Fast and Slow Changes Constrained Spatio-temporal Subpixel Mapping

Chengyuan Zhang, Qunming Wang, Ping Lu, Yong Ge, Peter M. Atkinson

Abstract— **Subpixel mapping (SPM) is a technique to tackle the mixed pixel problem and produce land cover and land use (LCLU) maps at a finer spatial resolution than the original coarse data. However, uncertainty exists unavoidably in SPM, which is an ill-posed downscaling problem. Spatio-temporal SPM methods have been proposed to deal with this uncertainty, but current methods fail to explore fully the information in the time-series images, especially more rapid changes over a short-time interval. In this paper, a fast and slow changes constrained spatio-temporal subpixel mapping (FSSTSPM) method is proposed to account for fast LCLU changes over a short-time interval and slow changes over a long-time interval. Namely, both fast and slow change-based temporal constraints are proposed and incorporated simultaneously into the FSSTSPM to increase the accuracy of SPM. The proposed FSSTSPM method was validated using two synthetic datasets with various proportion errors. It was also applied to oil-spill mapping using a real PlanetScope-Sentinel-2 dataset and Amazon deforestation mapping using a real Landsat-MODIS dataset. The results demonstrate the superiority of FSSTSPM. Moreover, the advantage of FSSTSPM is more obvious with an increase in proportion errors. The concepts of the fast and slow changes, together with the derived temporal constraints, provide a new insight to enhance SPM by taking fuller advantage of the temporal information in the available time-series images.**

Index Terms—**Land cover and land use (LCLU), subpixel mapping (SPM), super-resolution mapping, spatio-temporal dependence, downscaling, Hopfield neural network (HNN).**

I. INTRODUCTION

Land cover and land use (LCLU) patterns are a manifestation of land surface activities on *in situ* properties and, thus, reflect local balances between environmental and ecological value and economic development [1], [2]. Continuous monitoring of LCLU at fine spatial resolution can provide significant information for extensive applications such as urban planning, disaster assessment and agricultural yield forecasting [3]-[6]. However, rapid changes on the surface of the Earth have increased the requirement for a more timely and precise LCLU change detection. An increasing number of satellites have been launched to acquire remote sensing images for global

monitoring. However, the trade-off between fine spatial resolution and fine temporal resolution of the images has brought great challenges for timely and precise monitoring [7]. For example, the Landsat sensors can provide 30 m images but revisit the same area every 16 days. On the contrary, the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor covers the same scene daily, but with a coarser spatial resolution of 500 m [8]. As a result, mixed pixels (pixels containing more than one type of land cover class) exist ubiquitously in the coarse imagery.

LCLU mapping has been an important topic for long and various methods have been proposed [9]-[11]. Conventional LCLU mapping is implemented by hard classification at the pixel level [12]. That is, each pixel is assigned to a single land cover class. Hence, hard classification fails to fully represent the information in mixed pixels [13]. In the recent decades, spectral unmixing and subpixel mapping (SPM) have been developed and viewed as effective solutions to tackle the mixed pixel problem. Specifically, spectral unmixing estimates the class composition (i.e., known as proportions) of mixed pixels, while SPM predicts the spatial distribution of classes within each mixed pixel by separating each coarse pixel into $s \times s$ (s is the zoom factor) subpixels and assigning each subpixel a specific land cover class. Therefore, SPM provides a useful solution for continuous monitoring of LCLU at a finer spatial resolution [14], which can inherit the fine temporal resolution of the input images (e.g., daily access to MODIS images) simultaneously.

