>TGRS-2015-00405<

The following publication Wang, Q., Shi, W., & Atkinson, P. M. (2016). Spatiotemporal subpixel mapping of time-series images. IEEE Transactions on Geoscience and Remote Sensing, 54(9), 5397-5411 is available at https://doi.org/10.1109/TGRS.2016.2562178.

Spatio-temporal Sub-pixel Mapping of Time-series Images

Qunming Wang, *Member, IEEE,* Wenzhong Shi, and Peter M. Atkinson

***Abstract*—Land-cover/land-use (LCLU) information extraction from multi-temporal sequences of remote sensing imagery is becoming increasingly important. Mixed pixels are a common problem in Landsat and MODIS images that are used widely for LCLU monitoring. Recently developed sub-pixel mapping (SPM) techniques can extract LCLU information at the sub-pixel level by dividing mixed pixels into sub-pixels to which hard classes are then allocated. However, SPM has rarely been studied for time-series images (TSIs). In this paper, a spatio-temporal SPM approach was proposed for SPM of TSIs. In contrast to conventional spatial dependence-based SPM methods, the proposed approach considers simultaneously spatial and temporal dependences, with the former considering the correlation of sub-pixel classes *within* each image and the latter considering the correlation of sub-pixel classes *between* images in a temporal sequence. The proposed approach was developed assuming the availability of one fine spatial resolution map which exists amongst the TSIs. The SPM of TSIs is formulated as a constrained optimization problem. Under the coherence constraint imposed by the coarse LCLU proportions, the objective is to maximize the spatio-temporal dependence, which is defined by blending both spatial and temporal dependences. Experiments on three datasets showed that the proposed approach can provide more accurate sub-pixel resolution TSIs than conventional SPM methods. The SPM results obtained from the TSIs provide an excellent opportunity for LCLU dynamic monitoring and change detection at a finer spatial resolution than the available coarse spatial resolution TSIs.**

***Index Terms*—Spatio-temporal dependence, land-cover/land-use monitoring, time-series images, sub-pixel mapping, super-resolution mapping.**

I. INTRODUCTION

Monitoring the spatial distribution of land-cover/land-use (LCLU) through time is important for establishing links between policy decisions, regulatory actions and subsequent LCLU activities [1]. Such monitoring has long been recognized as a significant scientific goal since LCLU is a critical variable that describes, and impacts upon, many aspects of urban, rural and natural environments [2]. Satellite remote sensing images provide a major source of LCLU data and have the advantages that satellites can revisit the Earth's surface regularly and that the digital format is suitable for further computer processing. Over the past decades, a growing number of methods have been developed and applied for LCLU mapping from time-series images (TSIs), such as Bayesian classification [3], compound classification [4]-[6], spatio-temporal Markov random fields [7]-[9], domain adaption [10] and spatio-temporal segmentation [11]. The fundamental goal of these techniques is pixel-level LCLU classification of all the images in the time-series, but they are based on a recognition and explicit use of the temporal correlation between images (in the form of, for example, transition probabilities or joint probabilities between LCLU classes).

The Landsat and MODIS sensors are common sources of imagery used for LCLU monitoring due to their free availability, regular revisit capabilities and wide swath. However, they provide coarse spatial resolutions relative to the requirements of certain applications, for example, Landsat 30 m relative to changes in small residential buildings. It is often necessary to monitor LCLU at a fine spatial resolution to provide sufficient detail for specific applications. For the coarse spatial resolution image, each regular gird (i.e., pixel) covers a large area and generally contains more than one LCLU class. This type of pixel is termed a mixed pixel in the context of remote sensing. As one of the most popular mixed pixel analysis techniques, spectral unmixing has been investigated for decades to extract LCLU information within mixed pixels. This technique can estimate the *proportions* of LCLU classes constituting the mixed pixel, and has been applied for the goal of mapping TSIs [12], [13]. The unmixing outputs derived from TSIs, however, can inform users only of how the proportion of each LCLU class changes at the pixel-level, and cannot provide detailed change information at a finer spatial resolution. There is, therefore, a need for techniques that can produce continuous, fine spatial resolution maps from coarse spatial resolution TSIs.

In this paper, sub-pixel mapping (SPM) is suggested for continuous LCLU monitoring at a finer spatial resolution than that of the input TSIs. SPM, also termed super-resolution mapping in remote sensing, is a technique that can be achieved through the post-processing of spectral unmixing [14], [15]. By SPM, each coarse pixel is first divided into multiple sub-pixels and the number of sub-pixels for each class is determined by the spectral unmixing outputs and zoom factor. The sub-pixel classes are then predicted based on maximizing spatial dependence with the assumption that the land cover is spatially dependent both within and between pixels (i.e., compared to more distant pixels, neighboring pixels are more likely to be of

Manuscript received April 22, 2015; revised October 2, 2015 and January 22, 2016; accepted April 2, 2016. This work was supported in part by the Research Grants Council of Hong Kong under Grant PolyU 15223015 and 5249/12E, in part by the National Natural Science Foundation of China under Grant 41331175, in part by the Leading talent Project of National Administration of Surveying under grant K.SZ.XX.VTQA, and in part by the Ministry of Science and Technology of China under Grant 2012BAJ15B04 and 2012AA12A305. *(Corresponding author: W. Shi.)*

Q. Wang was with Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong. He is now with the Lancaster Environment Centre, Lancaster University, Lancaster LA1 4YQ, UK (e-mail: wqm11111@126.com).

W. Shi is with The Hong Kong Polytechnic University, Hong Kong, and also with Wuhan University, Wuhan 430072, China (e-mail: lswzshi@polyu.edu.hk).

P.M. Atkinson is with the Faculty of Science and Technology, Lancaster University, Lancaster LA1 4YR, UK; School of Geography, Archaeology and Palaeoecology, Queen's University Belfast, BT7 1NN, Northern Ireland, UK; and also with Geography and Environment, University of Southampton, Highfield, Southampton SO17 1BJ, UK (e-mail: pma@lancaster.ac.uk).

the same class). SPM transforms pixel level unmixing outputs (i.e., coarse LCLU proportions) into a finer spatial resolution hard classification, and this allows a hard classification technique to be applied at the sub-pixel level.

Over the past few decades, SPM has received increasing attention in the remote sensing community and a variety of approaches have been developed. Some existing SPM algorithms include genetic algorithms [16], particle swarm optimization [17], the pixel swapping algorithm (PSA) [18], [19], Hopfield neural network [20], [21], maximum *a posteriori* method [22], sub-pixel/pixel spatial attraction model (SPSAM) [23], [24], back-propagation neural network [25], [26], Kriging [27], indicator cokriging [28]-[30], Markov random field [31]-[33], contouring method [34], and the newly developed soft-then-hard SPM framework [35], [36]. In these algorithms, spatial dependence is described in different ways.

Recently, SPM has been applied to bi-temporal LCLU mapping [37]-[41]. In [37], with two 300 m Medium Resolution Imaging Spectrometer (MERIS)-like images as inputs, the HNN was employed to detect forest changes in Brazil at a 30 m spatial resolution. With the availability of a fine spatial resolution map (FSRM) on one date, Ling *et al.* [38] and Xu and Huang [39] modified the PSA for SPM of the coarse image on the other date, by borrowing thematic information in the FSRM. In [40], with the aid of a FSRM, a Markov random field model was developed to detect bi-temporal forest changes in the Brazilian Amazon Basin at a 30 m spatial resolution. Wang *et al.* [41] utilized a FSRM to modify the initialization of a Hopfield neural network to achieve more accurate and faster bi-temporal change detection at a sub-pixel resolution.

To the best of our knowledge, very little work has been reported on SPM of coarse spatial resolution TSIs for the purpose of continuous sub-pixel resolution LCLU monitoring. This goal may be achieved straightforwardly by employing directly existing SPM algorithms for the SPM of each coarse image in the TSI in turn. Such a scheme, however, fails to account for the temporal correlation between images. As widely acknowledged, temporal correlation is likely to exist between TSIs covering the same scene. It is always favorable to account for temporal correlation between images when performing LCLU classification of TSIs, as indicated by existing studies on pixel level multi-temporal mapping [3]-[11]. It is of great interest to develop SPM algorithms for continuous LCLU mapping at a sub-pixel resolution, which accounts for spatial *and* temporal dependences simultaneously.

In this paper, a new spatio-temporal SPM algorithm is proposed for multi-temporal LCLU mapping from coarse TSIs. SPM of coarse TSIs is formulated as a constrained optimization problem: The objective is to maximize the spatio-temporal dependence in the TSIs, under the coherence constraint imposed by the coarse proportions of each LCLU class in each image. The spatio-temporal dependence at the sub-pixel scale is defined by fusing the spatial dependence with the temporal dependence. Existing SPM algorithms based on spatial dependence provide effective ways to characterize spatial dependence, which can be described either by the relationship between the sub-pixel and its spatially neighboring sub-pixels or by the relationship between the sub-pixel and its neighboring pixels.

In pixel level multi-temporal classification, temporal links can be described by class transition or joint probabilities [4]-[6], [8], [9]. However, SPM involves scale transformation and, thus, the temporal dependence needs to be depicted at the sub-pixel level. In the proposed spatio-temporal SPM algorithm, one of the main problems to be addressed is related to the definition of an effective mathematical model for temporal dependence characterization. Based on the assumption of temporal dependence, the LCLU information covered by each image of the TSIs is deemed to resemble each other, and the similarity becomes obvious when the images are temporally proximate. In this paper, we propose to quantify the temporal dependence by measuring the similarity in LCLU (but at the sub-pixel level) between images. The temporal dependence is combined with spatial dependence to define the new spatio-temporal dependence.

The SPM problem is always ill-posed, with many multiple plausible solutions that can lead to an equally coherent reproduction of the input coarse proportion images. It is, thus, necessary to borrow information from auxiliary data, such as finer spatial resolution multi-source data [42]-[46] and shape information [47], [48]. The FSRMs are generally convenient to acquire during the period of the TSIs. The FSRM carries reliable LCLU information at the target fine spatial resolution. The proposed spatio-temporal SPM approach is, thus, designed based on the availability of at least one FSRM, which provides reliable fine spatial resolution temporal information for the TSIs. In this paper, the FSRM is assumed to be a "correct" starting point for SPM of the coarse image sequences using a cascade approach.

The proposed spatio-temporal SPM approach holds the following advantages.

1) By fusing spatial and temporal dependences, the two types of dependence are complementary. That is, the spatial dependence accounts for the correlation of LCLU of sub-pixels within each image, while the temporal dependence accounts for the correlation of sub-pixel classes between images in the sequence of TSIs. Thus, information encapsulated in the TSIs is exploited more deeply.
2) The incorporation of a FSRM in the given period can decrease the uncertainty in the SPM problem. The thematic LCLU information in the FSRM is propagated through from the closest to the farthest image in the TSIs. Such information helps to decrease the solution space in the SPM of each image, thereby increasing the SPM accuracy.
3) The temporal and spatial dependences are fused with weights that can be estimated without manual intervention. The weights are estimated by a fitting process, in which the FSRM is treated as a training image. Therefore, quantification of spatio-temporal dependence is completely automatic.
4) The spatial dependence characterization is flexible. The spatial dependence can be described either by the relationship between sub-pixels or by the relationship between a sub-pixel and its neighboring pixels.
5) The approach offers an excellent opportunity for LCLU dynamic monitoring and change detection at a finer spatial resolution than the available coarse TSIs. For example, by applying it to SPM of the coarse MODIS TSIs that inherently have a fine temporal resolution, fine

spatio-temporal resolution LCLU monitoring can be achieved.

