

Global Land Surface Evapotranspiration Estimation From Meteorological and Satellite Data Using the Support Vector Machine and Semiempirical Algorithm

Meng Liu, Ronglin Tang, Zhao-Liang Li, *Senior Member, IEEE*, Yunjun Yao, and Guangjian Yan

Abstract—Evapotranspiration (ET) is the combination process of the surface evaporation and plant transpiration, which occur simultaneously, and it links the terrestrial water cycles, carbon cycles, and energy exchange. In this study, based on the observations from 242 global FLUXnet sites, with daily mean temperature, relative humidity, net radiation, wind speed, incoming shortwave radiation, maximum temperature, minimum temperature, normalized difference vegetation index, altitude, difference in temperature, and observed ET as input data, we used a support vector machine and a semiempirical algorithm to estimate the land surface daily ET at nine different vegetation-type sites. Subsequently, based on the meteorological reanalysis data combined with remote sensing data, we estimated regional land surface ET of China during 1982–2010. The results showed that, for all vegetation-type sites, when the predicted ET was validated with the eddy covariance measurements, the support vector machine algorithm undervalued ET while the semiempirical algorithm overvalued ET. When five indicators and the second classification method were selected, the semiempirical algorithm probably could explain 56%–76% of the land surface ET change, whereas the support vector machine algorithm probably could explain 71%–85%. The regional values of annual daily average ET varied from 5.8 to 110.5 W/m², and the land surface ET overall trend decreased from the southeast to the northwest in China.

Index Terms—Evapotranspiration (ET), support vector machine (SVM).

Manuscript received September 15, 2016; revised June 30, 2017; accepted December 22, 2017. This work was supported in part by the National Natural Science Foundation of China under Grant 41571351, Grant 41571367, Grant 41401659, and in part by the International Science and Technology Cooperation Program of China under Grant 2014DFE10220. (*Corresponding author: Ronglin Tang.*)

M. Liu and R. L. Tang are with the State Key Laboratory of Resources and Environment Information System, Institute of Geographic Science and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China, and also with the University of Chinese Academy of Sciences, Beijing 100049, China (e-mail: xinyin_liumeng@163.com; trl_wd@163.com).

Z.-L. Li is with the Key Laboratory of Agri-Informatics, Ministry of Agriculture/Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, China, and also with the State Key Laboratory of Resources and Environment Information System, Institute of Geographic Science and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China (e-mail: lizhaoliang@caas.cn).

Y. Yao and G. Yan are with the State Key Laboratory of Remote Sensing Science, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China (e-mail: boyyunjun@163.com; gjyan@bnu.edu.cn).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JSTARS.2017.2788462

I. INTRODUCTION

EVAPOTRANSPIRATION (ET) is the combination process of the surface evaporation and plant transpiration, which occur simultaneously, and is defined as the process of liquid water contained in land surface being converted to water vapor in the atmosphere [1]. It links the terrestrial water cycles, carbon cycles, and energy exchange, and it is the main process parameter of water and energy exchange in hydrosphere, atmosphere, and biosphere [2], [3]. It is an important factor in heat balance and water balance of land, lake, or river surface and is central to earth system science. Nowadays, lots of explorations and research work have been carried out about the ET algorithm and regional application [4]–[6]. Traditional methods that use satellite remote sensing data to estimate land surface ET include the one-source models [7]–[9] and two-source models [10] which are based on surface energy balance, the soil temperature–vegetation index methods [11], the Penman–Monteith models [12], [13], the empirical models [14], and the data assimilation methods [15], [16]. Based on the existing research work, many traditional ET estimation algorithms were based on physical models, which have large regional differences, so coefficient calibrations are required in different regions. In recent years, with the continuous development of computer technology, some machine learning models appear. Due to their advantages in exploring the complex nonlinear relationship between different factors, machine learning models have been widely used in hydrological forecasting, groundwater research, and other fields. ET can be influenced by many factors, and its estimation can be seen as the nonlinear relationship regression analysis between meteorological factors and other factors. Several machine algorithms have been used to estimate ET or reference ET in previous studies [17]–[22]. These studies used fewer sites (several or more than ten) observations to estimate ET at site scale and validate the accuracy of the algorithm in relatively small areas. These studies have not obtained universal models to estimate ET at global site scale or in regional large scale and mesoscale and have not produced large regional ET in a country or continent. Support vector machine (SVM) was a widely used machine learning model in some areas, such as meteorological forecast and hydrological assessment [23]. In this paper, taking its advantages on regression into consideration, an SVM was