Fog prediction using artificial intelligence: A case study in Wamena Airport

Ristiana Dewi1*, Prawito1 and Hastuadi Harsa2

1Departement of Physics, Universitas Indonesia, Depok, Indonesia
2Agency for Meteorology Climatology and Geophysics, Jakarta, Indonesia

*ristiana.dewi.81@sci.ui.ac.id

Abstract. Fog is one of the atmospheric phenomena that affect airport operations. It can reduce visibility which impacts flight operations (taxiing, take-off, landing). Therefore, fog prediction is needed to support flight safety. The biggest challenge in making weather predictions is the chaotic and complicated process of the atmosphere. This research tries to use artificial intelligence (AI) to predict fog events at Wamena Airport. Design of model prediction using hourly synoptic data set from January 2015 till May 2018. Variables input such as dry ball temperature, wet ball temperature, dew point, relative humidity, cloud cover, wind direction, wind speed, visibility, and present weather for the past six hours ago are used to predict fog or no fog events. We performed a grid search parameter tuning on five algorithms such as Distributed Random Forest (DFR), Deep Learning (DL), Gradient Boosting Machine (GBM), Generalized Linear Model (GLM), and Extreme Randomized Tree (XRT). The best model is obtained from the ensemble model Stacked Ensemble (SE) with an accuracy of above 90% for the fog forecast from one to three hours later.

1. Introduction

According to WMO [1], fog is a suspension of very small, usually microscopic water droplets in the water, reducing visibility at the Earth's surface to less than one kilometer. Visibility information is needed for pilots to take off and landing [2]. Even though modern airports have the Instrument Landing System (ILS) installed, pilots still need a visual reference to touch down. Low visibility has the potential to cause plane crashes. If visibility suddenly deteriorates when the plane is about to land, the pilot must decide to go round and round waiting for weather conditions to improve or landings to be diverted to another airport due to fuel considerations. This causes delays and flight cancellations which are detrimental to airlines and passengers.

One of the big challenges in making fog forecasts is a very complex and chaotic atmospheric process. In these conditions, a forecaster must make accurate forecasts. Numerical Weather Prediction (NWP) is one of the most widely used methods by scientists in the world to weather forecasts. Many researchers have tried to evaluate the performance of high-resolution NWP models to predict fog events [3, 4, 5]. However, the fog remains difficult to predict because of the local scale, and the limited knowledge of the atmosphere that is so complex that many factors that control the formation of fog have not been sufficiently well simulated by the NWP model [6,7].

Development of technology, weather forecasts begin to be made by applying data science using artificial intelligence. The Artificial Neuron Network algorithm is used in some fog forecast studies.
from synoptic weather observations [2, 8]. Other studies try to use the Decision Tree algorithm from weather data from the NWP model to make decisions in making fog forecast [9]. Several regression-based algorithms such as Gaussian, SVR, MLP, and Wavelet have been tested to find the best algorithm in predicting visibility [10]. Forecast visibility at an airport has been carried out [11] by building a regression model with the deep learning algorithm using synoptic weather observation data. The advantage of the forecast model with artificial intelligence is that its objective nature does not depend on the forecaster subjectivity. Also, the process of making forecasts is faster. This is very helpful for flight operations that require accuracy and speed of time in providing weather information services.

2. Experiment

2.1. Study Area
Geographically, Jayawijaya Regency is located on a stretch of the Baliem valley, an alluvial valley that stretches to an altitude of 1500-2000 m above sea level. The Baliem valley is surrounded by the Jayawijaya Mountains which are famous for their eternal snow peaks, including the Trikora Peak (4,750 m), Mandala Peak (4,700 m) and Yamin Peak (4,595 m). The steep mountain slopes and narrow and steep river valleys are the hallmarks of these mountains. This makes the aircraft that will land must pass through the cracks of the mountains. When the weather deteriorates and the fog starts to fall, the gap will close and obstruct the flight path.