The remainder of this paper is organized as follows. Section 2 first presents the problem formulation of spatio-temporal SPM of the TSIs in Section 2.1, and then the approach to spatial dependence characterization in Section 2.2 (including two categories of method to describe spatial dependence) and proposed temporal dependence characterization in Section 2.3. Section 2.4 introduces the proposed spatio-temporal dependence model, followed by the two important considerations for SPM of TSIs (i.e., the starting image and the manner in which sub-pixel information is propagated temporally) in Section 2.5. The algorithm to solve the constrained optimization problem is introduced in Section 2.6. The last sub-section describes the approach to automatic weight estimation. Section 3 provides the experimental results for three case studies. Further discussion is given in Section 4, and Section 5 concludes the paper.

II. METHODS

A. Problem formulation

Let R be the number of TSIs, S be the zoom factor (i.e., each coarse pixel is divided into S by S sub-pixels), P_j^t ($j=1,2,...,M$, M is the number of pixels in each coarse image) be a coarse pixel in the t-th image $\mathbf{I}_t$ ($t=1,2,...,R$) and $F_k(P_j^t)$ be the coarse proportion of the k-th ($k=1,2,...,K$, K is the number of classes) class for pixel P_j^t . Based on physical processes, the coarse proportions estimated by spectral unmixing usually meet the abundance sum-to-one constraint and the abundance non-negativity constraint.

For a particular pixel in each image $\mathbf{I}_t$, say P_j^t , the number of sub-pixels for the k-th class, $E_k(P_j^t)$, is

$$E_k(P_j^t) = \text{round}(F_k(P_j^t)S^2) \tag{1}$$

where round$(\bullet)$ is a function that takes the integer nearest to $\bullet$. The sum of the numbers of sub-pixels for all K classes is S^2 . Let p_{ij}^t (i=1,2,...,MS^2) be a sub-pixel within coarse pixel P_j^t in image $\mathbf{I}_t$, and $B_k(p_{ij}^t)$ be the binary class indicator for the k-th class at sub-pixel p_{ij}^t

$$B_k(p_{ij}^t) = \begin{cases} 1, & \text{if sub-pixel } p_{ij}^t \text{ belongs to class } k \\ 0, & \text{otherwise} \end{cases} . \tag{2}$$

In the SPM result of each image in the TSIs, each sub-pixel should be assigned to only one class and the number of sub-pixels for each class should be consistent with the coarse proportion data, which are described as

$$\begin{aligned} &\sum_{k=1}^{K} B_k(p_{ij}^t) = 1, \quad i=1,2,...,S^2;\ j=1,2,...,M \\ &\sum_{i=1}^{S^2} B_k(p_{ij}^t) = E_k(P_j^t), \quad k=1,2,...,K;\ j=1,2,...,M \end{aligned} . \tag{3}$$

The task of SPM of TSIs is to obtain the binary class indicators for all sub-pixels in all R coarse images in the TSIs. In this paper, they are predicted based on spatio-temporal dependence. In the proposed spatio-temporal dependence-based SPM method, the objective for the SPM problem is formulated as

$$\max \sum_{t=1}^{R}\sum_{j=1}^{M}\sum_{i=1}^{S^2} A(i,j;t) \tag{4}$$

where $A(i,j;t)$ is the spatio-temporal dependence for sub-pixel p_{ij}^t in image $\mathbf{I}_t$. The proposed SPM method aims to maximize the sum of spatio-temporal dependence for all sub-pixels in all TSIs, under the coherence constraint in (3). $A(i,j;t)$ consists of two parts: spatial dependence $D_S(i,j;t)$ and temporal dependence $D_T(i,j;t)$. The two types of dependence are described below.

B. Spatial dependence

Based on the ubiquity of spatial dependence in the environment, at least at some scale, the LCLU is assumed to be spatially dependent within and between pixels; compared to more distant pixels, neighboring pixels are more likely to be of the same class (note this assumption may not be valid for small-sized objects, such as small residential buildings relative to Landsat 30 m). SPM exploits this property by setting the goal of SPM as maximizing the spatial dependence in the predicted image. This is the primary assumption that has underpinned SPM. There are two types of SPM methods to characterize the spatial dependence. One models the relationship between a sub-pixel and its spatially neighboring sub-pixels, while the other models the relationship between a sub-pixel and its neighboring pixels. The popular PSA is a typical SPM method for the former type [18]. With respect to the latter type, we consider three methods, including SPSAM, Kriging and radial basis function (RBF) interpolation [49]. In this paper, the two types of SPM methods are considered to describe the spatial dependence. For simplicity, we denote the spatial dependence quantified by the first and second types as $D_S^{SS}(i,j;t)$ and $D_S^{SP}(i,j;t)$, where "S" and "P" denote "sub-pixel" and "pixel".

1) Spatial dependence described by the relationship between sub-pixels: The PSA assumes that there is attractiveness between sub-pixels. The greater the attractiveness, the greater the spatial dependence. The PSA works by attracting sub-pixels of the same class to cluster spatially under the constraint of coherence with the original pixel-level class proportions. We, therefore, use sub-pixel attractiveness to describe the spatial dependence. Specifically, for a sub-pixel p_{ij}^t , the attractiveness between it and its spatially neighboring sub-pixels is quantified by

$$D_S^{SS}(i,j;t) = \frac{1}{N_{SS}} \sum_{m=1}^{N_{SS}} \sum_{k=1}^{K} w_{SS}(p_{ij}^t, p_m^t) B_k(p_{ij}^t) B_k(p_m^t) \tag{5}$$

where p_m^t is a spatially neighboring sub-pixel of p_{ij}^t in image $\mathbf{I}_t$ and N_{SS} is the number of spatial neighbors. The sub-pixels of the same LCLU class within the spatial neighborhood (i.e., the term $\sum_{k=1}^{K} B_k(p_{ij}^t) B_k(p_m^t)$ takes the value 1) will result in a larger attractiveness value, indicating greater spatial dependence. In (5),

$w_{SS}(p_{ij}^t, p_m^t)$ is a distance-dependent weight for the spatial dependence between sub-pixels p_{ij}^t and p_m^t

$$w_{SS}(p_{ij}^t, p_m^t) = \frac{1}{d_{SS}(p_{ij}^t, p_m^t)^{\psi}} \quad (6)$$

in which $d_{SS}(p_{ij}^t, p_m^t)$ is the spatial (Euclidian) distance between sub-pixels p_{ij}^t and p_m^t, and ψ is a non-linear parameter. The spatial dependence decreases with increasing spatial distance.

2) Spatial dependence described by the relationship between sub-pixels and pixels: In the SPSAM, Kriging and RBF interpolation methods, the relationship between a sub-pixel and neighboring pixels is used to estimate the soft class value at each sub-pixel. Let $F_k(p_{ij}^t)$ be the soft class value for the *k*-th class at sub-pixel p_{ij}^t. Accordingly, the spatial dependence $D_S^{SP}(i, j; t)$ is calculated as

$$D_S^{SP}(i, j; t) = \sum_{k=1}^{K} F_k(p_{ij}^t) B_k(p_{ij}^t) \quad (7)$$

where $F_k(p_{ij}^t)$ depends on the coarse class proportions within the neighboring pixels of p_{ij}^t in image $\mathbf{I}_t$ and the spatial distances between sub-pixel p_{ij}^t and its neighboring pixels. The approach to prediction of $F_k(p_{ij}^t)$ for the three methods can be found in [23], [27], [49].

C. Temporal dependence

It is well known that temporal dependence exists between TSIs. However, how best to describe mathematically the temporal dependence at sub-pixel resolution is a key problem. Temporal dependence has been used widely in pixel-level LCLU mapping. In the existing literature [4]-[6], [8], [9], temporal dependence was modeled by transition or joint probability matrices between LCLU classes. The transition or joint probabilities can be estimated from training data, if such information is available. Commonly, this type of training information can be difficult to acquire, as the training pixels at the different times should have the same coordinate that corresponds to the same points on the ground and should be statistically representative of all the transitions in the whole scene. To release the dependence on such training data, some iterative techniques were developed for estimation of transition or joint probabilities in [4]-[6]. These iterative methods, however, involve computationally costly processes.

For SPM of TSIs, when there is access to high quality training data at the desired fine spatial resolution, they can be used readily to estimate the transition or joint probabilities. With respect to iterative techniques in [4]-[6], although they are directed at pixel level mapping, they undeniably provide informative references for estimation of the probabilities at the sub-pixel level in the future. In this paper, as a simpler alternative and building on the concept of spatial dependence used commonly in SPM, the temporal dependence at sub-pixel resolution is proposed to be characterized by the similarity in LCLU (in terms of class labels) between temporally close images. Based on temporal dependence, the LCLU maps of the TSIs are considered to resemble each other when they are temporally proximate. By maximizing the temporal dependence, the differences in LCLU between the TSIs can be minimized. In temporal space, for each coarse pixel P_j^t, the objective is a constrained optimization problem.

$$\max \sum_{t=1}^{R} \sum_{i=1}^{S^2} D_T(i, j; t). \quad (8)$$

The coherence constraint is the same as that in (3). Theoretically, such a scheme can help to separate more of the real LCLU changes (i.e., signal) from noise. Compared to more temporally distant images, neighboring images have greater similarity in LCLU class. The greater the similarity, the greater the temporal dependence. This assumption is analogous to that for spatial dependence, in which the class label of the sub-pixel is assumed to resemble its spatial neighbors. Therefore, the temporal dependence for each sub-pixel can be described as

$$D_T(i, j; t) = \frac{1}{N_T} \sum_{r=1}^{N_T} \sum_{k=1}^{K} w_T(p_{ij}^t, p_{ij}^r) B_k(p_{ij}^t) B_k(p_{ij}^r) \quad (9)$$

where p_{ij}^r is a sub-pixel in image $\mathbf{I}_r$ that is acquired on a date close to that for image $\mathbf{I}_t$. The temporally neighboring sub-pixel p_{ij}^r has the same spatial coordinate with p_{ij}^t corresponding to the same points on the ground. N_T is the number of temporally neighboring images. $w_T(p_{ij}^t, p_{ij}^r)$ is a weight for the temporal dependence between sub-pixels p_{ij}^t and p_{ij}^r. It depends on the time interval between p_{ij}^t and p_{ij}^r

$$w_T(p_{ij}^t, p_{ij}^r) = \frac{1}{d_T(p_{ij}^t, p_{ij}^r)^{\phi}} \quad (10)$$

where $d_T(p_{ij}^t, p_{ij}^r)$ is the time interval between p_{ij}^t and p_{ij}^r, and measured by the acquisition time intervals between two images, and ϕ is a non-linear parameter. As the time interval increases, the temporal dependence decreases. The binary class indicator of sub-pixel p_{ij}^t (i.e., $B_k(p_{ij}^t)$) is compared to that of p_{ij}^r (i.e., (i.e., $B_k(p_{ij}^r)$)) to measure the similarity in LCLU between temporally close images. If the two sub-pixels belong to the same class, the term $\sum_{k=1}^{K} B_k(p_{ij}^t) B_k(p_{ij}^r)$ takes 1; otherwise, the term takes the value 0, indicating weaker temporal dependence. Thus, the greater the similarity in binary class indicators, the greater the temporal dependence.

