2018 IEEE 14th International Conference on e-Science

Visibility Prediction based on kilometric NWP Model Outputs using Machine-learning Regression

Driss BARI
National Research and System Information Department, CNRMSI/SMN
Direction de la Météorologie Nationale
Casablanca, Morocco
bari.driss@gmail.com

***Abstract*—Low visibility conditions have a strong impact on air and road traffics and their prediction is still a challenge for meteorologists, particularly its spatial coverage. In this study, an estimated visibility product over the north of Morocco, from the operational NWP model AROME outputs using the state-of-the art of Machine-learning regression, has been developed. The performance of the developed model has been assessed, over the continental part only, based on real data collected at 37 synoptic stations over 2 years. Results analysis points out that the developed model for estimating visibility has shown a strong ability to differentiate between visibilities occurring during daytime and nighttime. However, the KDD-developed model have shown low performance of generality across time. The performance evaluation indicates a bias of -9m, a mean absolute error of 1349m with 0.87 correlation and a root mean-square error of 2150m.**

***Keywords*-Visibility; Machine Learning; NWP model; AROME; Regression;**

I. Introduction

The occurrence of adverse visibility conditions that restricted the flow of air traffic in major airport terminals is one of the main causes of aircraft delays and is crucial for air traffic safety and economic issues. For example, a dense fog event (visibility below 200 m) obliged the diversion of 21 aircrafts in 2008, which supposed to land at Mohamed V international airport (GMMN), Casablanca, Morocco, to other national airports. Such flights diversion results in considerable cost to airlines (e.g. extra fuel consumption and hotel accommodation for the passengers). This indicates the need to improve visibility forecasting.

To deal with the aforementioned deficiency, meteorologists already use some physical numerical weather prediction (NWP) models. The latter provide guidance for operational with useful insight into future atmospheric conditions. In this context, researches at the Moroccan National department of Meteorology (DMN), have used three-dimensional (Bari *et al.*, 2015 [1]) and one-dimensional (Bergot *et al.*, 2005 [2]) NWP approaches. However, generating accurate visibility forecasts from these models remains a challenge, especially over airports, due to the complex interaction between the physical processes during the life cycle of such low visibility conditions.

As an alternative, Data mining is emerging as a suitable method for extracting patterns from extensive sets of heterogeneous data related to detection and prediction of meteorological phenomena (e.g. Bari and El Khlifi, 2015 [3], Bartokova et al., 2015 [4]). Using machine-learning regression, Cornejo-Bueno *et al.* (2017) [5] developped a model for low visibility events prediction at Valladolid airport, Spain. The proposed model focuses on runway visual range at this airport. Kneringer *et al.* (2018) [6] developped an ordered logistic regression model to nowcast the probabilities of the low visibility procedure (LVP) categories at Vienna International Airport during the cold season. All the previous studies are site specific, particularly at airport. Thus, there is a lack in the literature regarding the evaluation of the performance of such machine learning techniques for visibility prediction over a large domain.

In an attempt to overcome the aforementioned deficiency, the present study used some machine learning regression techniques and applied them to kilomteric NWP model outputs in order to examine suitability of Knowledge from Database Discovery (KDD) methods for spatial visibility occurrence forecasting.

II. Data and methods

A. Data preperation

The dataset used in this work, is extracted from output of the operational 3D NWP model AROME (Seity *et al.*, 2011 [7]). It is a non-hydrostatic model with 2.5km at horizontal resolution and 90 vertical levels where the first one is located at 10m. The study period covers 2 years from March 2015 to March 2017. The tri-hourly forecasted meteorological parameters, from midnight run, were made available for use in the development of visibility forecast algorithms. The constructed database contains predictors representing the atmospheric boundary layer and includes among others, atmospheric pressure, air temperature, relative humidity, liquid water content (LWC), wind speed and direction, covering the northern part of Morocco (Fig. 1a). This

978-1-5386-9156-4/18/$31.00 ©2018 IEEE
DOI 10.1109/eScience.2018.00048

domain, which is a fog-prone region (Bari *et al.* (2016)[8]), contains the main national airports and is characterized by a high density of population and road traffic. In the very-short forecasting framework, the proposed machine-learning algorithms were derived for 3-, 6-, 9-, and 12-h forecasts through a data-mining process. The performance of the developed model has been assessed, over the continental part only, based on real data collected at 37 synoptic stations over the study period, spatially distributed over topographically heterogeneous terrain (Fig. 1b).

