Journal of Integrative Agriculture 2016, 15(0): 60345-7

Available online at www.sciencedirect.com

ScienceDirect

RESEARCH ARTICLE

Estimating grassland LAI using the random forests approach and Landsat imagery in the meadow steppe of Hulunber, China

LI Zhen-wang, XIN Xiao-ping, TANG Huan, YANG Fan, CHEN Bao-rui, ZHANG Bao-hui

National Hulunber Grassland Ecosystem Observation and Research Station, Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, P.R.China

Abstract

Leaf area index (LAI) is a key parameter for describing vegetation structures and is closely associated with vegetative photosynthesis and energy balance. The accurate retrieval of LAI is important when modeling biophysical processes of vegetation and the productivity of earth systems. The random forests (RF) method aggregates an ensemble of decision trees to improve the prediction accuracy and demonstrates a more robust capacity than other regression methods. This study evaluated the RF method for predicting grassland LAI using ground measurements and remote sensing data. Parameter optimization and variable reduction were conducted before model prediction. Two variable reduction methods were examined: the variable importance value method and the principal component analysis (PCA) method. Finally, the sensitivity of RF to highly correlated variables was tested. The results showed that the RF parameters have a small effect on the performance of RF, and a satisfactory prediction was acquired with a root mean square error (RMSE) of 0.1956. The two variable reduction methods for RF prediction produced different results; variable reduction based on the variable importance value method achieved nearly the same prediction accuracy with no reduced prediction, whereas variable reduction using the PCA method had an obviously degraded result that may have been caused by the loss of subtle variations and the fusion of noise information. After removing highly correlated variables, the relative variable importance remained steady, and the use of variables selected based on the best-performing vegetation indices performed better than the variables with all vegetation indices or those selected based on the most important one. The results in this study demonstrate the practical and powerful ability of the RF method in predicting grassland LAI, which can also be applied to the estimation of other vegetation traits as an alternative to conventional empirical regression models and the selection of relevant variables used in ecological models.

Keywords: leaf area index, random forests grassland, remote sensing, Hulunber

1. Introduction

Grassland covers approximately one third of the global terrestrial surface (Fan *et al.* 2003; Lemaire *et al.* 2005), and plays an important role in the interactions among earth's atmosphere, hydrosphere and continental surface. The study of grassland is necessary to understand global climatic change and terrestrial carbon cycling (Lemaire *et al.* 2005). Leaf area index (LAI) is a key parameter for describing veg-

Received 15 October, 2015 Accepted 20 January, 2016
LI Zhen-wang, Tel: +86-10-82109618, E-mail: lizhenwang10@hotmail.com; Correspondence ZHANG Bao-hui, Tel: +86-10-82109618, E-mail: zhangbaohui@caas.cn

doi: 10.1016/S2095-3119(15)61303-X