SPM is undertaken on the basis of the assumption of spatial dependence (i.e., spatially neighboring observations are considered more possible to be of the same LCLU class). In general, current SPM methods can mainly be summarized into two types: subpixel-to-subpixel-based and subpixel-to-pixel-based methods. According to the description of spatial dependence between observations (i.e., pixels and subpixels), spatial dependence for the subpixel-to-subpixel group is depicted by the relation between the subpixel and its surrounding subpixels. Solutions in this category contain the pixel swapping algorithm (PSA) [15], genetic algorithm [16], Hopfield neural network (HNN) [17] and maximum a posteriori (MAP) [18]. For the subpixel-to-pixel group, the spatial dependence refers to the relation between the subpixel and its adjacent pixels. Common methods of this type include radial basis function (RBF) interpolation [19], subpixel/pixel spatial attraction model [20], area-to-point kriging [21], back-propagation neural network [22], indicator co-kriging [23] and double-calculated spatial attraction model [24]. Moreover, both types of spatial dependence can be combined for SPM [25].

It is noted that the coarse proportion is used as coherence constraint for most SPM methods (both subpixel-to-subpixel-based and subpixel-to-pixel-based methods mentioned above) that are performed as the

C. Zhang, Q. Wang and P. Lu are with the College of Surveying and Geo-Informatics, Tongji University, 1239 Siping Road, Shanghai 200092, China (e-mail: wqm11111@126.com).

Y. Ge is with the State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

P.M. Atkinson is with the Faculty of Science and Technology, Lancaster University, Lancaster LA1 4YR, UK, with Geography and Environment, University of Southampton, Highfield, Southampton SO17 1BJ, UK and with the Chinese Academy of Sciences, 11A Datun Rd, Beijing, China.

post-processing of spectral unmixing. That is, in SPM predictions, the number of subpixels for each class needs to be in agreement with the coarse proportions. Since spectral unmixing remains an open problem, the proportion error is propagated inevitably into the proportion-maintained SPM results. To mitigate the heavy reliance on proportions, spatial-spectral SPM models have been developed by considering the spectral-based coherence constraint and the spatial dependence-based goal simultaneously [26]. The Markov random field (MRF) [27] is a typical spatial-spectral SPM method. However, it remains a challenge to estimate automatically the size of spatial neighborhood [28] and the parameter that controls the spatial and spectral parts, since the accuracy of SPM is considerably sensitive to these parameters.

SPM is a typical ill-posed problem: multiple land cover maps at the target fine spatial resolution can fulfill the coherence constraint exerted by the coarse proportions [29]. To address this issue, various auxiliary data have been considered to decrease the uncertainty in SPM, such as vector data [30], color images [31], panchromatic images [32], training images [33] and subpixel shifted images [34], [35] and point data [36]. Learning-based SPM methods have been also developed by using fine spatial resolution training data [37-40]. However, with respect to the strong demand to produce similar spatial structure to the area of interest, these ancillary data are generally laborious to acquire.

Benefiting from the periodic revisit ability of satellites, multi-temporal remote sensing images are available from various platforms. To borrow the information from temporally close images, several spatio-temporal SPM methods have been developed, which are promising solutions for continuous LCLU mapping at fine spatial resolution. In these spatio-temporal SPM methods, the spatial term (i.e., spatial dependence) is characterized based on either subpixel-to-subpixel-based or subpixel-to-pixel-based model. Ling et al. [41] first introduced a subpixel land cover change mapping algorithm, which incorporates a pre-date land cover map at the target fine spatial resolution. Xu and Huang [42] extended conventional PSA to a spatio-temporal PSA (STPSA) model. Li et al. [43] presented a spatio-temporal HNN (STHNN) method. Using a pre-date fine spatial resolution map, STHNN identifies the changed coarse proportions and incorporates the temporal information of unchanged pixels in the HNN. Wang et al. [14] developed a general spatio-temporal SPM framework for time-series coarse image, which is suitable for most SPM methods (i.e., both subpixel-to-subpixel-based or subpixel-to-pixel-based methods). He et al. [44] incorporated a pre-date fine spatial resolution map-derived temporal regularization term in a spectral-spatio-temporal MAP-based model. The above spatio-temporal SPM methods used one pre-image or post-image. For more reliable LCLU mapping of historical images, several studies have been developed by using both pre- and post-images. Li et al. [45] used two images in the temporal constraint term in spatio-temporal SPM. Using pre-date and post-date coarse and fine spatial resolution proportion pairs, Zhang et al. [46] introduced a spatio-temporal proportion map fusion method to predict the fine spatial resolution image on the middle date.