2.2. Data
We used 10 meteorological variables as described in table 1. Those variables were recorded hourly. Initially, the data were composed of 29175 records x 10 columns. We then expanded those variables using their lags. There were six lags added to each variable. Therefore, each variable had seven-column eventually. We excluded the hour variable for this process. This was because the hour parameter had described all six additional columns for every record. Hence, the dimension of the data were 29175 rows x 64 columns. These were the composition of independent variables for the model to predict the response variables later. The response variables were the weather condition having the status of it was fog or not (denoted as FOG and NO FOG). There were three response variables obtained from the observation. These three response variables were the extension of the present weather condition. The present weather condition was expanded to three hourly lead condition. Therefore, they constructed three columns as response variables in the data. The final dataset dimension was 29175 rows x 67 columns. From this point, the objective function of this paper is the development of three hours ahead of fog prediction using the AI model.

| Parameter name            | Code | Type     | Units     |
|---------------------------|------|----------|-----------|
| dry bulb temperature      | tt   | numeric  | °C        |
| wet bulb temperature      | tw   | numeric  | °C        |
| dew point                 | td   | numeric  | °C        |
| visibility                | vv   | numeric  | km        |
| relative humidity         | rh   | numeric  | %         |
| observation hour          | hour | factor   | hour      |
| cloud cover               | n    | factor   | okta      |
| wind direction            | dd   | factor   | °north    |
| wind speed                | ff   | factor   | knot      |
| weather condition         | ww   | actor    | fog, rain, nil |

The fog formation process requires stable air conditions, calm winds, and high relative humidity. Wind directions are considered by weather factors locally. Fog is a local scale weather phenomenon. A Clear sky is also a factor forming fog especially radiation fog. Low visibility is an impact of the fog phenomenon so the variable needs to be monitored. The weather conditions variable in this research is
categorized as rain, fog, and nil (sunny or cloudy). When it rains, the humidity is relatively high. This condition is the same as the fog. For the learning process, rain is considered in this research for the learning process model. Time observation is entered into the model because radiation fog often occurs in the early hours of the morning.

3. Methodology
The programming language used for this research is R. The other supporting applications using RStudio and H2O. In the model building process, we implemented a grid search strategy to find a better combination of the model’s parameters. We also gave a time constraint for the model to update their knowledge in a particular hyperspace model parameter location. The maximum time constraint was four minutes. The data were divided into three parts (training, validation, and testing). The number of the split was 70% for training, 20% for validation, and 10% for testing.

3.1. Data Pre-processing
Data transformation was useful to change the actual value of a variable into a certain range so it can be read by the program. In this study, data transformations use the standardization method for numeric types and one-shot encoding for factor types.

3.2. Model
The cross-validation method used five different data composition. There were five AI models utilized to perform the prediction task. These models were Distributed Random Forest (DFR), Deep Learning (DL), Gradient Boosting Machine (GBM), Generalized Linear Model (GLM), Extreme Randomized Tree (XRT), and Stacked Ensemble (SE).

3.2.1. Distributed Random Forest. DFR method is an algorithm used for data classification by combining many classification trees. This algorithm can handle large input variables, and balance errors in unbalanced datasets [12].

3.2.2. Deep Learning. DL is a development of machine learning based on Artificial Neural Network (ANN) with many hidden layers that have the ability to learn data features automatically [13]. Deep learning architecture consists of the input layer, hidden layer, and output layer where the weight of each unit of the perceptron is optimized using the backpropagation algorithm.

3.2.3. Gradient Boosting Machine. GBM is an ensemble learning method for regression or classification. This method works to improve weak learning machines to become powerful. The algorithm trains the decision tree by giving the same weight to each observation. After evaluating the first tree, the forecast is corrected by adding weight observations that are difficult to classify and reduce the weight that is easily classified. The second tree is built based on the new weight value. Calculation of misclassification from the ensemble of the two trees is used to grow the third tree and so on.

3.2.4. Generalized Linear Model. GLM is the development of a linear regression model with the assumption that the predictor has a linear effect but does not assume a particular distribution of response variables that are members of an exponential family [14].

