



Estimating Gross Primary Production using machine-learning algorithms based on eddy covariance measurements and remote sensing in forest ecosystem

Bora Lee, Eunsook Kim, Keunchang Jang, Nang Hyun Cho, Wookyung Song, Goeun Park, Chanwoo Park, and Jong-Hwan Lim

National Institute of Forest Science, Forest Science and Climate Change Division, Seoul, Korea, Republic Of

Gross primary production (GPP) is the amount of total carbon assimilated in gross carbon uptake or photosynthesis of chlorophyll containing organs. Forest ecosystems have an important role to determine global GPP because forests are the largest part of global carbon flux among terrestrial ecosystems. There have been several attempts to estimate GPP using mechanism-based models. However, mechanism-based models including biological, chemical, and physical processes are limited due to a lack of flexibility in predicting non-stationary ecological processes, which are caused by a local and global change. Mechanism-free methods are instead strongly recommended to estimate nonlinear dynamics that occur in nature like GPP. Therefore, we used the mechanism-free machine learning techniques to estimate the daily GPP. In this study, artificial neural network (ANN), Support vector machine (SVM), and random forest (RF) were developed and compared with the traditional multiple linear regression model (MLR). MODIS products and meteorological parameters from Eddy covariance data were employed to train the machine learning and MLR models from 2006 to 2013. GPP prediction models were compared with daily GPP from eddy covariance measurement in a deciduous forest in South Korea in 2014 and 2015. Statistical analysis including correlation coefficient (R), root mean square error (RMSE) and mean squared error (MSE) were used to evaluate the performance of models. In general, the models from machine-learning algorithms ($R = 0.85 - 0.93$, $p < 0.001$) were showed better performance than linear regression model ($R = 0.82 - 0.92$, $p < 0.001$). These results provide insight into high predictability and the possibility of expansion through the use of the mechanism-free machine-learning models and remote sensing for predicting non-stationary ecological processes such as seasonal GPP.