POSTER 02

Study on Frost Prediton Model with Machine Learning Method

Y. S. Kim*, K. M. Shim, M. P. Jeong, I. T. Choi, and K. K. Kang

Climate and Agroecology Division, National Institute of Agricultral Sciences, RDA, Korea

I. Introduction

Weather conditions for frost occurrence are that the skies are clear and calm, the temperature is falling fast, the wind is blowing from the north, the moisture in the atmosphere condenses to form water drops. These events may have serious consequences on crop production, so actions like irrigation or wind machines operation must be taken to minimize damaging effects. But it is no use without information about when to start those actions. To solve this problem, the studies on frost prediction have been carried out.

For example, Han *et al.*(2009) carried out a study on frost forecasting based on discriminant analysis of climate data in Naju, Korea. Verdes *et al.*(2000) carried out a study on frost prediciton using Artificial Neural Networks (ANN), Simple Bayes (SB) and k-Nearest Neighbors (k-NN). Mort and Robinson (1996) developed an ANN-based system to predict overnight frost formation in Sicily, Italy. The input variables were the previous day's minimum and maximum temperatures, amount of cloud, maximum wind speed and direction, humidity, wind speed and wind direction at 1900.

Also, this study carried out to compare accuracy of models to predict using climatc elements if frost would occur the next day.

II. Materials and Methods

The weather and frost event data for this study were obtained from the Korea Meteorological Administration (KMA) for the years 2004 through 2013. Three climate variables were used temperature at 2400 (T24), amount of cloud (Cloud) and amount of precipitation within 5 days (Pre5) prior to an event of frost or non-frost (Han *et al.,* 2009).

We used a randomly selected half of the 448 dataset for model training (‘training dataset’). The other half was used as the ‘test dataset’ to validate of model. We considered four different prediction modeds to prediction the occurrence of frost: ANN, Support Vector Machine (SVM), Random Forest (RF), Ensemble. We used the statistical program R along with the package ‘randomForest’ for RF model, 'nnet' for ANN model (Fig. 1) and ‘kernlab’ for SVM.

* Correspondence to : cyberdoli@korea.kr

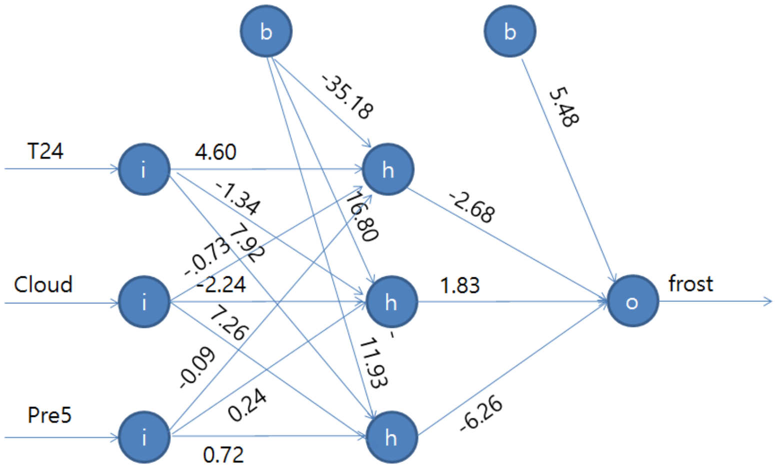

Fig. 1. Structure of ANN model.

III. Results

T24 of frost was lower than that of non-frost. and Cloud of frost was smaller than that of non-frost. but Pre5 of frost was similar that of non-frost (Table 1 and Fig. 2).

The meandecreaseGini plot shows average of all decreases in the Gini impurity in the forest that use a certain variable in the model, when values of the variable are randomly permuted. T24 was the most important variable for the RF model (Fig. 3).

And the results of the models are summarized in Table 2. Based on the percentage of model's accuracy, it can be seen that ANN model was best, followed by Ensemble.