3.2.5. Extreme Randomized Tree. XRT is similar to Random Forest but there are two differences that XRT does not resample observations when building a tree. XRT also do not use the best split. The SE models were constituted by some models (DRF, DL, GBM, GLM, and XRT). The models were ranked from the best to the worst using the Area Under the Curve (AUC) as the main scoring metric. Some other metrics were also employed as additional information about the performance of the models. They are Precision-Recall AUC (PRAUC), Gini coefficient, Mean Per Class Error (MPCE), Logloss, Root
Mean Squared Error (RMSE) and Mean Squared Error (MSE). In the model building process, we also implemented a grid search strategy to find a better combination of the model’s parameters.

4. Results and Discussion

The metric performances are displayed by tables 2, 3 and 4 for one, two and three-hour lead prediction respectively. The data displayed by the tables are the cross-validation metrics. The metrics are sorted descendingly for AUC, PRAUC, and Gini. MPCE, Logloss, RMSE, and MSE are sorted ascendingly.

### Table 2. Model performance for one hour lead

| Model                          | AUC  | PRAUC | Gini  | MPCE  | Logloss | RMSE  | MSE    |
|--------------------------------|------|-------|-------|-------|---------|-------|--------|
| Stacked Ensemble All Models    | 0.96500 | 0.89706 | 0.93000 | 0.12425 | 0.16674 | 0.20997 | 0.04409 |
| GBM grid 1                     | 0.96494 | 0.92877 | 0.92987 | 0.12733 | 0.16504 | 0.21470 | 0.04610 |
| GBM 1                          | 0.96459 | 0.95964 | 0.92917 | 0.12390 | 0.15746 | 0.20944 | 0.04386 |
| Stacked Ensemble Best Of Family| 0.96435 | 0.89087 | 0.92870 | 0.12282 | 0.16802 | 0.21142 | 0.04470 |
| GBM 2                          | 0.96404 | 0.92068 | 0.92807 | 0.13294 | 0.15936 | 0.20933 | 0.04382 |
| GBM 3                          | 0.96379 | 0.91262 | 0.92758 | 0.13066 | 0.16264 | 0.21042 | 0.04428 |
| GBM 5                          | 0.96375 | 0.93471 | 0.92749 | 0.12278 | 0.15849 | 0.20963 | 0.04395 |
| GBM 4                          | 0.96296 | 0.94143 | 0.92593 | 0.12232 | 0.17002 | 0.21103 | 0.04453 |
| DRF 1                          | 0.96174 | 0.45026 | 0.92347 | 0.12881 | 0.19651 | 0.21389 | 0.04575 |
| GLM grid 1                     | 0.95887 | 0.84830 | 0.91774 | 0.14314 | 0.16990 | 0.21972 | 0.04828 |
| Deep Learning 1                | 0.95797 | 0.56598 | 0.91595 | 0.14329 | 0.18100 | 0.22376 | 0.05007 |
| XRT 1                          | 0.95776 | 0.72525 | 0.91553 | 0.13113 | 0.18771 | 0.22013 | 0.04846 |
| Deep Learning grid 1           | 0.94996 | 0.34813 | 0.89992 | 0.14792 | 0.24293 | 0.23036 | 0.05307 |

### Table 3. Model performance for two hour lead

| Model                          | AUC  | PRAUC | Gini  | MPCE  | Logloss | RMSE  | MSE    |
|--------------------------------|------|-------|-------|-------|---------|-------|--------|
| Stacked Ensemble All Models    | 0.94229 | 0.90963 | 0.88458 | 0.20224 | 0.21764 | 0.24875 | 0.06188 |
| GBM 3                          | 0.94171 | 0.89943 | 0.88342 | 0.20551 | 0.21037 | 0.24917 | 0.06208 |
| Stacked Ensemble Best Of Family| 0.94170 | 0.91430 | 0.88340 | 0.21157 | 0.21810 | 0.24905 | 0.06203 |
| GBM 1                          | 0.94161 | 0.95653 | 0.88321 | 0.20523 | 0.20669 | 0.24893 | 0.06196 |
| GBM 2                          | 0.94108 | 0.92584 | 0.88216 | 0.19244 | 0.20814 | 0.24830 | 0.06165 |
| GBM grid 1                     | 0.94050 | 0.92674 | 0.88099 | 0.19541 | 0.21682 | 0.25226 | 0.06364 |
| GBM 4                          | 0.94023 | 0.94612 | 0.88045 | 0.20251 | 0.20839 | 0.24921 | 0.06210 |
| DRF 1                          | 0.93942 | 0.92768 | 0.87884 | 0.20985 | 0.20251 | 0.25218 | 0.06359 |
| GLM grid 1                     | 0.93141 | 0.87065 | 0.86281 | 0.23733 | 0.22106 | 0.25889 | 0.06702 |
| Deep Learning 1                | 0.93058 | 0.64246 | 0.86116 | 0.22956 | 0.22832 | 0.26235 | 0.06883 |
| Deep Learning grid 1           | 0.91764 | 0.50741 | 0.83528 | 0.24131 | 0.29539 | 0.27587 | 0.07610 |
| Deep Learning grid 2           | 0.91175 | 0.51789 | 0.82351 | 0.24801 | 0.28375 | 0.27997 | 0.07839 |

