Remote Sensing of Environment 213 (2018) 162–181

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Spatial-temporal fraction map fusion with multi-scale remotely sensed images

Yihang Zhang[a], Giles M. Foody[b], Feng Ling[a,*], Xiaodong Li[a], Yong Ge[c], Yun Du[a], Peter M. Atkinson[d]

[a] *Key Laboratory of Monitoring and Estimate for Environment and Disaster of Hubei Province, Institute of Geodesy and Geophysics, Chinese Academy of Sciences, Wuhan 430077, China*
[b] *School of Geography, University of Nottingham, University Park, Nottingham NG7 2RD, UK*
[c] *State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences & Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China*
[d] *Lancaster Environment Center, Faculty of Science and Technology, Lancaster University, Lancaster LA1 4YQ, UK*

ARTICLE INFO

Keywords:
Land cover
Fraction maps
Spatial-temporal fusion
Spectral unmixing
Super-resolution mapping

ABSTRACT

Given the common trade-off between the spatial and temporal resolutions of current satellite sensors, spatial-temporal data fusion methods could be applied to produce fused remotely sensed data with synthetic fine spatial resolution (FR) and high repeat frequency. Such fused data are required to provide a comprehensive understanding of Earth's surface land cover dynamics. In this research, a novel Spatial-Temporal Fraction Map Fusion (STFMF) model is proposed to produce a series of fine-spatial-temporal-resolution land cover fraction maps by fusing coarse-spatial-fine-temporal and fine-spatial-coarse-temporal fraction maps, which may be generated from multi-scale remotely sensed images. The STFMF has two main stages. First, FR fraction change maps are generated using kernel ridge regression. Second, a FR fraction map for the date of prediction is predicted using a temporal-weighted fusion model. In comparison to two established spatial-temporal fusion methods of spatial-temporal super-resolution land cover mapping model and spatial-temporal image reflectance fusion model, STFMF holds the following characteristics and advantages: (1) it takes account of the mixed pixel problem in FR remotely sensed images; (2) it directly uses the fraction maps as input, which could be generated from a range of satellite images or other suitable data sources; (3) it focuses on the estimation of fraction changes happened through time and can predict the land cover change more accurately. Experiments using synthetic multi-scale fraction maps simulated from Google Earth images, as well as synthetic and real MODIS-Landsat images were undertaken to test the performance of the proposed STFMF approach against two benchmark spatial-temporal reflectance fusion methods: the Enhanced Spatial and Temporal Adaptive Reflectance Fusion Model (ESTARFM) and the Flexible Spatiotemporal Data Fusion (FSDAF) model. In both visual and quantitative evaluations, STFMF was able to generate more accurate FR fraction maps and provide more spatial detail than ESTARFM and FSDAF, particularly in areas with substantial land cover changes. STFMF has great potential to produce accurate time-series fraction maps with fine-spatial-temporal-resolution that can support studies of land cover dynamics at the sub-pixel scale.

1. Introduction

With the capabilities of broad spatial coverage and temporally repeated imaging from Earth observation sensors, remote sensing has considerable potential to provide time-series satellite images for studying land surface dynamics (Townshend et al., 1991; Yang and Lo, 2002). In heterogeneous areas, land surface dynamics, such as urban expansion, flooding and deforestation, often occur at a fine spatial scale and within a short period. It is, therefore, necessary to collect fine-spatial-temporal-resolution remote sensing images to monitor fine scale land cover changes in a timely manner. Due to the common trade-off between the spatial resolution and the temporal repeat frequency of satellite sensing systems, there is so far no single satellite sensor that can provide remote sensing images with both fine spatial and temporal resolutions (Gao et al., 2006; Li et al., 2017; Zhu et al., 2016). Generally, fine spatial resolution (FR) satellite images are acquired infrequently and have a relatively coarse temporal resolution, making it hard to monitor rapid land cover changes. On the contrary, coarse