Table 1. Statistics summary of variables

Variable	**Event**	**Statistics Summary**			
		First Quartile	**Median**	**Mean**	**Third Quartile**
T24 (°C)	Non-Frost	5.2	8.0	7.8	10.5
	Frost	2.5	4.7	4.3	6.3
Cloud	Non-Frost	2.9	4.9	4.9	6.7
	Frost	0.8	1.9	2.5	3.5
Pre5 (mm)	Non-Frost	0.0	2.0	8.4	12.1
	Frost	0.0	2.5	8.8	12.3

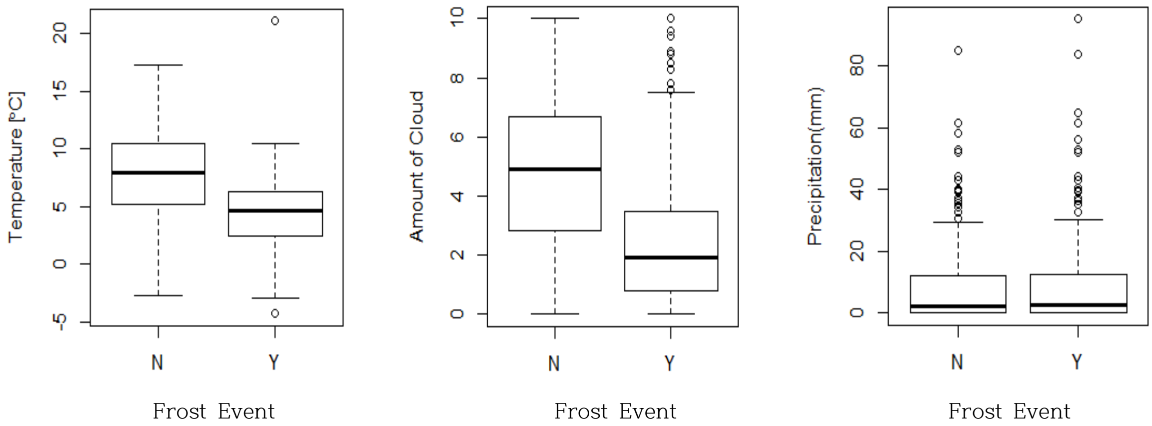

Fig. 2. Box plot of variables. (N: Non-Frost ; Y: Frost)

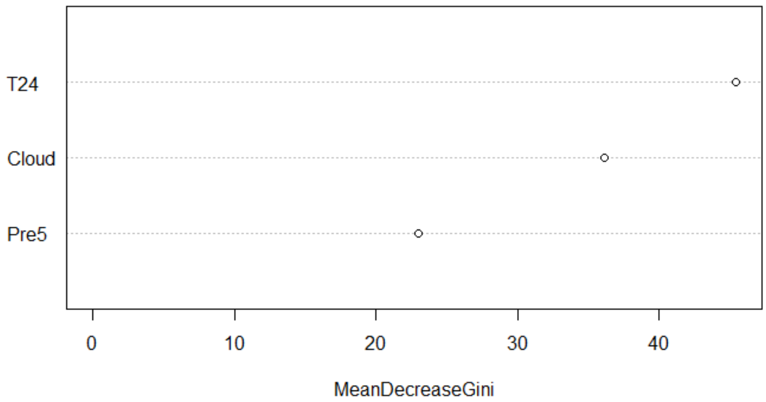

Fig. 3. The variable importance plot for RF model.

Table 2. The cross-table of matches for the prediction and observation

Observation \ Prediction	Non-Frost	Frost	Accurcy (%)
Non-Frost	81	31	
Frost	19	93	
			77.7

ANN

Observation \ Prediction	Non-Frost	Frost	Accurcy (%)
Non-Frost	77	35	
Frost	26	86	
			72.8

RF

Observation \ Prediction	Non-Frost	Frost	Accurcy (%)
Non-Frost	83	29	
Frost	37	75	
			70.5

SVM

Observation \ Prediction	Non-Frost	Frost	Accurcy (%)
Non-Frost	80	32	
Frost	21	91	
			76.3

Ensemble

Acknowledgement

This research was carried out with the support of "Research Program for Agricultural Science & Technology Development (Project No: PJ01000702)", Rural Development Administration, Republic of Korea.

References

Han, J. H., J. J. Choi, U. Chung, K. S. Cho, and J. P. Chun, 2009: Frostfall forecasting in the naju pear production area based on discriminant analysis of climatic data. *Korean Joural of Agricultural and Forest Meteorology* **11**(4), 135-142.

Mort, N., and C. Robinson, 1996: A neural network system for the protection of citrus crops from frost damage. *Computer and Electronics in Agriculture* **16**(3): 177-187.

Verdes, P. F., P. M. Granitto, H. D. Navone, and H. A. Ceccatto, 2000: Frost prediction with machine learning techniques. *Argentine Congress on Computer Science*, 1423-1433.