D. Spatio-temporal dependence

In the proposed spatio-temporal dependence-based SPM, the sub-pixel class depends not only on the spatial information in the studied image for SPM, but also the thematic information in the temporally neighboring images of the TSIs. The goal is to maximize the spatial autocorrelation in the image for SPM and at the same time the similarity in LCLU between TSIs, under the coherence constraint imposed by the coarse proportions (see (3)). That is, the spatial and temporal dependences need to be maximized simultaneously to achieve SPM. It is essential to choose a suitable fusion approach to combine these two types of dependence. In [50], several existing approaches have been summarized for multisource data fusion, including an approach subdividing the data into subsets of sources and then analyzing

each subset, an ambiguity reduction approach, a supervised relaxation labeling approach and a stacked-vector approach. They have significant limitations as general approaches for multisource data fusion [50].

We select the consensus fusion approach developed in [50] to fuse spatial and temporal dependences. Appreciating the property of finding consensus among members of a group of experts, consensus theory has been applied widely in statistics and management science [50], [51]. An appealing advantage of this fusion approach is that flexible weights can be assigned to different types of dependence and, thus, the contributions of different sources of dependence can be controlled according to specific requirements. As one of the most commonly used consensus rules, the linear opinion pool is employed in this paper. Following this rule, the spatial and temporal dependence is combined linearly to characterize the spatio-temporal dependence. Consequently, the spatio-temporal dependence for a single sub-pixel is

$$A(i,j;t)=\lambda_1(t)D_S(i,j;t)+\lambda_2(t)D_T(i,j;t) \quad (11)$$

where $\lambda_1(t)$ and $\lambda_2(t)$ ($0\le\lambda_1(t),\lambda_2(t)\le 1$) are two weights controlling the influence of the two types of dependence for image $\mathbf{I}_t$, and $\lambda_1(t)+\lambda_2(t)=1$. Both $D_S(i,j;t)$ and $D_T(i,j;t)$ fall within the interval [0, 1], thus, making it easier to choose appropriate weights between 0 and 1. How to determine the optimal weights is a key issue in the consensus fusion approach. The weights cannot be determined analytically. If a training set at the fine spatial resolution is available, the optimal weights can be determined by a training procedure. We treat the FSRM as the training image to estimate the optimal weights. The detailed process is illustrated in Section 2.7.

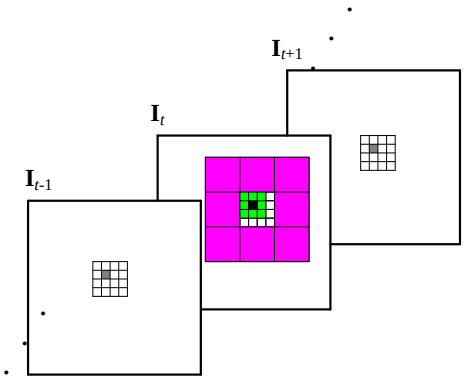


Fig. 1. Spatio-temporal neighbors for a single sub-pixel (marked in black in image $\mathbf{I}_t$). The spatial neighbors are either the green sub-pixels or deep red pixels. The temporal neighbors are the gray sub-pixels in the temporally closest images $\mathbf{I}_{t-1}$ and $\mathbf{I}_{t+1}$.

As seen from Section 2.2, the spatial dependence $D_S(i,j;t)$ can be selected as $D_S^{SS}(i,j;t)$ or $D_S^{SP}(i,j;t)$, as defined in (5) or (7). Consequently, there are two approaches for modeling spatio-temporal dependence, which are denoted as SST and SPT. Fig. 1 shows an example for definition of spatio-temporal neighbors for a sub-pixel. In this example, by using D_S^{SS}, the spatial dependence is described by the relationship between the black sub-pixel and its spatially neighboring sub-pixels (marked in green in image $\mathbf{I}_t$); For D_S^{SP}, the spatial dependence is described by the relationship between the black sub-pixel and its spatial neighboring pixels (marked in deep red in image $\mathbf{I}_t$). The temporal dependence D_T can be described by the relationship between the black sub-pixel and its corresponding sub-pixels in the temporally closest images (marked in gray in images $\mathbf{I}_{t-1}$ and $\mathbf{I}_{t+1}$).

E. Spatio-temporal SPM of TSIs

We assumed access to at least one FSRM at the desired fine spatial resolution in the TSIs. The FSRM can be obtained from a GIS database or by hard classification of a fine spatial resolution remote sensing image (e.g., a Landsat image amongst coarse MODIS TSIs), under the condition that the source of FSRMs is temporally close to the studied TSIs. If the spatial information of the FSRM is coarser than the desired fine spatial resolution for SPM, an additional downscaling process will be required to provide the FSRM at the desired fine spatial resolution. The thematic LCLU information from the FSRMs can be used in the temporal dependence characterization. This section introduces the approach to SPM of TSIs using the concept of spatio-temporal dependence. The SPM process is performed for each image one-by-one. For convenience of illustration, we consider the case of one FSRM. The approach to SPM of TSIs introduced in this section can also be extended to the case of multiple FSRMs. When conducting SPM of TSIs, there are two important considerations. One is the starting image, while the other is the manner in which the temporal information is propagated.

1) Starting point: As for the starting image, an intuitive option is the image at the earliest time. In this way, the starting point is a SPM solution of the earliest image and involves the inevitable uncertainty of the scale transformation (i.e., downscaling). By utilizing the temporal information, the uncertainty in the SPM of the starting image may be propagated to the SPM process of later images. To avoid such uncertainty, we select the FSRM as the starting point.

The FSRM is a thematic LCLU map at the target fine spatial resolution, and can be regarded as a highly reliable SPM result at that time. From this viewpoint, the whole SPM process of the TSIs starts from the time closest to that of the FSRM to make the most use of the fine spatial resolution LCLU information in the FSRM and decrease the uncertainty in the SPM. Fig. 2 shows an example for illustration of this point. Suppose the FSRM is the x-th ($x\in\{1,2,...,R\}$) image in the TSIs, and the zoom factor S for SPM of each coarse image is four. The SPM process begins from the fine spatial resolution thematic map $\mathbf{I}_x$ looking to both of its sides: SPM of images $\mathbf{I}_{x-1}$ and $\mathbf{I}_{x+1}$ is carried out first. The LCLU distribution in the FSRM (such as red, yellow and blue pixels in the coarse pixel in Fig. 2) can be included in the temporal dependence characterization and used to aid the SPM of the corresponding coarse pixels in the temporally closest images $\mathbf{I}_{x-1}$ and $\mathbf{I}_{x+1}$.

2) Propagation of temporal information: In multi-temporal image classification (but at pixel-level), two main approaches have been suggested for propagation of temporal information, that is, the cascade and mutual approaches [3], [8]. The main difference between the two approaches is choice of temporally neighboring images. For SPM of each image in the TSIs, the mutual approach borrows temporal information from images before it and after it. It repeats the SPM of TSIs to decrease the uncertainty and allow enough iteration for the process to

converge on a satisfactory solution. Such an iterative scheme is generally computationally expensive, especially when the number of TSIs (i.e., R) is large. In contrast to the mutual approach, the cascade approach is a single-pass scheme and, thus, non-iterative. In this paper, we use the non-iterative cascade approach for propagation of temporal information as it allows a significant simplification of the proposed spatio-temporal SPM for TSIs.

In the cascade approach, once the image on a given date has been classified by SPM, the resultant map is considered as the source of temporal information for the next image that is temporally closest to it. For example, in Fig. 2, after SPM of image $\mathbf{I}_{x-1}$ is completed by using the temporal information from the closest image (i.e., the FSRM), the SPM result along with the FSRM provides the temporal information for the next image $\mathbf{I}_{x-2}$ (see (11)), and so on. This is also the case for the images at the other side of $\mathbf{I}_x$ (i.e., $\mathbf{I}_{x+1}$, $\mathbf{I}_{x+2}$,…, $\mathbf{I}_R$). The arrows in Fig. 2 show the direction of temporal information propagation. The whole process is terminated when the SPMs of all coarse TSIs are predicted once.

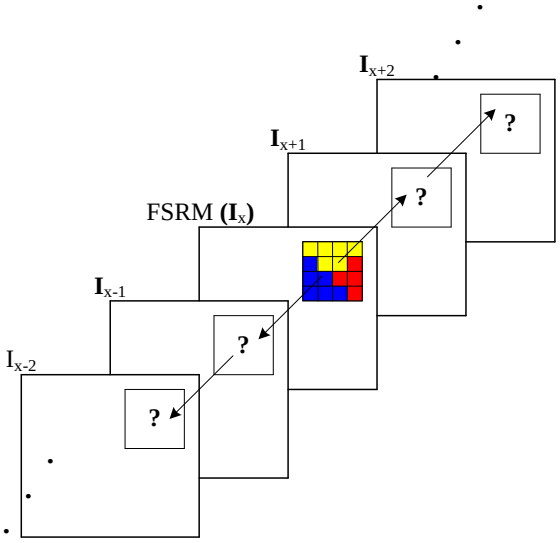


Fig. 2. SPM of the TSIs, in which the FSRM is considered as the starting point. The red, yellow and blue colors represent three LCLU classes in a coarse pixel containing 4 by 4 sub-pixels.

F. Model optimization

The proposed SPM method for coarse spatial resolution TSIs is implemented by maximizing the overall spatio-temporal dependence, as defined in the optimization problem in (4). The coherence constraint in (3) is imposed in the SPM of all TSIs. This section introduces the approach to solve the optimization problem. Ideally, the most suitable distribution of all sub-pixel classes within each coarse pixel can be obtained by evaluating all possible configurations and selecting the one that meets the constraint in (3) and maximizes the objective function in (4). This assumption works mainly for cases involving small-size images and a small zoom factor. For a large zoom factor, the number of combinations of possible sub-pixel spatial distributions increases dramatically and the computational load may become unrealistic. This necessitates the application of an effective optimization algorithm to solve the optimization problem. The simulated annealing algorithm is employed for this purpose [51]. Readers may refer to Atkinson [51] for details on this algorithm.

The input is a set of proportion images for all TSIs, and the whole solving process contains two stages: initialization and update. The whole flowchart is shown in Fig. 3.

Stage 1: Initialization. According to the constraint in (3), in each image, sub-pixels for each class are allocated. As a straightforward scheme, sub-pixel classes can be allocated randomly. However, to achieve a faster convergence rate, this paper adopts the corresponding basic SPM algorithm (i.e., SPSAM, Kriging or RBF method) to produce the initial SPM maps. For the SST method, in which the PSA is essentially the basic SPM algorithm, the simple and fast SPSAM is utilized for initialization. After initialization, only the spatial locations of the sub-pixels can vary, and the number of sub-pixels for each class within each coarse pixel is fixed.

Stage 2: Update. As mentioned in Section 2.5, the SPM process is started from the FSRM and implemented for each coarse spatial resolution image one-by-one, that is, based on the cascade approach.