---

* Corresponding author.
E-mail address: lingf@whigg.ac.cn (F. Ling).

https://doi.org/10.1016/j.rse.2018.05.010
Received 12 October 2017; Received in revised form 1 May 2018; Accepted 4 May 2018

spatial resolution (CR) satellite sensors acquire data with a high repeat frequency. However, their spatial resolutions are often too coarse to allow the detection of land cover changes occurring in small areas. Therefore, to deal with this dilemma, methods for spatial-temporal data fusion are highly desirable for application to both kinds of remotely sensed imagery to provide remote sensing data with fine spatial and temporal resolutions for studying land surface dynamics (Gao et al., 2006; Gong et al., 2013; Hansen and Loveland, 2012; Li et al., 2015; Ling et al., 2016; Ling et al., 2011; Zhu and Woodcock, 2014).

Recently, the spatial-temporal super-resolution mapping (STSRM) method proposed by Ling et al. (2011) has become a promising spatial-temporal fusion method to extract fine spatial and temporal resolution land cover change information (Li et al., 2016; Ling et al., 2016; Wang et al., 2015; Wu et al., 2017; Xu et al., 2017). STSRM aims to predict a FR land cover map from CR fraction maps, assuming that another FR land cover map, acquired at previous time for the same area, is available. STSRM can be considered as an extension of the traditional super-resolution mapping approach applied to a mono-temporal image, by incorporating information about the land cover changes through time. The key of STSRM is the multi-scale land cover change principle that is using coarse-to-fine resolution change detection between current CR fraction maps and previous FR land cover map to predict the potential locations of current land cover labels of FR land cover map (Ling et al., 2011). The multi-scale land cover change principle in STSRM was further analyzed and assessed by using existing land cover maps, and it has been demonstrated consistently that the principle could be suitable for most current satellite sensors (Ling et al., 2016). Some popular super-resolution mapping algorithms applied on mono-temporal remote sensing images were also extended to the spatial-temporal domain, leading to various STSRM models (He et al., 2016; Li et al., 2015; Li et al., 2017; Wang et al., 2016; Xu and Huang, 2014; Zhang et al., 2017). Compared with the traditional super-resolution mapping methods applied to mono-temporal remote sensing imagery, STSRM can provide details about the spatial distribution of different land cover classes and their changes over time. It is a promising means to produce fine spatial and temporal resolution land cover maps from multi-scale remote sensing imagery.

It is noteworthy that in all existing STSRM models the FR pixels are treated as pure units. That is, the fine pixels within the input and the resultant FR land cover maps are all considered as pure pixels, and each of them is labeled as representing an area comprised of one and only one land cover class. This assumption is reasonable in some cases because the proportion of mixed pixels in an image is typically positively related to pixel size. However, the limitation of this assumption is also obvious, as mixing may still exist in FR image pixels, especially if the land cover mosaic is highly fragmented and heterogeneous.

In practice, the satellite sensor's instantaneous field-of-view often includes more than one land cover feature irrespective of the scale of measurement. Indeed, the mixed pixel problem is widely observed in remote sensing images across different spatial scales (Keshava and Mustard, 2002). It is well known that CR remote sensing data, such as those obtained from the Advanced Very High Resolution Radiometer (AVHRR), MEdium Resolution Imaging Spectrometer (MERIS) and MODerate resolution Imaging Spectroradiometer (MODIS) images, contain a large number of mixed pixels. However, the mixed pixel problem is also evident in medium and high spatial resolution satellite sensor images, such as Landsat (Lu and Weng, 2004; Powell et al., 2007), ASTER (Weng et al., 2009), IKONOS (Lu and Weng, 2009) and Quickbird (Lu et al., 2010), and spectral unmixing techniques may still be needed to obtain fraction maps to enhance the representation of land cover. In this situation, the assumption that all FR pixels are pure in STSRM models may be unreasonable in some real applications.