B. Modelling

1) Data mining: A wide variety of data-mining tools and algorithms are available in the literature. These methods include two classes: unsupervised methods like clustering and associations and supervised methods like decision trees/rules, neural network and k nearest neighbor. In this research work, boosting tree-based regression and deep learning are used as supervised algorithms to train the data.

The boosting method combine an ensemble of decisions trees and can model nonlinear data features. With boosting, new decision trees always grow on forecast information of previously grown trees, since the new tree is fitted to residuals of the previous ones. To develop the boosted trees, XGBoost (Extreme Gradient Boosting), which is a scalable machine learning system for tree boosting, has been used (Chen and Guestrin, 2016 [9]). In this approach, we are learning functions (trees) instead of learning numerical weights by additive training (boosting).

Deep learning (Schmidhuber (2015) [10]) is based on a multi-layer feedforward artificial neural network that is trained with stochastic gradient descent using back-propagation. This method has been used by Lei et al. (2017) [11] to predict visibility at the Urumqi international airport in China. The network can contain a large number of hidden layers consisting of neurons with an activation functions. Many configurations of the network arichitecture has been tested and the best one which has been retained is denoted as 20-15-5-1. This means that the deep learning network is composed from two layers of hidden nodes, one with 15 nodes and the other with 5, with each node in a layer being connected to every node in the next layer.

2) Experiment methodology: In this work, the training (70% of all data) and testing (30% of all data) sets have been created by a random split of the whole dataset. Over the study period, the reduced visibility below 5km is more frequent during night-time (77.69%) than during daytime (22.31%) over the study domain. Then, several experiment permutations were performed in order to evaluate the generality of our results over the whole day. Thus, the following configurations (Table I) have been performed over three sets of data covering: (1) the whole day, (2) the daytime only and (3) the night-time only.

Table I: Experiment configurations setup

		Training		
		daytime	night-time	whole day
Test	daytime	X	X	X
	night-time	X	X	X
	whole day			X

III. RESULTS

To evaluate the potential of KDD-produced algorithms of estimating the continuous parameters, the results of the datamining experiments on testing dataset are presented below. Correlation coefficient (CC) is used to measure the relationship of the algorithm output with observation; mean absolute error (MAE); root-mean-square error (RMSE) and bias (BIAS) are used to measure the accuracy of the developed algorithms.