1) For each coarse image, SPM is conducted in units of coarse pixels.
2) For a current coarse image $\mathbf{I}_t$, within a particular coarse pixel P_j^t, the following steps are implemented.
   a) The sum of spatial dependence for all S^2 sub-pixels is calculated by using $D_S^{SS}(i,j;t)$ in (5) or $D_S^{SP}(i,j;t)$ in (7). Then, with the temporal neighbors in images from the FSRM to $\mathbf{I}_{t-1}$ (if the time of $\mathbf{I}_t$ is after the FSRM) or $\mathbf{I}_{t+1}$ (if the time of $\mathbf{I}_t$ is before the FSRM), the sum of temporal dependence for all S^2 sub-pixels is calculated by using $D_T(i,j;t)$ in (9). For all S^2 sub-pixels, the sum of spatio-temporal dependence is calculated according to (11).
   b) A pair of sub-pixels with different class labels is selected randomly and their spatial locations are swapped. The sum of spatio-temporal dependence for all S^2 sub-pixels in the new configuration is calculated again. If the overall spatio-temporal dependence increases, the swap is accepted; Otherwise, the swap is allowed with a certain probability determined according to the current "temperature". Such a probability decreases with the decreasing temperature at each iteration.
3) For each coarse pixel, steps a) and b) are implemented.
4) For the current image $\mathbf{I}_t$, the swap process is repeated until the pre-defined number of iterations is reached.
5) For each coarse image in the TSIs, steps 1)-4) are implemented.

When calculating $D_S^{SS}(i,j;t)$, the class labels of the neighboring sub-pixels are used, see (5). However, they are updated after each iteration. Thus, this type of spatial dependence needs to be calculated at each iteration. For $D_S^{SP}(i,j;t)$, however, it is calculated using the fixed coarse proportions, see (7). For each sub-pixel, the spatial dependence of all cases (one case corresponds to one class) can be quantified according to (7) in advance. The calculation is conducted only once and the generated values can be utilized in all iterations.

Therefore, the SPT approach is deemed more computationally efficient than the SST approach.

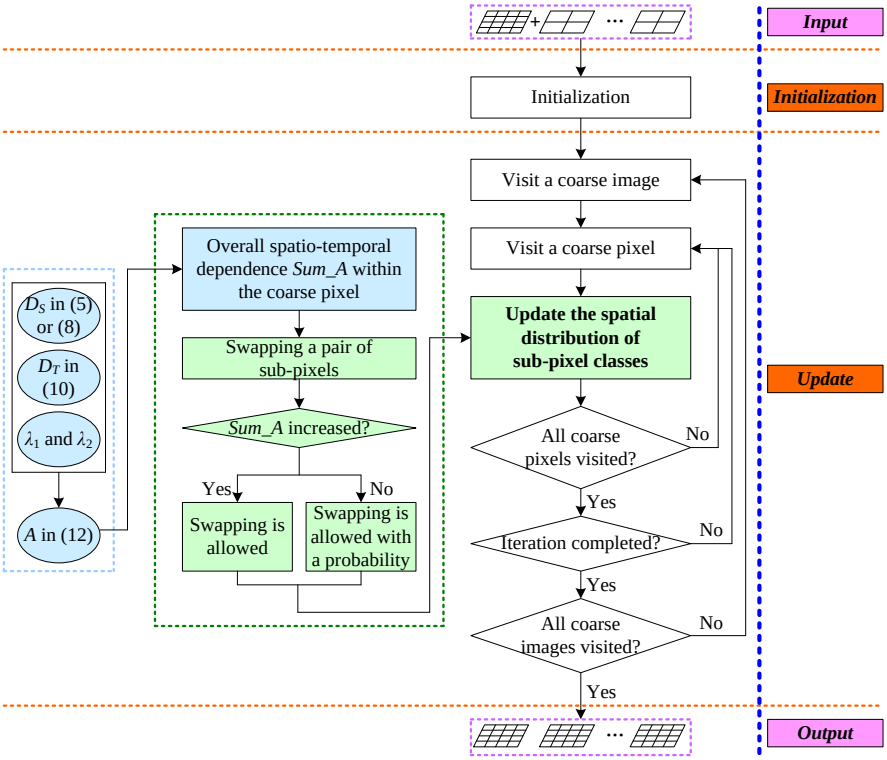


Fig. 3. Flowchart of the proposed spatio-temporal SPM algorithm.

G. *Estimation of optimal weights*

The weights in (11) (i.e., $\lambda_1(t)$ and $\lambda_2(t)$) control the influence of the spatial and temporal dependences. This section introduces a new approach for completely automatic estimation of the optimal weights. As mentioned earlier, the FSRM is regarded as a highly reliable thematic LCLU map at the target fine spatial resolution. We therefore adopt the FSRM as a training image for weight estimation using a fitting procedure. The weights need to be estimated for each coarse image $\mathbf{I}_t$ ($t=1,2,...,R$) in the TSIs (except the FSRM). Essentially, only one weight, either $\lambda_1(t)$ or $\lambda_2(t)$, needs to be estimated for each coarse image, as $\lambda_1(t)+\lambda_2(t)=1$.

Suppose S is the spatial resolution (zoom) ratio between the coarse images and FSRM, the FSRM is the x-th image in the TSIs (see Fig. 2) and the current image is $\mathbf{I}_{x+n}$. The FSRM is first applied to spatio-temporal SPM of the coarse images from $\mathbf{I}_{x+1}$ to $\mathbf{I}_{x+n}$ along the single direction, and then their SPM results are applied to spatio-temporal SPM of the degraded FSRM backwards. The original FSRM is used to examine each weight. The detailed processes are described as follows.

Step 1: A weight pool is set for $\lambda_1(x+n)$: $\{\lambda_{1,1}(x+n),\lambda_{1,2}(x+n),...,\lambda_{1,L}(x+n)\}$. In this paper, $\lambda_1(x+n)$ was varied from 0.1 to 0.9 with a step of 0.1, that is, the pool set is $\{0.1,0.2,...,0.9\}$.

Step 2: A weight $\lambda_{1,l}(x+n)$ ($l\in\{1,2,...,L\}$) is selected from the pool and the following procedures are conducted.

1) Regarding the FSRM as a starting point, spatio-temporal SPM of coarse images $\mathbf{I}_{x+1}$, $\mathbf{I}_{x+2}$,…, $\mathbf{I}_{x+n}$ is performed with a zoom factor of S. In this process, the temporal information from the FSRM is propagated from $\mathbf{I}_{x+1}$ to $\mathbf{I}_{x+n}$, as illustrated in Section 2.6.
2) The FSRM is degraded with the factor of S to simulate the coarse images at that time.
3) SPM of the simulated coarse images for FSRM using the spatio-temporal model, in which the SPM results of $\mathbf{I}_{x+1}$, $\mathbf{I}_{x+2}$,…, $\mathbf{I}_{x+n}$ are considered as temporally neighboring images.
4) The original FSRM is used for supervised assessment of the corresponding SPM result, and an accuracy value is recorded for the selected parameter.

Step 3: Step 2 is implemented for all weights in the pool and L accuracy values are obtained as a result.

Step 4: The weight leading to the greatest accuracy is determined as the optimal one.

Step 5: Steps 1-4 are performed for the next coarse image $\mathbf{I}_{x+n+1}$ to estimate the corresponding weight $\lambda_1(x+n+1)$. The whole procedure is terminated after all coarse images are visited.

Fig. 4 is a flowchart of the weight estimation method. In this example, FSRM is assumed to be $\mathbf{I}_0$ and SPM goes from $\mathbf{I}_1$ to $\mathbf{I}_t$ directly. When the FSRM is not $\mathbf{I}_0$, SPM of each side follows the rule in Fig. 4. We can see from the procedure that the functions of FSRM in the proposed spatio-temporal approach are twofold: it not only provides valuable fine spatial resolution temporal information for the TSIs, but also acts as a training image to obtain the optimal weight.

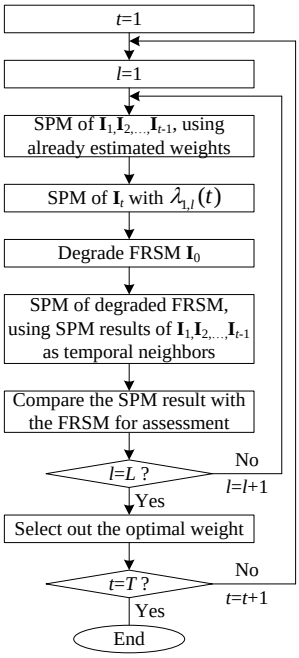


Fig. 4. Flowchart of optimal weight estimation approach, where the FSRM is assumed to be $\mathbf{I}_0$.

III. Experiments

Two synthetic datasets and one real dataset were used in the experiments to examine the proposed spatio-temporal SPM approach. As stated in Section 2, there are two approaches for modeling spatial dependence, that is, D_S^{SS} in (5) and D_S^{SP} in (7). PSA was used for D_S^{SS}, while three methods, SPSAM, Kriging, RBF, were used as for D_S^{SP}. The corresponding spatio-temporal dependence structures are referred to as SST and SPT. The four original SPM methods were considered as benchmark algorithms in this section. For the SPSAM, Kriging and RBF methods (whether or not they are coupled with temporal dependence), the window sizes of the neighborhood were set to 3, 5 and 5 [23], [27], [49]. The parameter in the basis function (i.e., Gaussian function) was set to 10 [49]. In addition, to illustrate the benefit of the SPM technique in LCLU mapping, traditional pixel level hard classification (HC) was performed, by which all sub-pixels within a coarse pixel are assigned to the dominant class. In total, nine methods were compared for SPM of TSIs.

SPM is essentially a hard classification technique (but at the sub-pixel scale). The performances of the SPM methods were evaluated quantitatively by the classification accuracy of each class and the overall accuracy (OA) in terms of the percentage of correctly classified pixels. In the experiments on synthetic datasets, synthetic coarse images were considered, which contain no uncertainty in the coarse proportions. For pure pixels, SPM assigns all sub-pixels within it to the same class to which the pure pixel belongs. This simple copy process will only increase the SPM accuracy statistics without providing any useful information on the actual performance of the SPM methods, as suggested by the existing literature [21]. Therefore, for the synthetic coarse images, we did not consider the non-mixed pixels in the accuracy statistics. For the real dataset, both mixed and non-mixed pixels were included in the accuracy statistics.

A. Synthetic datasets

Two synthetic datasets were used for validation in Sections 3.2 and 3.3. Specifically, the fine spatial resolution (i.e., 30 m in the experiments) TSIs are available and were degraded to synthesize the coarse spatial resolution TSIs. One of the 30 m thematic maps was considered as the FSRM. The coarse class proportion images were simulated by degrading the other 30 m thematic maps via an S by S mean filter. SPM methods were implemented to recreate the 30 m LCLU maps of the TSIs. The produced SPM results were compared to the corresponding reference maps for assessment. By using synthetic coarse images, the input proportions were known to be error free and represent greater control in the test. Moreover, the reference maps are known perfectly for SPM evaluation. The test is directed at the SPM algorithm itself which is appropriate at the method development stage [14].

