Misestimation of Growing Season Length Due to Inaccurate Construction of Satellite Vegetation Index Time Series

Cong Wang and Kai Zhu

Abstract—**Satellite-based vegetation index (VI) is widely used in monitoring land surface phenology (LSP). Currently, well-developed VI products utilize the maximum value composite (MVC) algorithm to produce a composite VI time series (TS). Many of these products, however, lack the actual acquisition date (AD) of the VI value. As an alternative, the median or mean date of a composite period is used to reconstruct the VI TS, which might lead to bias on LSP detection. This letter quantifies the LSP bias in the Northern Hemisphere by generating a 15-day composited normalized difference vegetation index (NDVI) TS from the land long-term data record daily NDVI products using the MVC method. The results show that the AD of the NDVI value is usually later than the mean date of a composite period in spring and earlier in fall, effectively leading to a total overestimation of the growing season length of 5.91 days on average across the Northern Hemisphere (north of 30° N). This bias has a significant spatial pattern with high values observed in Northeastern China, Central North America, and high-latitude areas. However, the temporal trend is not largely influenced overall. Accordingly, we suggest the research community using accurate temporal information, whenever possible, in extracting LSP from VI TS.**

Index Terms—**End of the growing season (EOS), land long-term data record normalized difference vegetation index (LTDR NDVI), land surface phenology (LSP), maximum value composite (MVC), start of the growing season (SOS).**

I. Introduction

CHANGE in land surface phenology (LSP) is one of the fingerprints of ecological responses to climate change [1], [2]. Mounting literature has reported the shift of vegetation phenology worldwide with advancing spring events and extending growing seasons due to anthropogenic global warming [1], [3], [4].

To monitor this change, satellite remote sensing provides a unique approach and has been widely applied in phenological research at large spatial scales in recent decades [5]. Conventionally, time series (TS) of satellite-based vegetation index (VI), such as the well-known normalized difference vegetation index (NDVI) and the newly developed normalized difference phenology index [6] are used to retrieve the LSP. Compared with field observations, remote sensing data have much greater spatial coverage, but the data quality is limited by a variety of noise sources such as cloud, snow, and orbital drift, which all impose uncertainty on the LSP estimate [7].

Most of the current VI products have a coarse temporal frequency. For instance, Global Inventory Modeling and Mapping Studies (GIMMS) NDVI data set is half-month composited, and Moderate Resolution Imaging Spectroradiometer (MODIS) VI products provide both 8-day and 16-day composited VI TS. Daily NDVI data are generally sampled to a coarse temporal resolution using the maximum value composite (MVC) algorithm, such as for the GIMMS NDVI [8] and VGT 10S products [9]. This procedure is accomplished by selecting the highest NDVI value as the output of a composite period. Because MVC favors clear sky pixels, the output NDVI is less affected by the atmosphere and cloud cover.

The application of the MVC-based temporal composited VI products could potentially impose a substantial bias in LSP detection. To construct an accurate annual TS of NDVI, researchers need not only the NDVI value but also the exact actual acquisition date (AD) of the value. The AD of NDVI, although not essential for vegetation biomass estimation, is the key information for LSP detection which focuses on the temporal transition of vegetation status. Some MODIS products include the AD and have been utilized to rebuild the VI TS in [6]. However, information of AD is lacking in many VI products, especially the AVHRR-based data sets. A common way to construct the VI TS when using these data sets is to choose a reference date (RD), such as the median or mean date of a composite period [10]. It will result in a bias in the LSP detection [11], and importantly, the bias is nonstochastic. In most regions, NDVI increases rapidly in spring with the vegetation growth and decreases in fall with the leaf senescence. Considering that MVC selects the highest NDVI value during a composite period, pixels of late date in spring and pixels of early date in fall are more likely to be selected, which means that the actual VI TS may be "thinner" than the estimated NDVI TS using the median or mean date as the x-axis (Fig. 1). Therefore, the growing season length (GSL) may be overestimated with an earlier start of the growing season (SOS) and a later end of the growing season (EOS) when using these MVC-based VI products.

Manuscript received November 11, 2018; revised January 22, 2019; accepted January 23, 2019. Date of publication February 12, 2019; date of current version July 18, 2019. This work was supported by the Faculty Research Grant awarded by the Committee on Research from the University of California, Santa Cruz. *(Corresponding author: Kai Zhu.)*

C. Wang was with the Department of Environmental Studies, University of California at Santa Cruz, Santa Cruz, CA 95064 USA. He is now with the University of Illinois at Urbana–Champaign, IL 61820 USA (e-mail: wangcongrs@gmail.com).

K. Zhu is with the Department of Environmental Studies, University of California at Santa Cruz, Santa Cruz, CA 95064 USA (e-mail: kai.zhu@ucsc.edu).

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

Digital Object Identifier 10.1109/LGRS.2019.2895805