Amongst the existing studies on SPM of time-series images, although the subpixel map is predicted by referring to pre-image

or post-image or both images, temporal correlation is explained by assuming that LCLU change is a consecutive process. However, in view of the period of land cover changes in time-series images, sometimes the change in LCLU can occur rapidly within a short period (i.e., no LCLU changes in other periods) instead of a long period (i.e., from the beginning to the end of the investigated period). Taking the slow change case as an example, a road in a developed urban area may remain unchanged for several years (e.g., from pre-image to post-image). Instead, in the case of fast changes, land cover may change in the first short period and then remain unchanged in the second short period, such as bare land appearing in the weeks during deforestation (e.g., from pre-image to target image) but failing to recover during subsequent seasons (e.g., from target image to post-image). Therefore, for SPM, it is appealing to exploit the unchanged information for the temporal constraints in both long and short-time intervals.

In this paper, a fast and slow changes constrained spatio-temporal SPM method (FSSTSPM) is proposed to enhance SPM by characterizing slow and fast LCLU changes in the long and short-time intervals simultaneously. FSSTSPM is constructed based on the HNN model with the availability of fine spatial resolution maps acquired on the dates before and after the coarse image. The fine spatial resolution maps can be acquired conveniently from a variety of data sources, for example, classification of remote sensing images available in the investigated period. Different from the STHNN proposed by Li et al. [43] that integrates a pre-date fine spatial resolution land cover map, FSSTSPM employs fast change-based temporal constraint in short-time interval and slow change-based temporal constraint in long-time interval at the subpixel resolution simultaneously, which is also the most important advantage over existing studies that commonly use both pre- and post-images. Another advantage inherited from the HNN is that FSSTSPM can mitigate the heavy reliance on the coarse proportion constraint and preserve the small-sized features simultaneously. This can help to address the uncertainty in the coarse proportions.

The main contributions of this paper are four-fold:
1) Theoretically, the fast change-based temporal constraint is proposed for spatio-temporal SPM, which uses unchanged information within a short period. Meanwhile, the slow change-based temporal constraint is integrated with the fast change-based temporal constraint to produce a more reliable temporal constraint in SPM.
2) Technically, the idea mentioned above is realized by the FSSTSPM method. FSSTSPM is performed directly on each subpixel. Moreover, it inherits the advantage of HNN, which can reduce the uncertainty in the coarse proportions.
3) For application, FSSTSPM was used to monitor events of great concern. Specifically, FSSTSPM was undertaken to fine spatial and temporal resolution monitoring of the well-known deforestation in the Amazon rainforest and the most recent oil-spill incident on the Ambarnaya River which occurred on May 29, 2020.
4) For utilization of data, FSSTSPM was examined using various datasets from different satellite sensors, such as PlanetScope data, which has received increasing attention recently [47]-[50].

The remainder of this paper is organized as follows. In section II, the mechanism of HNN is described briefly, and then the proposed energy function of the FSSTSPM model is introduced explicitly, followed by the details of the fast and slow change-based temporal constraints provided by the multi-temporal images. Section III provides experimental results for synthesized and real datasets. In Section IV, open issues on FSSTSPM and its applications are discussed, followed by the conclusion in Section V.