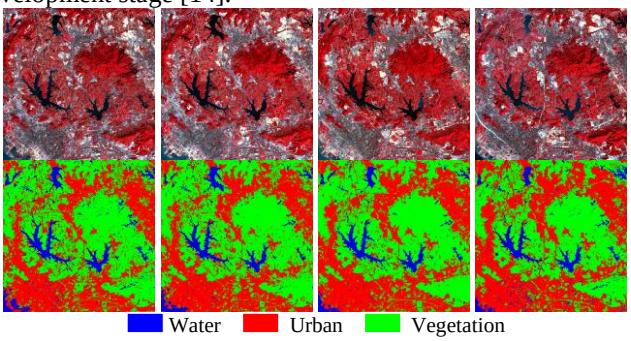


Fig. 5. Four Landsat images of Shenzhen, China on four dates. From left to right: $\mathbf{I}_1$ in Nov 2001, $\mathbf{I}_2$ in Nov 2002, $\mathbf{I}_3$ in Nov 2004 and $\mathbf{I}_4$ in 23 Nov 2005. Line 1: Color image (Bands 4, 3 and 2 as RGB). Line 2: Hard classified LCLU maps used as reference.

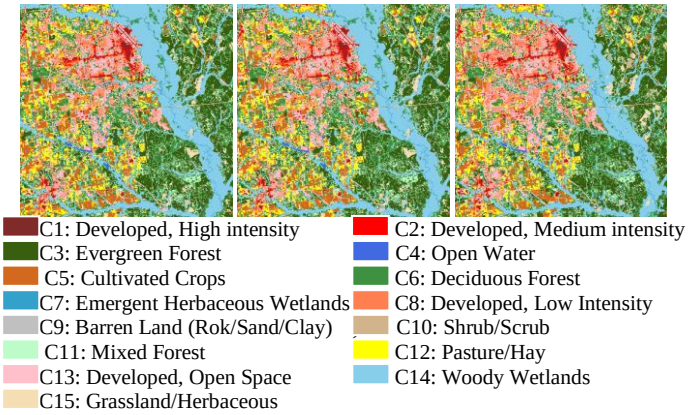


Fig. 6. Three NLCD maps in Georgia, US at three times. From left to right: NLCD 2001, NLCD 2006 and NLCD 2011.

The first dataset includes four 30 m Landsat images covering an area in Shenzhen, China. They were acquired in Nov 2001 ($\mathbf{I}_1$), Nov 2002 ($\mathbf{I}_2$), Nov 2004 ($\mathbf{I}_3$) and Nov 2005 ($\mathbf{I}_4$). Registration and radiometric correction (using the LEDAPS tool) were applied to the Landsat images. The study area is a heterogeneous region covered by 600 by 600 pixels in which three main LCLU classes can be identified, including water, urban and vegetation. The four images were classified using K-means-based unsupervised classification to generate the four 30 m reference LCLU maps. Fig. 5 shows the four images and the classified LCLU maps.

The second dataset includes three maps from the National Land Cover Database (NLCD) 2001, 2006 and 2011. The NLCD dataset is a raster-based classification with a 30 m spatial resolution covering all 50 US states and Puerto Rico. The study area covers an area in Georgia, and has a size of 1000 by 1000 pixels and ground extent of 30 km by 30 km. As shown in Fig. 6, 15 classes are presented in the maps, which are labeled as C1-C15. This dataset aims to examine the proposed approach for a large region with a large number of LCLU classes.

B. Experiment on the Shenzhen Landsat images

In this section, we used the 30 m reference map in 2001 as the FSRM. The other three 30 m maps were degraded with an 8 by 8 mean filter to synthesize 240 m MODIS-like TSIs (R=3). Fig. 7 shows the 240 m proportion images of the three classes for the image in 2002, which can be treated as error-free spectral unmixing results. Through visual inspection, due to the ambiguous boundaries between classes, the LCLU information presented in these proportion images was found to be insufficient for interpretation. Three sets of proportion images were taken as input for SPM. With a zoom factor of eight (i.e., S=8), three 30 m LCLU maps of the TSIs were reproduced. We took the results of the 2002 image as an example for visual inspection.

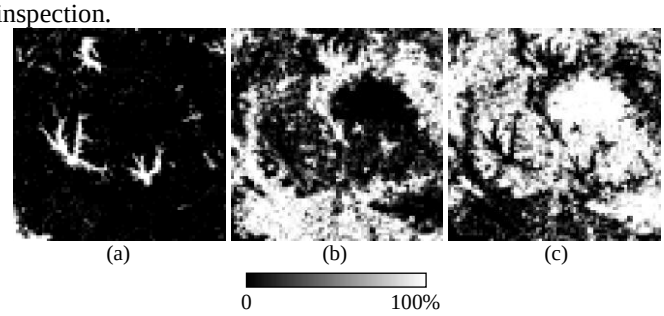


Fig. 7. Synthesized 240 m proportion images of the 2002 Shenzhen Landsat image. (a) Water. (b) Urban. (c) Vegetation.

We first show the influence of the weights (see (11)) in the proposed spatio-temporal approach in Fig. 8. Both the SPT and SST methods produce a stable accuracy when the weights change from 0.1 to 0.6. The approach presented in Section 2.7 is able to determine an appropriate weight for characterizing spatio-temporal dependence, as marked by the asterisk. Fig. 9 shows the SPM results of the nine methods for the 2002 image. The HC result in Fig. 9(a) was dominated by the jagged boundaries that provide limited LCLU information at the 30 m spatial resolution. The other eight SPM methods produced more detailed LCLU information than the HC method and the boundaries in Fig. 9(b)-Fig. 9(i) were characterized by more fine (i.e., 30 m) pixels. This reveals the obvious benefit of SPM in LCLU mapping. Comparing the results of the four proposed spatio-temporal SPM methods in Fig. 9(f)-Fig. 9(i) to those of the original methods in Fig. 9(b)-Fig. 9(e), the proposed methods produced much more satisfying results than the original methods. The original SPM methods (i.e., SPSAM, Kriging and RBF), based only on the spatial dependence between sub-pixels and neighboring coarse pixels, produced many linear artifacts, particularly for the SPSAM method. This phenomenon can be illustrated by the distribution of the urban class in the results. For the original PSA method, which described the spatial dependence at sub-pixel level, the result was over-smooth (see, e.g., the boundaries of the river class), with many disconnected and hole-shaped patches (e.g., the restoration of the urban class). The four proposed methods, by accommodating temporal information propagated from the FSRM, restored many linear features and small size patches, and the results were similar to the 2002 reference map in Fig. 5.

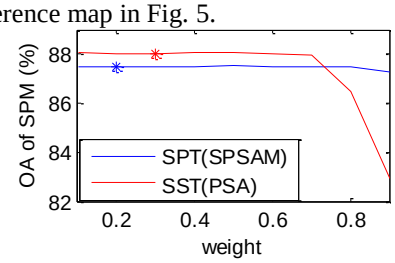


Fig. 8. Influence of weights in (11) for SPM of the 2002 Shenzhen Landsat image, where the estimated optimal weights in each case are marked by the asterisk.

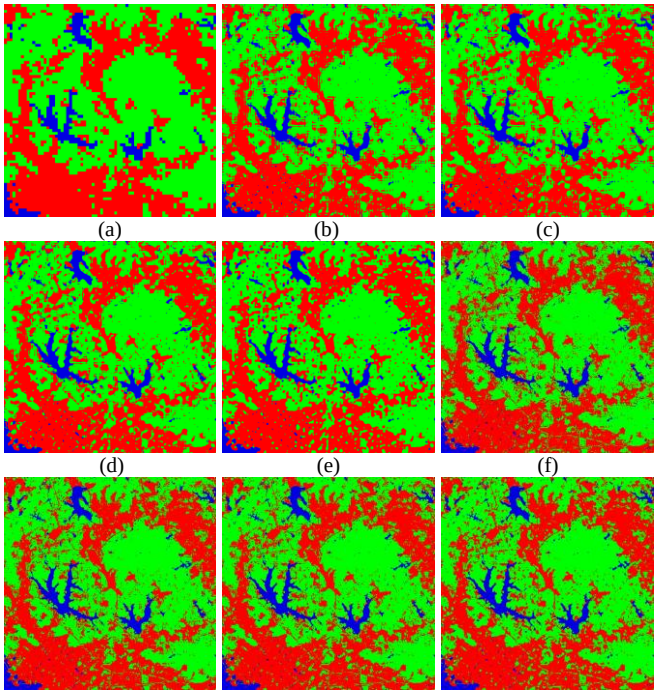


Fig. 9. Results of the 2002 Shenzhen Landsat image. (a) HC. (b) SPSAM. (c) Kriging. (d) RBF. (e) PSA. (f)-(h) SPT results of SPSAM, Kriging and RBF. (i) SST (PSA).

Table 1 lists the accuracies of the nine methods for all three images in the TSIs. Checking the class accuracies as well as the OAs for all images, the proposed spatio-temporal approaches were superior to HC and the four original SPM methods (the differences in OAs are statistically significant at the 95% level of confidence based on the McNemar test). For the HC and four original SPM methods, they produced close OAs (around 79% for all three images). Using the proposed spatio-temporal approaches, all four original SPM methods were enhanced. More precisely, for the 2002 image, using the proposed approaches, the accuracy gains of the water, urban and vegetation classes were about 15%, 8% and 8%, respectively, and the gains of OA were about 9%. Focusing on the values of the 2004 image, the increases in accuracy were smaller than those for the 2002 image. The accuracies of the water, urban and vegetation classes increased by around 13%, 4% and 5%, respectively, and the OAs increased by around 5%. Regarding

the 2005 image, the OAs of the proposed approaches were about 4% larger than those of the original SPM methods. Therefore, for spatio-temporal SPM of a coarse image, the SPM accuracy decreased when the acquisition time interval between the FSRM and the coarse image increased. An interesting observation is that the accuracy increase for water was much greater than that for the urban and vegetation classes. Moreover, it is worth noting that the SST and SPT approaches have similar performances in SPM: the OAs of the four new methods are close and the differences are insignificant at the 95% level of confidence.

C. Experiment on the NLCD maps

In the experiment on the NLCD maps, the NLCD 2001 map was selected as the FSRM. The NLCD 2006 and 2011 were degraded with an 8 by 8 mean filter to simulate the 240 m coarse TSIs. The synthesized two sets of coarse proportion images were considered as spectral unmixing results. SPM was performed with S=8 to restore the 30 m LCLU maps for the TSIs. The results of the NLCD 2011 map are shown in Fig. 10. For clear visual inspection, we present zoomed results of a sub-area, with a size of 100 by 100 pixels and marked in Fig. 10(a).

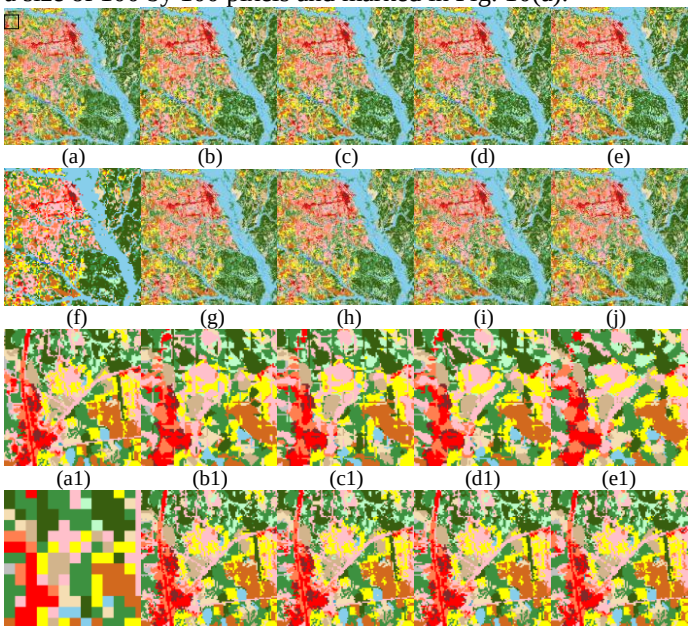


Fig. 10. Results for the NLCD 2011 map. (a) Reference. (b) SPSAM. (c) Kriging. (d) RBF. (e) PSA. (f) HC. (g)-(i) SPT results of SPSAM, Kriging and RBF. (j) SST (PSA). (a1)-(j1) Results of the sub-area.

