Regression-Based Modeling of the Relationship between Climate Indices to Predict Tea Yield in Sri Lanka

Jayatilake M.D.C.D.S.[1][*], and Rankothge W.[2]

[1]Faculty of Geomatics, Sabaragamuwa University of Sri Lanka, Sri Lanka
[2]Institute of Information Technology, Sri Lanka
[*]dilshan@geo.sab.ac.lk

Research indicates that the regression-based tea–weather prediction models for the tea production in Sri Lanka, based on climate parameters namely, rainfall, relative humidity, minimum, maximum temperature, average wind speed, and sunshine hours. Agro climate geographical regions of UVA province, which contributes more of the country's tea production, are used for this research. The significance of climate parameters on tea production was explored using the random forest algorithm, determining each variable's importance. The results indicate that the minimum relative humidity, rainfall and the maximum temperature during the tea plantation period are the most influential climate indices. Machine learning implementations of the Random Forest (RF), Linear Regression (LR), Multiple Linear Regression (MLR), and Support Vector Machine (SVM) were applied for the tea prediction model. According to the results, RF is the most reliable and accurate model for the prediction of tea production in Sri Lanka. UVA province prediction model accuracy is 88.79% of the eight agro climate districts and region-wise prediction tea-production model accuracy is low parentage of the results. Further MLR and SVM, Machine Learning implementation trained and validated for the same dataset and although the results were low percentage compared to the RF implementation model accuracy. The research, regression analysis already applied for RF, SVM and LR for the region-wise of the UVA province. Final outcome of the results indicates that the same process can be applied to the Tea- weather prediction model for all of the tea growing areas in the country.

Keywords: *Agro climate, Regression model, Random forest, Tea production, Tea-weather*