

Comparison of three machine learning techniques for detecting tropical cyclone genesis from WindSat

Min Sang Kim¹, Myung-Sook Park¹, and Myong-In Lee¹
Jungho Im²

¹ *Climate-Environment Modeling Laboratory, Ulsan National Institute of Science and Technology, Ulsan, South Korea*

² *Intelligent Remote Sensing and Geospatial Information System, Ulsan National Institute of Science and Technology, Ulsan, South Korea*

The WindSat polarimetric radiometer retrievals on 10-m wind were used to develop and evaluate an objective algorithm for detecting tropical cyclone genesis in the Western North Pacific. Three machine learning techniques were tested – decision trees (DT), random forest (RF), and support vector machines (SVM) – using the data during 2005 – 2009. A variety of indices were invented and tested, which were to quantitatively describe the potential of intensification of tropical cyclones – wind intensity, wind pattern, wind organization, rain intensity, and rain organization. The indices related to intensity used the average and weighted-average values of wind over the area of wind speed ≥ 15 m/s and rain rate ≥ 5 mm/day as a threshold. The wind pattern index utilized the circular variance, which provides a good measure of the symmetry in the low-level circulation. The FRGSTATS programs were used to compute the degree of organization in wind and rainfall distribution. To train and validate each machine learning technique, a total of 1325 samples (352 developing and 973 nondeveloping) were selected, where the 2005-2007 year samples (206 developing and 519 nondeveloping) were used for training and the remaining samples (146 developing and 454 nondeveloping) were used for validation. The DT method showed the highest performance of 95.3% hit rate (HR) and 28.4% of false alarm rate (FAR), RF produced 88.3% of HR and 10.3% of FAR, and SVM produced 99.2% of HR and 18.2% of FAR, respectively. The lead times of tropical depression based on DT, RF, and SVM were -0.3h, 11.4h, and 1.0h after genesis.

Key words: Tropical cyclone genesis, circular variance, FRAGSTATS, machine learning, WindSat