Remote Sensing of Environment 206 (2018) 403–423

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

An integrated model for generating hourly Landsat-like land surface temperatures over heterogeneous landscapes

Jinling Quan[a,b,*], Wenfeng Zhan[c], Ting Ma[a,b], Yunyan Du[a,b], Zheng Guo[d], Bangyong Qin[e]

[a] State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China
[b] College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100190, China
[c] Jiangsu Provincial Key Laboratory of Geographic Information Science and Technology, International Institute for Earth System Science, Nanjing University, Nanjing, Jiangsu 210046, China
[d] National Satellite Meteorological Center, Beijing 100081, China
[e] Key Laboratory of Space Utilization, Technology and Engineering Center for space Utilization, CAS, Beijing 100094, China

ARTICLE INFO

Keywords:
Data fusion
Land surface temperature
Landsat
MODIS
Geostationary satellite
Heterogeneity

ABSTRACT

The trade-off between spatial and temporal resolutions in remote sensing has greatly limited the availability of concurrently high spatiotemporal land surface temperature (LST) data for wide applications. Although many efforts have been made to resolve this dilemma, most have difficulties in generating diurnal fine-resolution LSTs with high spatial details for landscapes with significant heterogeneity and land cover type change. This study proposes an integrated framework to BLEnd Spatiotemporal Temperatures (termed BLEST) of Landsat, MODIS and a geostationary satellite (FY-2F) to one hour interval and 100 m resolution, where (1) a linear temperature mixing model with conversion coefficients is combined to better characterize heterogeneous landscapes and generate more accurate predictions for small and linear objects; (2) residuals are downscaled by a thin plate spline interpolator and restored to the primary fine-resolution estimations to include information about land cover type change; and (3) separate operations at annual and diurnal scales with nonlinear temperature modeling are designed to neutralize the hybrid impacts of large scale gap and land cover type change. BLEST was tested on both simulated data and actual satellite data at annual, diurnal and combined scales, and evaluations were conducted with the simulated/actual fine-resolution data, in-situ data, and with three popular fusion methods, i.e., the spatial and temporal adaptive reflectance fusion model (STARFM), the Enhanced STARFM (ESTARFM) and the spatiotemporal integrated temperature fusion model (STITFM). Results show higher accuracy by BLEST with more spatial details and pronounced temporal evolutions, particularly over heterogeneous landscapes and changing land cover types. BLEST is proposed to augment the spatiotemporal fusion system and further support diurnal dynamic studies in land surfaces.

1. Introduction

Land surface temperature (LST) is a crucial variable in surface energy processes, hydrological balance, and climate change (Li et al., 2013). It varies significantly in both space and time (Prata et al., 1995). With the advance of remote sensing, amounts of satellite-derived LST products have been publicly available for characterizing the spatiotemporal variations in LST at spatial resolutions of 60 m–10 km and temporal resolutions of 15 min–26 days (Quan et al., 2014a). However, due to limitations of techniques and budgets (Zhu et al., 2010), current instruments have yet been able to provide concurrently high spatial and temporal resolutions (Zhan et al., 2013), which has greatly constrained the potential applications of satellite-derived LSTs in various fields. For example, polar-orbiting satellite sensors (e.g., Landsat TM LST: 120 m and 16 day resolutions) probably miss the optimal observation time, particularly for rapidly changing areas, while geostationary satellite sensors (e.g., MSG SEVIRI LST: 3 km and 15 min resolutions) lose spatial details over heterogeneous landscapes (Sobrino et al., 2012).

Regarding this issue, various methods aiming for high spatiotemporal resolutions have been proposed. They can be divided into two categories: using sole sensor LST or multi-sensor LST pairs. The first category includes spatial downscaling of geostationary satellite data and temporal interpolation of polar-orbiting satellite data, while the second includes endmember unmixing-, sparse representation-, and weight function- based fusion methods for multi-resolution LST pairs at different times.