### Table 4. Model performance for three hour lead

| Model                          | AUC  | PRAUC | Gini  | MPCE  | Logloss | RMSE  | MSE    |
|--------------------------------|------|-------|-------|-------|---------|-------|--------|
| Stacked Ensemble All Models    | 0.92577 | 0.93048 | 0.85154 | 0.23116 | 0.24666 | 0.26919 | 0.07247 |
| Stacked Ensemble Best Of Family| 0.92513 | 0.92362 | 0.85027 | 0.23946 | 0.24685 | 0.26961 | 0.07269 |
| GBM 1                          | 0.92500 | 0.92621 | 0.85000 | 0.26448 | 0.23566 | 0.27017 | 0.07299 |
| GBM 2                          | 0.92415 | 0.92877 | 0.84830 | 0.25369 | 0.23789 | 0.27036 | 0.07309 |
| GBM 5                          | 0.92336 | 0.92871 | 0.84671 | 0.24975 | 0.23698 | 0.27018 | 0.07299 |
| GBM grid 1                     | 0.92295 | 0.93485 | 0.84589 | 0.25804 | 0.24602 | 0.27257 | 0.07430 |
| GBM 4                          | 0.92190 | 0.95303 | 0.84379 | 0.26555 | 0.24344 | 0.27261 | 0.07432 |
| DRF 1                          | 0.91776 | 0.56466 | 0.83552 | 0.24339 | 0.27674 | 0.27277 | 0.07440 |
| GLM grid 1                     | 0.91229 | 0.93598 | 0.82458 | 0.27178 | 0.25021 | 0.27872 | 0.07769 |
| XRT 1                          | 0.90628 | 0.74579 | 0.81256 | 0.28015 | 0.27453 | 0.28453 | 0.08096 |
| Deep Learning 1                | 0.90418 | 0.73876 | 0.80836 | 0.28141 | 0.26328 | 0.28539 | 0.08145 |
| Deep Learning grid 1           | 0.89825 | 0.52245 | 0.79650 | 0.30473 | 0.30949 | 0.29283 | 0.08575 |
We used the sequence of metrics to break the ties in deciding which model was better than the others. Based on the trial and error process, there were thirteen best models in each lead hour. The best model is found in the first row. For all prediction days, the Stacked Ensemble models yielded the best performance among other models. We took only the best model and disregard the rest.

After the best forecast model is obtained through the process, the next step is to test the model. The confusion matrix is one method that can be used to measure the performance of a classification method. The confusion matrix for the cross-validation of all ensembles models is summarized in tables 5, 6, and 7. The confusion matrix for the test data is summarized in tables 10, 11, and 12. Based on the error value on the total column, we can find out the accuracy of the model.