II. METHODS

A. HNN-based SPM

The proposed FSSTSPM method is constructed on the basis of HNN [51]. In the HNN, each subpixel is viewed as a neuron and each neural network layer represents a LCLU class. Referring to the coarse proportion images (i.e., spectral unmixing predictions), the input of the network layer is initialized as the input signal u_{ijk}, where i and j denote the row and column of the subpixels in the layer, respectively. Each neuron is updated iteratively to produce the final prediction v_{ijk}, which can be viewed as a self-optimization process without the need of any extra training data, as shown in Fig. 1.
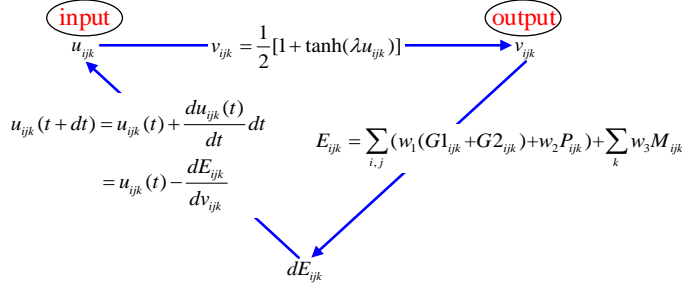

Fig. 1. The iteration process of neuron (i, j, k) for HNN.

With the input signal u_{ijk}, the output signal v_{ijk} is determined as

$$v_{ijk} = \frac{1}{2}[1 + \tanh(\lambda u_{ijk})] \quad (1)$$

where parameter λ determines the steepness of the tanh function, and is always set to 10 to ensure that the $\tanh(\cdot)$ term ranges [-1, 1]. More details can be found in [17]. In each iteration, the output v_{ijk} is used to update dynamically the network energy change dE_{ijk} of the neuron. For the t-th iteration, the input of the neuron (i, j, k) is updated as

$$u_{ijk}(t+dt) = u_{ijk}(t) + \frac{du_{ijk}(t)}{dt}dt = u_{ijk}(t) - \frac{dE_{ijk}}{dv_{ijk}} \quad (2)$$

where dt is an iteration step length, and the network energy function in HNN is characterized as

$$E_{ijk} = \sum_{i,j}(w_1(G1_{ijk} + G2_{ijk}) + w_2 P_{ijk}) + + \sum_k w_3 M_{ijk} \quad (3)$$

where w_1, w_2, and w_3 are three weights, $G1_{ijk}$ and $G2_{ijk}$ are two spatial constraint functions characterized by the spatial dependence assumption (or *prior* spatial structure information, if available). Specifically, $G1_{ijk}$ and $G2_{ijk}$ are determined by the

output signal of the eight neighboring neurons of v_{ijk}. If the average of the outputs of the neighboring neurons is greater than 0.5, the output of the center neuron is pushed towards 1. In contrast, if the average of the outputs of the neighbors is smaller than 0.5, the output of the center neuron is pushed towards 0.

In Eq. (3), P_{ijk} is the proportion constraint and is expressed as

$$\frac{dP_{ijk}}{dv_{ijk}} = \frac{1}{2s^2}\sum_{d=x\cdot s}^{x\cdot s+s-1}\sum_{e=x\cdot s}^{y\cdot s+s-1}\left[1 + \tanh\left(v_{dek} - 0.5\right)\lambda\right] - F_k(V) \quad (4)$$

where x and y denotes the coordinates of the coarse pixel V that subpixel at (i, j) falls in. $F_k(V)$ is the coarse proportion of class k for pixel (x, y), and s is the zoom factor. If the coarse proportion calculated in the current version is larger than $F_k(V)$, v_{ijk} decreases to make the current coarse proportion approach an ideal coarse proportion.

Moreover, M_{ijk} in Eq. (3) is the multi-class constraint, which is defined to make the sum of v_{ijk} $(k = 1, 2, ..., K)$ for all classes equal to 1. When the energy change dE_{ijk} reaches zero or is extremely small, the output v_{ijk} is assumed to converge to a stable solution. Afterwards, the subpixel located at (i, j) is assigned to class k if the output v_{ijk} is 1 (or very close to 1). Alternatively, the subpixel is allocated to another class if the output is 0.