* Corresponding author at: No.11A Datun Road, Beijing 100101, China.
E-mail addresses: quanjl@lreis.ac.cn (J. Quan), zhanwenfeng@nju.edu.cn (W. Zhan), mting@lreis.ac.cn (T. Ma), duyy@lreis.ac.cn (Y. Du), guozheng@cma.gov.cn (Z. Guo), qinby@csu.ac.cn (B. Qin).

https://doi.org/10.1016/j.rse.2017.12.003
Received 23 February 2017; Received in revised form 26 October 2017; Accepted 3 December 2017

The spatial downscaling method is to downscale a coarse spatial resolution, typically for geostationary satellites, by associating with visible and near-infrared scale factors at a high spatial resolution (Zakšek and Oštir, 2012; Weng and Fu, 2014a; Sismanidis et al., 2015). It generally requires measurements of scale factors in the same day (except for the case that neglects land cover change), which cannot always be satisfied due to the long revisiting period of the scale factor sensors. Also, few studies are able to downscale the geostationary satellite LSTs to a spatial resolution finer than 1 km due to large spatial scale differences (Kumar et al., 2012). One remarkable exception is the approach proposed by Bechtel et al. (2012) that incorporates hundreds of scale factors to downscale SEVIRI LSTs to a resolution of ~100 m. The results are rather promising, but the spatial variation details are underperformed mainly because of the regression-based reconstruction. Moreover, this approach may not be well suited to wide coverage due to the large auxiliary data requirement and the spatial relationship change between LST and scale factors. Herein, we mainly focus on downscaling geostationary satellites considering their high temporal frequencies. Complete reviews of thermal downscaling for all kinds of satellites can be found in Zhan et al. (2013) and Chen et al. (2014b).

The temporal interpolation method is to interpolate temporally discrete observations, typically for polar-orbiting satellites, by associating with a diurnal/annual temperature cycle (DTC/ATC) or a surface energy balance model (Duan et al., 2012; Weng and Fu, 2014b). At a diurnal scale, it generally requires at least four LST observations a day, which can only be satisfied by a small number of MODIS observations under clear skies (Quan et al., 2014a), and the spatial resolution after interpolation is often unchanged, i.e., 1 km (Duan et al., 2012). At an annual scale, many details of day-by-day variations are lost and diurnal LST dynamics are unresolved. It should be mentioned that Zhan et al. (2016) proposed a method combining spatial downscaling and temporal interpolation, i.e., downscaling the DTC and ATC parameters using scale factors, such as vegetation index and albedo, to estimate diurnal 250 m-resolution LSTs throughout a year. Nevertheless, problems remain, such as the minimum four-observations/day requirement and loss of spatial and day-by-day variation details.

The endmember unmixing-based method regards temporal variations at a coarse spatial resolution as the mixture of component temporal variations at a fine spatial resolution (Wu et al., 2012; Amorós-López et al., 2013; Gevaert and García-Haro, 2015; Zhu et al., 2016). The fine-scale temporal variations are then derived assuming negligible intra-component variance and constant component fractions. It can be regarded as an unmixing-based downscaling method (Zhukov et al., 1999) extended from one sensor and one time to multiple sensors and different times. Its good performance is highly dependent on accurate fraction/classification maps (Zhang et al., 2015) which may not always be available (Zhu et al., 2016), and its spatial pattern reconstruction may be unsatisfactory in the case of a large scale difference (Quan et al., 2013).

