



## **Development of Prediction Model for The Occurrence of Heavy Snow Using Statistical**

Jin Jeon, Jeong-Eun park, and Hye-jin Kee

Gwangju Regional Meteorological Administration, Forecast Division , Korea, Republic Of (swimpyo@korea.kr)

In this study, we have developed a model that predicts the occurrence of heavy snow by using statistical methods that have recently been spotlighted. Statistical methods used in model development are Logistic Regression, Random Forest, and Artificial Neural Network.

Logistic Regression analysis is a method of classifying two groups using the relationship between dependent and independent variables when the dependent variable has two categories. Random Forest creates multiple decision trees and votes to determine the outcome of a majority decision. Artificial neural networks are machine learning algorithms that mathematically model brain learning methods. In this study, single-hidden-layer neural network(NNET in R) are used.

First, we developed a prediction model for the occurrence of heavy snow using one statistical method. As a result, the critical success index of each model was 0.48. Then, an ensemble technique was applied. The ensemble technique employed a method of using machine learning to select one of three models. The critical success index of the ensemble model is 0.54, which is more predictive than the individual model.

As well as, in order to verify the statistical model developed by applying the ensemble technique, we compared it with the numerical model through case studies. As a result, it can be confirmed that the ensemble statistical model predicted to be closer to the real situation when the numerical model does not anticipate, or wrongly predicted.

**Keyword:** Logistic Regression, Random Forest, Artificial Neural Network, Prediction model for the occurrence of heavy snow