Note that other than SPM methods which fulfill entirely the coarse proportions, the proportion constraint in the HNN is approached, but not perfectly satisfied. As a result, isolated error pixels (e.g., subpixels caused by proportion errors) can be mitigated by the HNN.

B. The proposed FSSTSPM energy function

Considering slow and fast LCLU change processes, FSSTSPM aims to take full advantage of unchanged information in the long and short-time intervals during the investigated period, by constructing fast and slow change-based temporal constraints. To fully exploit the information in the pre- and post-images, the following energy function is constructed in FSSTSPM

$$ME_{ij} = E_{ijk} + \sum_{i,j}\sum_k w_4(T1_{slow} + T2_{slow} + T3_{fast} + T4_{fast}) \quad (5)$$

where ME_{ijk} is the energy function of FSSTSPM, and it considers spatial and proportion constraints in the term of E_{ijk} defined in Eq. (3), and moreover, the proposed multi-temporal constraint. Specifically, $T1_{slow}$, $T2_{slow}$, $T3_{fast}$ and $T4_{fast}$ represent four terms of the proposed multi-temporal constraint (i.e., two terms of the slow change-based temporal constraint representing a slow LCLU change process in a long period and two terms of the fast change-based temporal constraint representing a fast LCLU change process in a short period), which will be introduced explicitly in the Sections II-C and D. w_4 is a weight coefficient.

The methodology of the proposed multi-temporal constraint in FSSTSPM is elaborated in Fig. 2. FSSTSPM adopts one coarse proportion image predicted by spectral unmixing on the target date and two fine spatial resolution land cover maps that pre-date and post-date the target image. The coarse proportion

images on the pre-date and post-date can be obtained by spatial degradation of the two land cover maps with the zoom factor s. Subsequently, FSSTSPM constructs slow change-based temporal constraint in the long period (i.e., from the pre-date to the post-date) and fast change-based temporal constraint in the short period (i.e., from the pre-date to the target date or the target to the post-date). More details about the formulation of the temporal constraints are elaborated in Sections II-C and D.
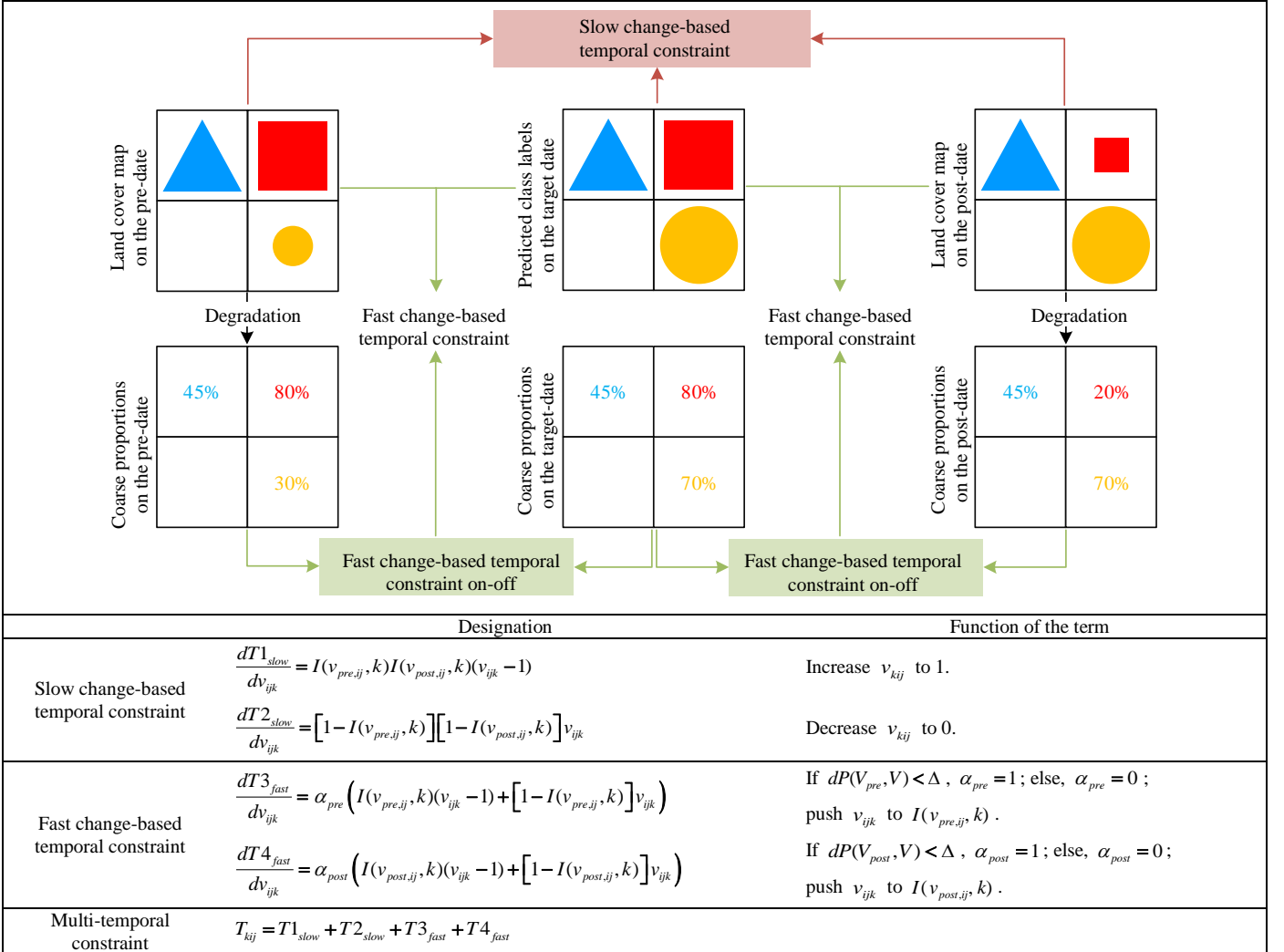