Another limitation of using the pure pixel assumption in STSRM model is that land cover change information used by it may be partial and possibly erroneous. With the assumption, only one land cover class can be associated with a pixel and hence the only change that can be characterized is that it represents a complete alteration in land cover class: a land cover conversion (e.g. a change from forest to grassland). However, many important land cover changes happed at the sub-pixel scale (finer than the spatial resolution of pixel) may not involve a change in class label. For example, a pixel may represent a forested region which may undergo a substantial change such as a major reduction in tree cover and yet still remain classed as a forest. Changes of the latter type, therefore, do not involve a change in label but a change in the character of the land cover: a land cover modification. Land cover modifications cannot be studied using methods that assume pure pixels but they, and the land cover conversions, can be studied if mixed pixels are allowed such as via the application of soft classification techniques (Foody, 2001).

Given the two limitations arising from the pure pixel assumption, error and uncertainty could be introduced in the resultant fine spatial and temporal resolution land cover maps produced by STSRM. Since land cover class fraction values produced by unmixing or soft classification analyses can be used to obtain more accurate land cover information at the sub-pixel scale than discrete land cover labels produced by hard classification (Foody, 2002; Foody and Doan, 2007), they may have a potential role to play in increasing the accuracy of the STSRM approach.

A different approach to the STSRM for fusing fine-spatial-coarse-temporal and coarse-spatial-fine-temporal remotely sensed images is the spatial-temporal reflectance fusion model. Unlike the STSRM approach that aims to predict land cover class labels at a fine resolution, the spatial-temporal reflectance fusion approach is used to blend reflectance values of remotely sensed images. Gao et al. (2006) first proposed the spatial and temporal adaptive reflectance fusion model (STARFM) to blend Landsat and MODIS reflectance images and produce daily 30 m synthetic Landsat-like reflectance images. Hilker et al. (2009) developed a spatial and temporal adaptive fusion model (STAARCH) to explore spatio-temporal pattern details of forest disturbance based on Landsat and MODIS images. Thereafter, STARFM was developed as an enhanced spatial-temporal adaptive reflectance fusion model (ESTARFM) (Zhu et al., 2010) and a flexible spatio-temporal data fusion (FSDAF) model (Zhu et al., 2016). Moreover, other image spatial temporal fusion models, such as the unmixing based fusion model (Gevaert and Garcia-Haro, 2015; Zhukov et al., 1999; Zurita-Milla et al., 2008), the sparse representation based fusion model (Huang and Song, 2012; Song and Huang, 2013) and spatial and temporal reflectance fusion considering the sensor difference (Shen et al., 2013), have also been proposed. Once the fine spatial and temporal resolution remote sensing images have been produced by the spatial-temporal reflectance image fusion method, a spectral unmixing approach can then be used to produce the corresponding fine spatial and temporal fraction maps. The effectiveness of this approach, however, depends greatly on the spatial-temporal reflectance fusion method, which often suffers from two major limitations when the final objective is to produce fraction maps. First, most spatial-temporal reflectance fusion methods do not account for land cover changes that may have occurred within the period represented by the time-series of remotely sensed images (Gevaert and Garcia-Haro, 2015; Zhu et al., 2016). Second, spatial-temporal reflectance fusion methods can generally deal with image pairs with similar spectral bands. Given that many satellite sensors produce images with unique spectral bands, the range of application of these spatial-temporal reflectance fusion methods is thus limited. In comparison, STSRM-based approaches are free from the assumption of sensor-based coherence and can accommodate information on class label change, but not the land cover fraction changes.

In this paper, a novel Spatial-Temporal Fraction Map Fusion (STFMF) model is proposed to generate fraction maps that have a fine resolution in both the spatial and temporal domains by fusing coarse-spatial-fine-temporal and fine-spatial-coarse-temporal remotely sensed images. Critically, the STFMF approach addresses limitations of other methods and hence forms an important contribution to the realization