The sparse representation-based method learns the correspondence between structures within coarse-fine resolution image pairs by means of sparse representation, and applies that correspondence to the coarse resolution at another time to predict the fine-resolution values (Aharon et al., 2006; Huang and Song, 2012; Song and Huang, 2013; Chen et al., 2016). It can be regarded as a more sophisticated version (accounting for structural features, including changes in phenology and land cover types) of regression-based downscaling, where the scale factors are replaced by the target variable (e.g., LST or LST dictionary), and the relationship is learned according to patches rather than pixels and applied to the coarse resolution at a different time. The difficulties of this method include the physical explanation of the correspondence, strong reliance on the image extent, good preservation of spatial details at a large scale difference, and complex model computation (Zhang et al., 2015; Zhu et al., 2016). Furthermore, neither of the aforementioned endmember unmixing- and sparse representation- based methods has been widely tested on LST data and therefore their feasibility for LST can hardly be asserted.

The weight function-based method has gained popularity since it was first proposed. The most representative is the spatial and temporal adaptive reflectance fusion model (i.e., STARFM by Gao et al. (2006)), which estimates daily Landsat-like reflectance by a weighted sum of reflectance changes in similar neighboring pixels assuming scale invariance of temporal change (Zhang et al., 2015). Regarding its low accuracy over landscapes with mixed and changing land cover types, Hilker et al. (2009) proposed a spatial temporal adaptive algorithm for mapping reflectance change (STAARCH) over vegetated surfaces by detecting changes and selecting optimal reference images, while Zhu et al. (2010) proposed an enhanced STARFM (ESTARFM) by combining a linear spectral mixing model to account for complex surface heterogeneity.

Although the STARFM-like methods were initially designed for reflectance, they have proved useful for generating daily Landsat-like thermal data such as evapotranspiration and LST (Anderson et al., 2012; Kim and Hogue, 2012; Huang et al., 2013a; Weng et al., 2014; Wu et al., 2015a; Shen et al., 2016a). The modifications mainly focus on the weight function design, temporal change modeling or thermal landscape representation to better restore the spatiotemporal LST patterns. To further compensate for the inability to generate diurnal Landsat-like LSTs, Wu et al. (2015b) integrated geostationary satellite LSTs (i.e., SEVIRI and GOES) with Landsat and MODIS LSTs, termed the spatiotemporal integrated temperature fusion model (STITFM). However, similarly to STARFM, STITFM experiences problems with complex surface heterogeneity and land cover type change.

Most methods reported in the literature were developed independently focusing one type of resolution, while a few studies formulated integrated frameworks for fusing multiple spatio–temporal–spectral resolutions, such as Huang et al. (2013b), Meng et al. (2015), and Shen et al. (2016b). Nonetheless, they have mainly focused on optical images and rarely been tested on LST data.

This study aims to integrate the main features and advantages of existing methods to develop an advanced spatiotemporal fusion model for LSTs. Specifically, we blend geostationary satellite (i.e., FY-2F), MODIS, and Landsat thermal observations to obtain diurnal ~100 m-resolution LSTs; adopt the weight functions to better preserve spatial details; use the ATC and DTC models to characterize nonlinear temporal patterns; combine a linear temperature mixing model (LTMM) and thin plate spline (TPS) downscaling (Chen et al., 2014a) to account for the impacts of landscape heterogeneity and land cover type change; and perform two (i.e., annual and diurnal) steps to bridge the large scale gap. We term this integrated framework the BLEnding of Spatiotemporal Temperatures (BLEST). BLEST is further described in Section 2. It is tested and evaluated using simulated data, actual satellite data, and filed observations in comparison with STARFM, ESTARFM, and STITFM (Sections 4 and 5). Discussion and conclusions are provided in Sections 5 and 6.

2. Method

Before presenting the details of the proposed method (BLEST), important variables and definitions are provided as follows for convenience and clarity.

(I) i	(total number of) components
j	prediction/central pixel
(K) k	(total number of) pixels
(S) s	(total number of) similar pixels
$F/M/C$	fine/medium/coarse spatial scale
d/t	day of year/time of day
l/n and m	subscripts of observation day and prediction day
o/q and p	subscripts of observation time and prediction time
T	LST
R	residual
$f/\dot{f}$	component fraction at a medium/coarse spatial scale
v	conversion coefficient
w	weight