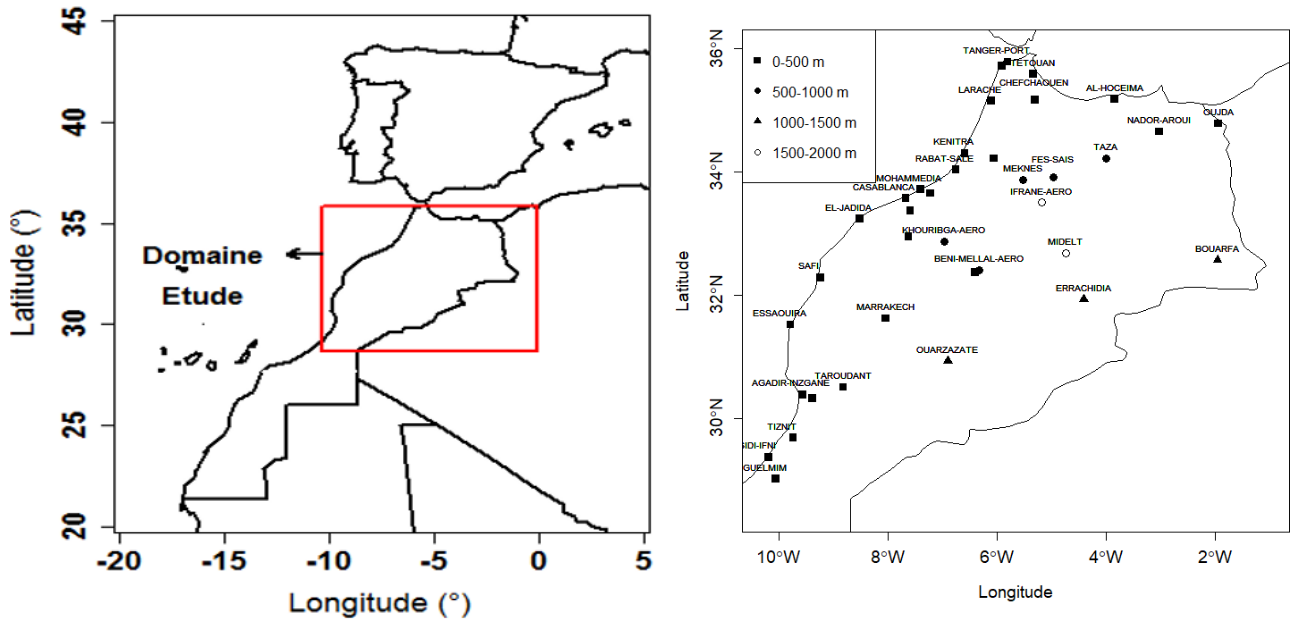
Tables II summarize the performance statistics of visibility estimation on the testing dataset, issued respectively from XGBoost- and Deep learning-based models. For reference, the performance of the persistence is added. Both KDD-developped models outperforms persistence except similarity of mean absolute error with deep learning algorithm. The best performance is related to algorithms using ensemble tree-based approach (XGBoost) with a bias of -9m, a mean absolute error of 1349m with 0.87 correlation and a root mean-square error of 2150m.

Table II: Performance statistics of visibility estimation on the testing dataset issued from models based on XGBoost, deep learning and persistence.

Method	Bias (m)	RMSE (m)	MAE (m)	CC
XGBoost	-9.05	2149.45	1348.97	0.87
Deep Learning	-7.5	2464.8	1694.87	0.83
Persistence	-7.2	3073.36	1601.15	0.75

To evaluate the diurnal performance dependance of the best developed model, the verification scores of the unified model, applied separately to daytime and nighttime data, have been investigated (Table III). This is done to check if it is enough to develop a single model trained on data covering the whole day or it is necessary to develop two models where the first was trained on data observed during daytime and the other one has been trained on nighttime data.

The analysis of the verification scores (Table III) shows that the developed models associated with the three configurations slightly underestimate the horizontal visibility with an average bias of about -9m associated with a strong linear correlation between the observed visibility and that estimated, exceeding 0.8. Overall, the performance



(a)

(b)

Figure 1: (a) Map showing the study domain. (b) Map showing the spatial distribution of the synoptic stations used in this study.

Table III: Performance statistics of visibility estimation by models trained and tested on data covering the same period of the day.

Method	Bias (m)	RMSE (m)	MAE (m)	CC
whole day	-9.05	2149.45	1348.97	0.87
Daytime	-10.09	2160.37	1354.52	0.87
Nighttime	-7.96	2168.35	1342.22	0.86

of the three models are similar. Thus, one can conclude that the performance of developed models does not depend on the time of day (daytime and nighttime); thus it is sufficient to develop a single model trained on data covering the whole day.

This leads us to evaluate the capacity of the unified developed model on the basis of all the data covering the whole day, to discriminate between visibilities occurring during the day and those occurring during the night.

Table IV: Performance statistics of visibility estimation issued from the application of the unified developed model on testing dataset occurring separately during daytime and nighttime.

Method	Bias (m)	RMSE (m)	MAE (m)	CC
Daytime	-4.18	1716.14	876.54	0.92
Nighttime	-16.19	1695.81	851.60	0.92

Table IV summarizes the the visibility estimation performance statistics of the unified developed model on the testing dataset occurring separately during daytime and nighttime. It is seen clearly that the verification scores are similar for both experiments with a strong correlation exceeding 0.9 and a mean absolute error of about 860m. This confirms the hypothesis that it is sufficient to develop a single machine learning model based on data covering the whole day.