		Designation	Function of the term
Slow change-based temporal constraint		$\dfrac{dT1_{slow}}{dv_{ijk}} = I(v_{pre,ij},k)I(v_{post,ij},k)(v_{ijk}-1)$	Increase v_{kij} to 1.
		$\dfrac{dT2_{slow}}{dv_{ijk}} = \left[1-I(v_{pre,ij},k)\right]\left[1-I(v_{post,ij},k)\right]v_{ijk}$	Decrease v_{kij} to 0.
Fast change-based temporal constraint		$\dfrac{dT3_{fast}}{dv_{ijk}} = \alpha_{pre}\left(I(v_{pre,ij},k)(v_{ijk}-1)+\left[1-I(v_{pre,ij},k)\right]v_{ijk}\right)$	If $dP(V_{pre},V)<\Delta$, $\alpha_{pre}=1$; else, $\alpha_{pre}=0$; push v_{ijk} to $I(v_{pre,ij},k)$.
		$\dfrac{dT4_{fast}}{dv_{ijk}} = \alpha_{post}\left(I(v_{post,ij},k)(v_{ijk}-1)+\left[1-I(v_{post,ij},k)\right]v_{ijk}\right)$	If $dP(V_{post},V)<\Delta$, $\alpha_{post}=1$; else, $\alpha_{post}=0$; push v_{ijk} to $I(v_{post,ij},k)$.
Multi-temporal constraint		$T_{kij}=T1_{slow}+T2_{slow}+T3_{fast}+T4_{fast}$	

Fig. 2. The full designation of the multi-temporal constraint for a neuron in FSSTSPM.