### Table 5. Confusion matrix for one hour lead test data

|       | FOG | NO FOG | Error (%) |
|-------|-----|--------|-----------|
| FOG   | 2477| 739    | 22.98     |
| NO FOG| 322 | 16884  | 1.87      |
| Totals| 2799| 17623  | 5.20      |

### Table 6. Confusion matrix for two hour lead test data

|       | FOG | NO FOG | Error (%) |
|-------|-----|--------|-----------|
| FOG   | 1972| 1217   | 38.16     |
| NO FOG| 394 | 16839  | 2.29      |
| Totals| 2366| 18056  | 7.89      |

### Table 7. Confusion matrix for three hour lead test data

|       | FOG | NO FOG | Error (%) |
|-------|-----|--------|-----------|
| FOG   | 273 | 175    | 14.06     |
| NO FOG| 63  | 2407   | 7.09      |
| Totals| 336 | 2582   | 8.16      |

The testing data shows that the model has an error value of 5.11% in the prediction of one hour. The error increases by up to 8.16% for the next two hours. The biggest error occurred in the prediction for 3 hours later 9.6%. This shows that the model will be not good for predictions over a long period.

### 5. Conclusion

In this study prediction of the occurrence of fog using artificial intelligence has been done. Based on the results by conducting trials and errors on several models such as DRF, DL, GBM, GLM, and XRT, it was concluded that the best model was obtained through an ensemble model (Stacked Ensemble). Stacked Ensemble provides the best performance for fog prediction. Based on the testing data, the model obtained an accuracy value of 94.89% for one hour later, 91.84% for 2 hours later, and 90.4% for 3 hours later.

The results of this study can help the forecaster to predict fog at Wamena Airport. This research using synoptic weather observation data at the meteorology station to build a prediction model. This will greatly help meteorological stations that have limited tools and facilities to build weather forecast models. This method is not only limited to fog forecast but can be applied to various weather elements.
References

[1] World Meteorological Organization 1992 *International Meteorological Vocabulary* (Geneva: WMO)

[2] Colabone R D O, Ferrari A L, Vecchia F A da S and Tech A R B 2015 Application of artificial neural networks for fog forecast *Journal of Aerospace Technology and Management* 7(2) 240–246

[3] Bartok J, Bott A and Gera M 2012 Fog Prediction for Road Traffic Safety in a Coastal Desert *Region Boundary-Layer Meteorology* 145(3) 485–506

[4] Payra S and Mohan M 2014 Multirule Based Diagnostic Approach for the Fog Predictions Using WRF Modelling Tool *Advances in Meteorology* 1–11

[5] Singh A, George J P and Iyengar G R 2018 Prediction of fog/visibility over India using NWP Model *Journal of Earth System Science* 127(2) 1–13

[6] Velde I R van der, Steeneveld G J, Schreur B G J, Wichers, and Holtslag A A M 2010 *Modeling and Forecasting the Onset and Duration of Fog Event during Frost Conditions* Paper presented at 5th International Conferences on Fog, Fog Collection and Dew (Munster: 25-30 July 2010)

[7] Abhishek K, Singh M P, Ghosh S and Anand A 2012 Weather Forecasting Model Using Artificial Neural Network *Procedia Technology* 4 311-318

[8] Fabbian D, Dear R D and Lellyett S 2007 Application of Artificial Neural Network Forecast to Predict Fog at Canberra International Airport *Weather and Forecasting* 22 372-381

[9] Bartokova I, Bott A, Bartok J and Gera M 2015 Fog Prediction for Road Traffic Safety in a Coastal Desert Region: Improvement of Nowcasting Skills by the Machine-Learning Approach *Boundary-Layer Meteorology* 157(3) 501–516

[10] Cornejo-Bueno L, Casanova-Mateo C, Sanz-Justo J, Cerro-Prada E and Salcedo-Sanz S 2017 Efficient Prediction of Low-Visibility Events at Airports Using Machine-Learning Regression *Boundary-Layer Meteorology* 165(2) 349–370

[11] Zhu L, Zhu G, Han L and Wang N 2017 The Application of Deep Learning in Airport Visibility Forecast *Scientific Research Publishing* 7 314-322

[12] Kurniawan F A, Adiwijaya and Kurniati A P 2011 Analisis dan Implementasi Random Forest dan Classification dan Regression Tree (CART) untuk Klasifikasi Pada Misuse Intrusion Detection System *Universitas Telkom*

[13] Hinton G, Bengio Y and LeCun Y 2015 Deep learning review *Nature* (521) 436-444

[14] Nelder J A and Wedderburn 1972 Generalized Linear Model. *Journal of the Royal Statistical Society Series A* 135 (3)