To evaluate the generalization error of the developed model across time, we compare the verification scores of the developed model trained on daytime data and tested on nighttime data with the opposite one (Table V).

Table V: Performance statistics of visibility estimation by the developed model trained on daytime data and tested on nighttime data with the opposite one.

Method	Bias (m)	RMSE (m)	MAE (m)	CC
TrainDay-TestNight	-12.39	2565.97	1762.07	0.81
TrainNight-TestDay	-560.16	3535.20	2335.97	0.64

Table V shows clearly that the developed model trained on daytime data only and tested on nighttime data outperforms the other model with a strong linear correlation (0.81 vs 0.64) between the estimated visibility and the observed one during the night. Besides, the best configuration indicates a good bias (-12.39m vs -560.16m) and mean absolute error of about 1760m. Thus, one can conclude that KDD-developed models is not able to generalize its performance across time. This deficiency could be due to many factors such as:

- the weak representativness of all stations in the training and testing datasets due to random split of the dataset.
- some synoptic stations do not operate during the night. Then, the availability of observed data during the day is higher than during the night.
- the uncertainties in AROME forecasts associated with model error.

To illustrate the derived product from the output of KDD-developed model, we presented in Fig. 2 an example of a situation with fog occurring over the coastal part of the study domain. Overall, this figure shows that the developed model captures well the spatial coverage of the fog patch along the coast in comparison with observation at synoptic stations and derived product from MSG satellite based on the brightness temperature difference between the infrared channels IR3.9 μm and IR10.8 μm at 0600UTC on 09 January 2017. Due to lack of observation over sea, the verification has been done over the continental part only.

IV. Conclusion

Motived by aeronautical requirements for more accurate assessment of visibility, an improved utilization of NWP model output data sources to estimate visibility, using Data mining methods, has been developed. In this research work, our main goal is to examine suitability of data mining methods for visibility forecasting over a large domain.

The performance of the developed model has been assessed, over the continental part only, based on real data collected at 37 synoptic stations over 2 years. Results analysis points out that the developed model for estimating visibility has shown a strong ability to differentiate between visibilities occurring during daytime and nighttime. However, the KDD-developed model have shown low performance of generality across time. The performance evaluation indicates a bias of -9m, a mean absolute error of 1349m with 0.87 correlation and a root mean-square error of 2150m.

References

[1] Bari, D., Bergot, T., and El Khlifi, M. ”Numerical study of a coastal fog event over Casablanca, Morocco”. Quarterly Journal of the Royal Meteorological Society, 141(690), 1894-1905. 2015.

[2] Bergot, T., Carrer, D., Noilhan, J., and Bougeault, P. ”Improved site-specific numerical prediction of fog and low clouds: A feasibility study”. Weather and Forecasting, 20(4), 627-646. 2005.

[3] D. Bari, and M. EL Khlifi. ”LVP conditions at Mohamed V airport, Morocco:local characteristics and prediction using neural networks”. International Journal of Basic and Applied Sciences, 4 (4), pp. 354-363, 2015.

[4] I. Bartoková, A. Bott, J. Bartok and M. Gera. ”Fog Prediction for Road Traffic Safety in a Coastal Desert Region: Improvement of Nowcasting Skills by the Machine-Learning Approach”. Springer Science+Business Media Dordrecht, 2015.

[5] L. Cornejo-Bueno, C. Casanova-Mateo, J. Sanz-Justo, E. Cerro-Prada and S. Salcedo-Sanz. ”Efficient Prediction of Low-Visibility Events at Airports Using Machine-Learning Regression”. Springer Science+Business Media B.V. , 2017.