Examining the results, again the proposed spatio-temporal approaches produced more accurate results than the other approaches. Specifically, the HC result has an unnatural blocky appearance, and many features are mis-represented. Although the four original SPM methods were able to reproduce more LCLU information, the configuration of the classes was considerably different from that in the reference in Fig. 10(a). For example, they failed to reproduce the linear features of the C2 class and the C11 class was over-compact. There exist many large patches and linear artifacts in the SPSAM, Kriging and RBF results, and many locally smooth and disconnected patches in the PSA result. With respect to the proposed methods, however, most of the fine pixels were correctly located. The configurations of the scattered C11 and C12 classes were generally accurately reproduced and the linear feature for the C2 class was also well restored. Referring to Fig. 10(a), the results of the proposed methods were very close to the reference.

The quantitative results of the nine methods are displayed in Table 2. Consistent with the abovementioned visual evaluation, the four proposed spatio-temporal SPM methods produced greater accuracy than the other methods for both the NLCD 2006 and 2011 maps. Examining the results for the 2006 map, the OA gains from the four original SPM methods compared to the four corresponding spatio-temporal SPM methods were around 35%. For the 2011 map, the OAs of the four original SPM methods increased from 59% to 90% for the proposed methods, with gains of 31%. Furthermore, inter-comparison of the four spatio-temporal SPM methods reveals that the two types of spatio-temporal dependence (i.e., SST and SPT) led to similar accuracies (the differences are insignificant at the 95% level of confidence). More precisely, they yielded accuracies of about 93.4% and 90% for the 2006 and 2011 maps, respectively.

Table 1 SPM accuracy (%) of the nine methods for the TSIs. The 30 m reference map in 2001 was used as the FSRM

2002									
	HC	SPSAM	Kriging	RBF	PSA	SPT			SST (PSA)
						SPSAM	Kriging	RBF	
Water	58.85	67.18	67.73	68.67	69.22	84.60	84.67	85.20	85.75
Urban	78.06	77.72	78.19	78.49	77.97	86.66	86.81	86.91	87.21
Vegetation	82.30	80.41	80.80	81.13	80.76	88.52	88.65	88.74	89.01
OA	79.07	78.46	78.90	79.25	78.87	87.49	87.63	87.75	88.04
2004									
	HC	SPSAM	Kriging	RBF	PSA	SPT			SST (PSA)
						SPSAM	Kriging	RBF	
Water	61.82	68.40	69.21	70.05	70.72	83.34	83.13	83.29	84.51
Urban	80.92	80.32	80.78	81.06	80.55	84.09	84.27	84.31	84.48
Vegetation	80.11	80.03	80.56	80.94	80.51	84.73	84.90	84.96	85.17
OA	79.32	79.43	79.94	80.30	79.91	84.35	84.50	84.56	84.81
2005									
	HC	SPSAM	Kriging	RBF	PSA	SPT			SST (PSA)
						SPSAM	Kriging	RBF	

Water	49.55	58.41	58.97	59.56	59.85	73.91	73.98	74.18	74.87
Urban	81.64	79.46	79.88	80.19	79.72	82.79	83.03	83.10	83.46
Vegetation	80.51	79.99	80.46	80.84	80.35	84.17	84.40	84.46	84.85
OA	78.78	78.19	78.64	79.00	78.58	82.80	83.02	83.10	83.50

Table 2 SPM accuracy (%) of the nine methods for the TSIs. The 30 m NLCD 2001 map was used as the FSRM

	HC	SPSAM	Kriging	RBF	PSA	SPT			SST
						SPSAM	Kriging	RBF	(PSA)
NLCD 2006	58.59	58.69	59.39	59.84	59.37	93.33	93.42	93.46	93.39
NLCD 2011	58.01	58.15	58.80	59.23	58.80	89.98	90.15	90.22	90.19

D. Influence of the FSRM and registration errors

The proposed spatio-temporal approach borrows fine spatial resolution information in the FSRM when accounting for temporal dependence. It acts as a starting point for SPM of the TSIs and its information is propagated through the whole time-series. This necessitates a study on the influence of the FSRM on the performance of the proposed method. In this section, we considered two factors when using the FSRM: the error of the FSRM and the registration error between the FSRM and the coarse time-series images.

1) Error of the FSRM: The NLCD maps were used to analyze the influence of the error of the FSRM on the proposed approach. The 30 m NLCD 2001 map was used as the FSRM. The NLCD 2006 and 2011 maps were degraded via an 8 by 8 mean filter. Taking the sythsized 240 m proportion images as input, SPM was performed with S=8 to restore the 30 m TSIs. We simulated FSRMs with different errors by changing the class labels of some pixels in the 30 m NLCD 2001 map. For example, to create a FSRM with an error of 10%, 10% of the pixels in the whole map were selected randomly and their class labels were changed to others. In this sub-section, FSRMs were simulated with errors of 5%, 10%, 15%, 20%, 25%, 30%, 40%, 50%, 60%, 70%, 80% and 90%. The performances of both SPT- and SST- based spatio-temporal approaches were tested. Here, the RBF method was selected as a representative for the SPT approach. Fig. 11 displays the SPM accuracies of the SPT and SST approaches in relation to the error of the FSRM. Meanwhile, the accuracies of two original SPM methods (i.e., RBF and PSA) were also included as benchmarks. From the figure, three observations can be made.

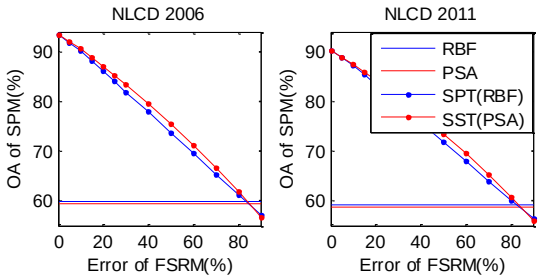


Fig. 11. Influence of the error of the FSRM on the proposed method.

First, the accuracies decreased with increasing error of the FSRM. For SPM of the coarse 2006 image, the OAs of two spatio-temporal approaches decreased from 93% to 56% when the error increased from 0 to 90%, while for the 2011 image, the OAs decreased from 90% to 55% correspondingly.

Second, when the error of the FSRM reached 80%, the accuracies of the spatio-temporal SPM approaches were lower than those of the original methods, suggesting that the FSRM could not help to increase the accuracy of the proposed approaches.

Third, when the FSRM has an error between 20% and 70%, the SST approach is more advantageous in comparison with the SPT approach. Particularly, when the error falls within [40%, 60%], for each image the OA of the SST approach was over 1% larger than that of the SPT approach.

2) Registration error: The Shenzhen Landsat images were used to analyze the registration error between the FSRM and the coarse time-series images. The 30 m reference map in 2001 was used as the FSRM, and the other three images were degraded with an 8 by 8 mean filter to simulate the 240 m coarse TSIs. SPM was conducted with S=8 for each coarse image. We simulated registration errors ranging from −4 to 4 sub-pixels. Fig. 12 shows the SPM accuracies of the original RBF and PSA approaches, and the proposed SPT (RBF) and SST (PSA) approaches in relation to the registration error. It can be observed clearly that the registration error imposes a negative effect on the proposed spatio-temporal SPM methods and the accuracies of both SPT and SST decrease with the increasing absolute value of registration error. Moreover, when the absolute value of error exceeds two sub-pixels, SPT and SST produce smaller accuracies than the original RBF and PSA methods. Therefore, reliable registration is critical in the proposed spatio-temporal SPM methods.

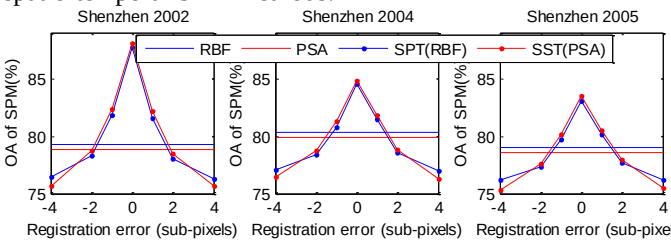


Fig. 12. Influence of registraiton error between the FSRM and coarse time-series images.

E. Computing Efficiency

All experiments were carried out on an Intel Core i7 Processor at 3.40 GHz with the MATLAB 7.1 version. In the spatio-temporal approaches, the number of iterations for simulated annealing was set to 3000. In the original PSA method, the iteration number was 1000. Table 2 is the average computing time for each image achieved using the SPM methods in the experiments. The computational cost of SPM is closely related to the size of coarse image, zoom factor and number of classes in the study area. As observed from the table, the original SPSAM, Kriging and RBF took much less time than the original PSA method. This is because the SPSAM, Kriging and RBF methods are implemented based on D_S^{SP}, that is, calculated using the fixed coarse proportions and, thus, are non-iterative. However,

the PSA method depends on the class labels of sub-pixels and requires several iterations to achieve an acceptable solution.

With respect to the four proposed spatio-temporal approaches, they were more time-consuming than the original methods. The reason is that the incorporation of temporal aspects in the proposed approaches complicated the process of model optimization, and more iterations were required to converge to a satisfactory solution. It is worth noting that by using D_S^{SP} for spatial dependence characterization, the SPT approaches are faster than the SST approach. The main reason is that D_S^{SP} is calculated only once and utilized in all iterations, while D_S^{SS} needs to be calculated at each iteration.

Table 3 Computing time (average for each image in the TSIs) of the SPM methods in the experiments

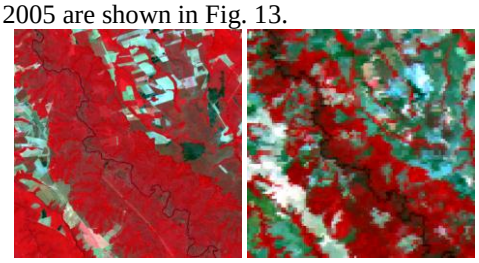
	Size of coarse image	Zoom factor *S*	Number of classes	SPSAM	Kriging	RBF	PSA	SPT			SST (PSA)
								SPSAM	Kriging	RBF	
Shenzhen	75×75	8	3	2s	17s	17s	65s	152s	167s	167s	200s
NLCD	125×125	8	15	30s	49s	48s	214s	490s	509s	508s	800s

Table 4 SPM accuracy (%) of the SPM methods for the real dataset

	SPSAM	Kriging	RBF	PSA	SPT			SST (PSA)
					SPSAM	Kriging	RBF	
OA	70.70	70.66	70.61	70.63	71.78	71.75	71.72	71.91

F. Application to the real dataset

The studied real dataset covers a 45 km by 45 km area of tropical forest in Brazil, within which two main land cover classes, forest and non-forest, were identified. We performed SPM of a MODIS image acquired in July 2005, using the Landsat image acquired in July 1988 as the source of the FSRM (the FSRM was produced with an unsupervised *k*-means classifier). The original seven-band MODIS image was re-projected into a Universal Transverse Mercator projection and resampled to 450 m using the nearest-neighbor algorithm. The task of SPM was to produce a 30 m spatial resolution land cover map in July 2005, and the zoom factor for SPM was set to 15. The Landsat image acquired in July 2005 was classified with an unsupervised *k*-means classifier to provide the reference for accuracy assessment. The Landsat image in 1988 and MODIS image in 2005 are shown in Fig. 13.



