at Valladolid airport (Spain). This way, our main interest is related to time series prediction (Chatfield 2000; Brockwell and Davis 2016), which is one of the most important tasks associated with time series datasets. Time series prediction consists of the prediction of the variable value based on the previous values and external data related to the problem being tackled. Auto-Regressive (AR) models (Brockwell and Davis 2013) are one of the most simple approaches for dealing with time series prediction. These models predict the target (future) value using a fixed number of lagged (past) values. The main disadvantage of AR models is the fact that the number of previous time series points to be considered, order of the model or window size, p, greatly influences the accuracy of the prediction. This number depends on the dynamics of the time series and can even depend on the specific part of the time series being considered. In order to adjust this parameter, different procedures exist in the state of the art. Partial Autocorrelation Function (PACF) analysis studies the partial correlation of a time series with its own lagged values. When exogenous time series are used as inputs, cross-correlation analysis compares each input time series with respect to the target one. Another possibility is applying a nested validation procedure: some training patterns are reserved to evaluate the error of a given order, the training process is repeated for different orders, and the final model is constructed considering the lowest validation error. The main problem of nested validation is the associated computational cost.

This paper considers a classification perspective for low-visibility events prediction. Specifically, the problem is tackled as ordinal classification (Gutiérrez et al. 2016), given that the categories involved in the problem (visibility condition at the airport) show a natural order. Note that time series categorical prediction (Guanche et al. 2014) has received less attention in atmospheric literature than standard regression scenarios (Lydia et al. 2016). Nonetheless, all these proposals still suffer from the difficulty of adjusting the order of the model. In order to avoid this issue, in this paper we consider a dynamic window instead of a fixed one. Basically, the past window used for predicting the next value is extended until a change in the class of the time series is found. After this, given that the number of inputs has to be constant, all the values in the window are summarized using two statistics (average and standard deviation). This approach has been successfully applied in Gutiérrez et al. (2015) for wave energy flux classification. One important problem with this dynamic window approach is that the only information about the time series exploded for the design of the window is the class label. Moreover, although dynamic windows can improve the performance of AR models, it is clear that, depending on the specific application, the loss of information produced by summarizing the window can decrease the performance of the final model. This is specially true for application fields where recent values highly influence future predictions, as there is a strong correlation for events close in time. In this way, it is advisable to combine fixed and dynamic windows, in order to obtain the benefits of both perspectives. This is the approach followed in this paper, where an autoregressive ordinal classifier is trained based on the combined use of a fixed window and a dynamic one. With these previous considerations in mind, the main contributions of this paper are the following: 1) to carry out a 24-h time-horizon prediction of low-visibility events at Valladolid airport (Spain), mainly driven by radiation fog situations; 2) to introduce a combination of dynamic and fixed windows to better summarize previous information of the visibility time series; 3) to use ordinal classification algorithms in the prediction of the categorical visibility time series.

The remainder of this paper is organized as follows: In Section 2, the dataset considered is introduced and explained, detailing the process for final data acquisition. The methodology used for solving the problem is presented in Section 3, including subsections for describing the input window proposals (Section 3.1), the ordinal classification algorithms (Section 3.2), the evaluation performance metrics (Section 3.3), a summary of the methodology considered (Section 3.4) and an enumeration of the different methods compared (Section 3.5). The experiments and results obtained in the prediction of low-visibility events at Valladolid airport are described in Section 4. Section 5 concludes the paper with some final remarks.

2. Dataset used

Given that the aim of this work is the prediction of low-visibility events due to the presence of fog, we have centred our study in one of the most affected Spanish airports in terms of number of reduced-visibility days. Valladolid Airport (latitute: 41.70 N; longitude: 4.88 W; elevation: 846 m) is located in a very extensive plateau, surrounded by