[6] Kneringer, P., Dietz, S. J., Mayr, G. J., and Zeileis, A. ”Probabilistic nowcasting of low-visibility procedure states at Vienna International Airport during cold season”. Pure and Applied Geophysics, 1-13. 2018

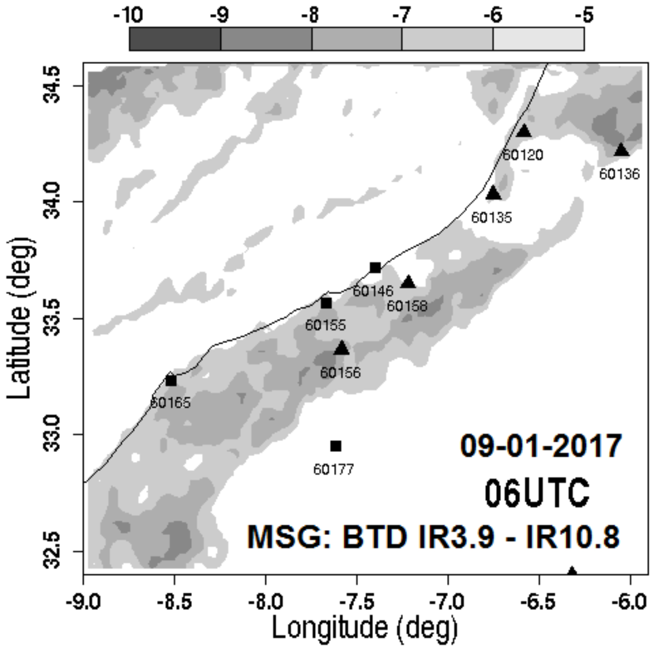
[7] Seity, Y., Brousseau, P., Malardel, S., Hello, G., Benard, P., Bouttier, F., Lac, C. and Masson, V. ”The AROME-France convective-scale operational model”. Monthly Weather Review, 139(3), pp.976-991. 2011.

[8] D. Bari, T. Bergot, and M. El Khlifi. ”Local meteorological and large-scale weather characteristics of fog over the grand casablanca region, Morocco”. Journal of Applied Meteorology and Climatology, 55(8), 1731-1745. 2016.

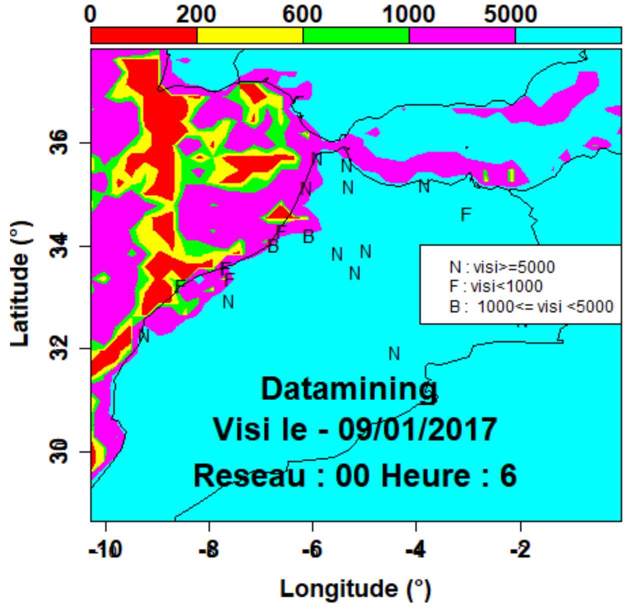
[9] T. Chen, and D. Guestrin.”Xgboost : A scalable tree boosting system”. In Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD ’16, New York, NY, USA, 2016.

[10] Schmidhuber, J. ”Deep learning in neural networks: An overview”. Neural networks, 61, 85-117. 2015.

[11] Z. Lei, Z. Guodong, H. Lei and W. Nan. ”The Application of Deep Learning in Airport Visibility Forecast”. Atmospheric and Climate Sciences, 7 pp. 314-322, 2017.



(a)



(b)

Figure 2: (a) Brightness temperature difference (BTD in K) images (shaded in grey scale) between the infrared channels IR3.9 and IR10.8 μm at 0600UTC on 09 January 2017. Triangles refer to airports. (b) KDD-developed model output for estimating visibility 6 hours later from run of midnight of NWP model on 09 January 2017. "F" refers to observed foggy conditions (visibility below 1km). "B" refers to observed mist (visibility between 1km and 5km). "N" refers to observed visibilities above 5km. The positions of all these symbols represents the locations of the synoptic stations.