(a) (b)

Fig. 13. The real dataset. (a) The Landsat image in July 1988 (1500 by 1500 pixels, bands 432 as RGB). (b) The MODIS image in July 2005 (100 by 100 pixels, bands 214 as RGB).

For the MODIS image, spectral unmixing was performed based on a fully constrained least squares linear spectral mixture analysis. The proportion images of forest and non-forest were used as input for SPM. Table 4 presents the accuracy of the eight SPM methods. Checking the results, the four proposed spatio-temporal SPM methods produce greater accuracies than the four conventional methods and the OA gains were around 1%. For the real dataset, there exist uncertainties in the unmixing algorithm, the point spread function of the MODIS sensor, registration, the FSRM and the reference map. To study the effect of these uncertainties in realistic circumstances, the 30 m reference map in July 2005 was degraded with a factor of 15 to simulate the 450 m spatial resolution proportions, and the SPM accuracy resulting from such a design was over 94% (20% larger than that of the OA values in Table 4).

IV. Discussion

This paper presents a spatio-temporal SPM approach to transfer coarse spatial resolution TSIs to a sequence of fine spatial resolution LCLU maps. The proposed approach was tested using two synthetic datasets and one real dataset. It was shown that, by accommodating temporal information, through a model of temporal dependence and initially propagated from the FSRM, the proposed method can produce greater SPM accuracy in comparison to conventional SPM methods. Through the experiment on the NLCD maps, it was demonstrated that the proposed approach also works well for large regions with many LCLU classes. This will promote operational applications of the SPM technique in downscaling of TSIs covering large areas.

In the proposed spatio-temporal SPM approach, the temporal dependence is quantified by the similarity between images. The temporal aspect pushes the SPM realization of each image towards its temporally neighboring images. However, if abrupt LCLU changes occur in the TSIs (e.g., a pure coarse pixel at one time is changed to a mixed pixel covering multiple new classes at another time), the temporal neighbors cannot provide useful information on SPM of these classes within the pixel. Therefore, the new method is more appropriate for cases where smooth changes occur between images in the TSIs. This can be more realistic if the temporal resolution of the TSIs is fine enough. For example, MODIS images revisit the Earth's surface on a daily basis. The temporally dense TSIs are reliable sources for monitoring smooth changes such as a changed coastline caused by melting glaciers and reduced vegetation caused by illegal deforestation. By applying the proposed SPM approach to the coarse MODIS TSIs, monitoring such LCLU changes at a fine spatio-temporal resolution becomes possible. Note that the spatio-temporal SPM problem is different from that defined in some studies where sub-pixel shifted coarse time-series images were used to enhance SPM. In those studies, it was assumed that *no* LCLU changes occur between the utilized coarse time-series images [52], [53].

With respect to spatial dependence, both D_S^{SS} and D_S^{SP} were extended to spatio-temporal dependence in this paper, and correspondingly, two versions of spatio-temporal SPM models (i.e., SPT and SST) were developed. The differences in their accuracies are insignificant. This leaves the door open for alternative approaches to spatial dependence characterization. The SPT approach, however, quantifies spatial dependence only once. This eases the computational burden to some extent, which is an advantage over the SST approach. Obviously, more SPT approaches can be developed by considering other methods in measuring D_S^{SP}, such as the existing back-propagation neural network [25], support vector machine [54] and indicator cokriging [28]. They are implemented based on the availability of prior spatial structure information (or some other alternatives [30]) on the LCLU classes at the desired fine spatial resolution. It is also worthwhile to consider fusing D_S^{SP} and D_S^{SS} in spatial dependence characterization. However, such a scheme will introduce a new parameter that balances the influence of the two parts. Certainly, by treating the FSRM as the training data, the parameter can be estimated in a similar way to that for the weight balancing spatial and temporal dependences (see Section 2.7), but such a process will quadratically increase the computational complexity. Another issue is how important is the reliability of the spatial dependence characterization in the new method, where the FSRM provides reliable temporal information for the whole time-series. In cases where the temporal correlation in the TSIs is not large, the FSRM may not propagate very useful information and it would be necessary to study approaches capable of describing spatial dependence accurately (e.g., shape information [48] and additional data [47]-[51]).

In this paper, the optimal weight in (11) was estimated by using the whole FSRM as training data. To reduce the computational burden in this process, users may select as subset of pixels from the whole FSRM as training data. The optimal weight for the subset can be similarly estimated according to the procedures presented in Section 2.7, which can then be used for the entire FSRM. The training data need to be representative of the spatial structure across the whole scene. It is not clear whether this scheme will adversely affect the SPM accuracy of TSIs, especially for the spatially non-stationary scene.

The SPM of TSIs is modelled as a constrained optimization problem, under the coherence constraint imposed by the coarse proportions. The optimization problem can be solved by other artificial intelligence algorithms, such as a genetic algorithm [21] and particle swarm optimization [22]. Artificial intelligence is an active scientific topic in the field of computer science. However, whether an optimization algorithm is a preferable choice for the problem depends on the accuracy obtained and computational cost. It will be of great interest to explore more effective and faster optimization algorithms for the proposed spatio-temporal SPM model.

Spectral unmixing is a critical pre-processing step of SPM and its uncertainty can be propagated directly to the SPM. For TSIs, this step needs to be performed independently for each coarse image. The spectral reflectance of some materials, such as vegetation, may change across time, especially when the TSIs cover a long period. This necessitates extraction of reliable endmembers for each coarse image or construction of a reliable spectral library over a long time. All these considerations motivate future research.

V. Conclusion

In this paper, a spatio-temporal SPM approach was proposed to extract sub-pixel resolution LCLU information from TSIs. In the proposed approach, the objective of the SPM problem is to maximize the spatio-temporal dependence. The temporal aspect aims at maximizing the temporal similarity in LCLU between images, while the spatial aspect aims at maximizing the spatial correlation of LCLU within each image. The spatial dependence was characterized in two different ways: the relationship between sub-pixels and the relationship between a sub-pixel and its neighboring pixels. The PSA method was used for the former, while the SPSAM, Kriging, RBF methods were used for the latter. Correspondingly, two types of spatio-temporal SPM approaches, SST and SPT, were defined. The proposed SPM approach incorporates information from the FSRM in the TSIs. The SPM approach has several advantages: it accounts for spatial and temporal dependences simultaneously to deeply exploit information encapsulated in the TSIs; it can readily incorporate multi-resolution and multi-source data (such as the FSRM in this paper) to enhance SPM; the spatio-temporal dependence is quantified automatically, with weights estimated without manual intervention; based on SPM of TSIs, LCLU dynamic monitoring and change detection at a fine spatio-temporal resolution can be achieved.

The proposed approach was tested with three datasets, and the conclusions are summarized as follows.

1) The proposed approach can provide more accurate TSIs at sub-pixel resolution than conventional methods. Both SST and SPT are effective spatio-temporal SPM approaches.
2) The FSRM imparts greater benefits for SPM of images temporally close to it. The closer to the FSRM, the greater the SPM accuracy achieved.
3) The reliability of the FSRM is crucial in the proposed approach, and when the error in it was increased to a very large value (above 80%), the proposed method could not increase the SPM accuracy compared to conventional spatial-only methods.
4) However, for remarkably large error values (up to 80%), the proposed approach was able to impart benefits compared to conventional spatial-only methods.
5) When the FSRM is in error, the SST method is generally more accurate than the SPT approach, but the advantage was small and the SST approach has a much larger computational burden than the SPT approach.

Acknowledgment

The authors would like to thank Dr. Y. Xu of The Chinese University of Hong Kong, Hong Kong, for providing the Shenzhen Landsat dataset, and the handling editor and anonymous reviewers for their valuable and constructive comments which greatly improved the work.

REFERENCES

[1] R. S. Lunetta, J. F. Knight, J. Ediriwickrema, J. G. Lyon, and L. D. Worthy, "Land-cover change detection using multi-temporal MODIS NDVI data," *Remote Sensing of Environment*, vol. 105, pp. 142–154, 2006.

[2] G. M. Foody, "Status of land cover classification accuracy assessment," *Remote Sensing of Environment*, vol. 80, pp. 185–201, 2002.

[3] P. H. Swain, "Bayesian classification in a time varying environment," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 8, no. 12, pp. 879–883, 1978.

[4] L. Bruzzone, and S. B. Serpico, "An iterative technique for the detection of land-cover transitions in multitemporal remote-sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, no. 4, pp. 858–867, 1997.

[5] L. Bruzzone, D. F. Prieto, and S. B. Serpico, "A neural-statistical approach to multitemporal and multisource remote-sensing image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 3, pp. 1350–1359, 1999.

[6] B. Demir, F. Bovolo, and L. Bruzzone, "Detection of land-cover transitions in multitemporal remote sensing images with active-learning-based compound classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 5, pp. 1930–1941, 2012.

[7] A. H. S. Solberg, T. Taxt, and A. K. Jain, "A Markov random field model for classification of multisource satellite imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 34, no. 1, pp. 110–113, 1996.

[8] F. Melgani and S. B. Serpico, "A Markov random field approach to spatio-temporal contextual image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 11, pp. 2478–2487, 2003.

[9] D. Liu, M. Kelly, P. Gong, "A spatial–temporal approach to monitoring forest disease spread using multi-temporal high spatial resolution imagery," *Remote Sensing of Environment*, vol. 101, pp. 167–180, 2006.

[10] K. Bahirat, F. Bovolo, L. Bruzzone, and S. Chaudhuri, "A novel domain adaptation Bayesian classifier for updating land-cover maps with class differences in source and target domains," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 7, pp. 2810–2826, 2012.

[11] Y. Tarabalka, G. Charpiat, L. Brucker, and B. H. Menze, "Spatio-temporal video segmentation with shape growth or shrinkage constraint," *IEEE Transactions on Image Processing*, vol. 23, no. 9, pp. 3829–3840, 2014.

[12] D. Lu, M. Batistella, and E. Moran, "Multitemporal spectral mixture analysis for Amazonian land-cover change detection," *Canadian Journal of Remote Sensing*, vol. 30, no. 1, pp. 87–100, 2004.

[13] V. Haertel, Y. E. Shimabukuro, and R. Almeida, "Fraction images in multitemporal change detection," *International Journal of Remote Sensing*, vol. 25, no. 23, pp. 5473–5489, 2004.

[14] P. M. Atkinson, "Issues of uncertainty in super-resolution mapping and their implications for the design of an inter-comparison study," *International Journal of Remote Sensing*, vol. 30, no. 20, pp. 5293–5308, 2009.

[15] P. M. Atkinson, "Downscaling in remote sensing," *International Journal of Applied Earth Observation and Geoinformation*, vol. 22, pp. 106-114, 2013.

[16] K. C. Mertens, L. P. C. Verbeke, E. I. Ducheyne, and R. De Wulf, "Using genetic algorithms in sub-pixel mapping," *International Journal of Remote Sensing*, vol. 24, no. 21, pp. 4241–4247, 2003.

[17] Q. Wang, L. Wang, and D. Liu, "Particle swarm optimization-based sub-pixel mapping for remote-sensing imagery," *International Journal of Remote Sensing*, vol. 33, no. 20, pp. 6480–6496, 2012.

[18] P. M. Atkinson, "Sub-pixel target mapping from soft-classified, remotely sensed imagery," *Photogrammetric Engineering and Remote Sensing*, vol. 71, no. 7, pp. 839–846, 2005.

[19] A. Villa, J. Chanussot, J. A. Benediktsson, C. Jutten, and R. Dambreville, "Unsupervised methods for the classification of hyperspectral images with low spatial resolution," *Pattern Recognition*, vol. 46, pp. 1556–1568, 2013.

[20] A. J. Tatem, H. G. Lewis, P. M. Atkinson, and M. S. Nixon, "Super-resolution land cover pattern prediction using a Hopfield neural network," *Remote Sensing of Environment*, vol. 79, pp. 1–14, 2002.

[21] A. M. Muad and G. M. Foody, "Impact of land cover patch size on the accuracy of patch area representation in HNN-based super resolution mapping," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 5, pp. 1418–1427, 2012.

[22] Y. Zhong, Y. Wu, X. Xu, and L. Zhang, "An adaptive subpixel mapping method based on MAP model and class determination strategy for hyperspectral remote sensing imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 3, pp. 1411–1426, 2015.

[23] K. C. Mertens, B. D. Basets, L. P. C. Verbeke, and R. De Wulf, "A sub-pixel mapping algorithm based on sub-pixel/pixel spatial attraction models," *International Journal of Remote Sensing*, vol. 27, no. 15, pp. 3293–3310, 2006.

[24] Z. Mahmood, M. A. Akhter, G. Thoonen, and P. Scheunders, "Contextual subpixel mapping of hyperspectral images making use of a high resolution color image," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no, 2, pp. 779–791, 2013.

[25] K. C. Mertens, L. P. C. Verbeke, T. Westra, and R. De Wulf, "Sub-pixel mapping and sub-pixel sharpening using neural network predicted wavelet coefficients," *Remote Sensing of Environment*, vol. 91, pp. 225–236, 2004.

[26] Y. Shao and R. S. Lunetta, "Sub-pixel mapping of tree canopy, impervious surfaces, and cropland in the laurentian great lakes basin using MODIS time-series data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 4, no. 2, pp. 336–347, 2011.

[27] J. Verhoeye and R. De Wulf, "Land-cover mapping at sub-pixel resolutions using linear optimization techniques," *Remote Sensing of Environment*, vol. 79, no. 1, pp. 96–104, 2002.

[28] A. Boucher and P. C. Kyriakidis, "Super-resolution land cover mapping with indicator geostatistics," *Remote Sensing of Environment*, vol. 104, no. 3, pp. 264–282, 2006.

[29] H. Jin, G. Mountrakis, and P. Li, "A super-resolution mapping method using local indicator variograms," *International Journal of Remote Sensing*, vol. 33, no, 24, pp. 7747-7773, 2012.

[30] Q. Wang, P. M. Atkinson, and W. Shi, "Indicator cokriging-based subpixel mapping without prior spatial structure information," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 1, pp. 309–323, 2015.

[31] T. Kasetkasem, M. K. Arora, and P. K. Varshney, "Super-resolution land-cover mapping using a Markov random field based approach," *Remote Sensing of Environment*, vol. 96, no. 3/4, pp. 302–314, 2005.

[32] V. A. Tolpekin and A. Stein, "Quantification of the effects of land-cover-class spectral separability on the accuracy of Markov-random-field based superresolution mapping," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 9, pp. 3283–3297, 2009.

[33] X. Li, F. Ling, Y. Du, and Y. Zhang, "Spatially adaptive superresolution land cover mapping with multispectral and panchromatic images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 5, pp. 2810-2823, 2014.

[34] Y. F. Su, G. M. Foody, A. M. Muad, and K. S. Cheng, "Combining pixel swapping and contouring methods to enhance super-resolution mapping," *IEEE Journal of selected topics in applied earth observation and remote sensing*, vol. 5, no. 5, pp. 1428-1437, 2012

[35] Q. Wang, W. Shi, and L. Wang, "Allocating classes for soft-then-hard subpixel mapping algorithms in units of class," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 5, pp. 2940–2959, 2014.

[36] F. Ling, Y. Du, X. Li, W. Li, F. Xiao, and Y. Zhang, "Interpolation-based super-resolution land cover mapping," *Remote Sensing Letters*, vol. 4, no. 7, pp. 629-638, 2013.

[37] G. M. Foody and H. T. X. Doan, "Variability in soft classification prediction and its implications for sub-pixel scale change detection and super-resolution mapping," *Photogrammetric Engineering and Remote Sensing*, vol. 73, no. 8, pp. 923-933, 2007.

[38] F. Ling, W. Li, Y. Du, and X. Li, "Land cover change mapping at the subpixel scale with different spatial-resolution remotely sensed imagery," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, no. 1, pp. 182–186, 2011.

[39] Y. Xu and B. Huang, "A spatio–temporal pixel-swapping algorithm for subpixel land cover mapping," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 2, pp. 474–478, 2014.

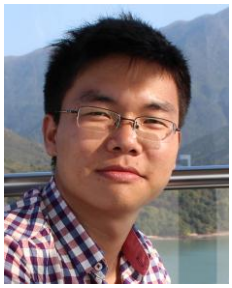
[40] X. Li, Y. Du, and F. Ling, "Super-resolution mapping of forests with bitemporal different spatial resolution images based on the spatial-temporal Markov random field," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 1, pp. 29-39, 2014.

[41] Q. Wang, W. Shi, P. M. Atkinson, and Z. Li, "Land cover change detection at subpixel resolution with a Hopfield neural network," *IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing*, vol. 8, no. 3, pp. 1339–1352, 2015.

[42] M. Q. Nguyen, P. M. Atkinson, and H. G. Lewis, "Superresolution mapping using a Hopfield neural network with LIDAR data," *IEEE Geoscience and Remote Sensing Letters*, vol. 2, no. 3, pp. 366–370, 2005.

[43] A. Boucher and P. C. Kyriakidis, "Integrating fine scale information in super-resolution land-cover mapping," *Photogrammetric Engineering and Remote Sensing*, vol. 73, no. 8, pp. 913-921, 2008.

[44] G. M. Foody, "Sharpening fuzzy classification output to refine the representation of sub-pixel land cover distribution," *International Journal of Remote Sensing*, vol. 19, no. 13, pp. 2593-2599, 1998.
[45] C. Huang, Y. Chen, and J. Wu, "DEM-based modification of pixel-swapping algorithm for enhancing floodplain inundation mapping," *International Journal of Remote Sensing*, vol. 35, no. 1, pp. 365–381, 2014.
[46] P. M. Atkinson, "Super-resolution mapping using the two-point histogram and multi-source imagery," in *GeoENV VI: Geostatistics for Environmental Applications*, pp. 307–321, 2008.
[47] P. Aplin and P. M. Atkinson, "Sub-pixel land cover mapping for per-field classification," *International Journal of Remote Sensing*, vol. 22, no. 14, pp. 2853–2858, 2001.
[48] F. Ling, X. Li, F. Xiao, S. Fang, and Y. Du, "Object-based sub-pixel mapping of buildings incorporating the prior shape information from remotely sensed imagery," *International Journal of Applied Earth Observation and Geoinformation*, vol. 18, pp. 283–292, 2012.
[49] Q. Wang, W. Shi, and P. M. Atkinson, "Sub-pixel mapping of remote sensing images based on radial basis function interpolation," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 92, 1–15, 2014.
[50] J. A. Benediktsson and P. H. Swain, "Consensus theoretic classification methods," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 22, no. 4, pp. 688–704, 1992.
[51] Y. Ge, "Sub-pixel land-cover mapping with improved fraction images upon multiple-point simulation," *International Journal of Applied Earth Observation and Geoinformation*, vol. 22, pp. 115–126, 2013.
[52] Y. Zhong, Y. Wu, L. Zhang, and X. Xu, "Adaptive MAP sub-pixel mapping model based on regularization curve for multiple shifted hyperspectral imagery," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 96, pp, 134–148, 2014.
[53] Y. Chen, Y. Ge, G. B. M. Heuvelink, J. Hu, and Y. Jiang, "Hybrid constraints of pure and mixed pixels for soft-then-hard super-resolution mapping with multiple shifted images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 5, pp. 2040–2052, 2015.
[54] Y. Zhang, Y. Du, F. Ling, S. Fang and X. Li, "Example-based super-resolution land cover mapping using support vector regression," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 4, pp. 1271–1283, 2014.



Qunming Wang (M'15) received the B.S. degree and M.S. degree from Harbin Engineering University, China, in 2010 and 2012, and the Ph.D. degree from The Hong Kong Polytechnic University, Hong Kong, in 2015.

He is now a Senior Research Associate in Lancaster Environment Centre, Lancaster University, U.K. From June to December 2013, he was a Visiting Ph.D. Student with Geography and Environment, University of Southampton, U.K. He has authored or coauthored over 25 peer-reviewed articles in international journals such as *Remote Sensing of Environment*, *IEEE Transactions on Geoscience and Remote Sensing*, and *ISPRS Journal of Photogrammetry and Remote Sensing*. His current research interests focus on remote sensing image analysis and geostatistics.

Dr. Wang serves as a reviewer for over ten international journals. He was awarded the hypercompetitive Hong Kong Ph.D. Fellowship to support his three-year Ph.D. study. He was a recipient of the Excellent Master Dissertation Award and the Excellent Graduates in Heilongjiang Province, China, in 2012.

Wenzhong Shi obtained the PhD degree from University of Osnabrück in Vechta, Germany, in 1994.

He is a Chair Professor in GIS and remote sensing, and the Head of Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University. His current research interests include GIS and remote sensing, uncertainty and spatial data quality control, image processing for high resolution satellite images. He has published over 130 SCI papers and 10 books. Prof. Shi received the State Natural Science Award from the State Council of China in 2007 and The Wang Zhizhuo Award from International Society for Photogrammetry and Remote Sensing in 2012.

Peter M. Atkinson received the BSc degree in Geography from the University of Nottingham in 1986 and the PhD degree from the University of Sheffield (NERC CASE award with Rothamsted Experimental Station) in 1990. More recently, he received the MBA degree from the University of Southampton in 2012.

He is currently the Dean of the Faculty of Science and Technology at Lancaster University. He has been Professor of Geography at the University Southampton (for the last 21 years; 13 as Professor), where he is currently Visiting Professor. He is also Visiting Professor at Queen's University Belfast. The main focus of his research is in remote sensing, GIS and spatial (and space-time) statistics applied to a range of environmental science and socio-economic problems. He has published around 200 peer-reviewed articles in international scientific journals. Prof. Atkinson is Associate Editor for *Computers and Geosciences* and sits on the editorial boards of several further journals including *Geographical Analysis*, *Spatial Statistics*, the *International Journal of Applied Earth Observation and Geoinformation*, and *Environmental Informatics*. He sits on various international scientific committees.