C. The proposed slow change-based temporal constraint

Following the assumption that the class tends to be unchanged for the image on the target date if the land cover changes slowly or remains unchanged in the long period (i.e., from pre-date to post-date of the target date), the two terms of the slow change-based temporal constraint are expressed as

$$\frac{dT1_{slow}}{dv_{ijk}} = I(v_{pre,ij},k)I(v_{post,ij},k)(v_{ijk}-1) \qquad (6)$$

$$\frac{dT2_{slow}}{dv_{ijk}} = \left[1-I(v_{pre,ij},k)\right]\left[1-I(v_{post,ij},k)\right]v_{ijk} \qquad (7)$$

in which $v_{pre,ij}$ and $v_{post,ij}$ are temporally neighboring subpixels of v_{ijk} located spatially at (i, j) in the pre- and post-images, respectively. Moreover, $I(v_{pre,ij},k)$ and $I(v_{post,ij},k)$ are two class indicator functions. Taking the former as an example, it is defined as

$$I(v_{pre,ij},k)=\begin{cases}1, & \text{if } v_{pre,ij} \text{ belongs to class } k \\ 0, & \text{otherwise}\end{cases}. \qquad (8)$$

In the slow change-based temporal constraint, $T1_{slow}$ or $T2_{slow}$ is activated only if $I(v_{pre,ij},k)=I(v_{post,ij},k)$, which means that the indicator of v_{ijk} in the long-time interval (from the pre-date to the post-date) is very likely to be unchanged, and the process can be further viewed as a slow change case. Specifically, for a subpixel at (i, j) and falling within a coarse pixel V, if the class labels of the subpixels in the pre- and post-fine land cover maps both belong to class k (i.e., $I(v_{pre,ij},k)=1$ and $I(v_{post,ij},k)=1$), the temporal constraint $T1_{slow}$ pushes v_{ijk} to 1 while $T2_{slow}$ exerts no effect on v_{ijk}. Otherwise, if $I(v_{pre,ij},k)=0$ and $I(v_{post,ij},k)=0$, the temporal constraint $T2_{slow}$ pushes v_{ijk} to 0 while $T1_{slow}$ exerts no effect

D. The proposed fast change-based temporal constraint

The land cover sometimes changes fast in a short period (i.e., from the pre-date to the target date or the target to the post-date) instead of a long period. That is, the land cover remains unchanged in other short periods during the entire period. Accordingly, the fast change-based temporal constraint is proposed, which is expressed as

$$\frac{dT3_{fast}}{dv_{ijk}} = \alpha_{pre}\left(I(v_{pre,ij},k)(v_{ijk}-1)+\left[1-I(v_{pre,ij},k)\right]v_{ijk}\right) \quad (9)$$

$$\frac{dT4_{fast}}{dv_{ijk}} = \alpha_{post}\left(I(v_{post,ij},k)(v_{ijk}-1)+\left[1-I(v_{post,ij},k)\right]v_{ijk}\right) \quad (10)$$

in which α_{pre} and α_{post} are two on-off switches guided by the temporal difference in proportions. More specifically, if the difference in coarse proportions between the coarse spatial resolution pixel V and the pre-date pixel V_{pre} at the same location is less than a threshold Δ (i.e., $dP(V_{pre},V)<\Delta$), the fast change-based temporal constraint in Eq. (9) is activated (i.e., $\alpha_{pre}=1$). This means that the land cover distribution in pixel V is very likely to be unchanged compared to that of V_{pre} (see, for example, the 'red rectangle' in Fig. 2). Otherwise, if $dP(V_{pre},V)\geq\Delta$, the fast change-based temporal constraint in Eq. (9) is not activated (i.e., $\alpha_{pre}=0$), which indicates that a fast change is occurred from the pre-date to the target date, such as that for the 'orange circle' in Fig. 2. This is also the same for the determination of the on-off switches in Eq. (10).

In the fast change-based temporal constraint, for a subpixel (i, j) falling in coarse pixel V, if the proportion difference $dP(V_{pre},V)$ is less than a threshold, v_{ijk} is assumed to resemble the land cover class label on the pre-date (i.e., no changes from the pre-date to the target date). That is, $T3_{short}$ takes a positive effect on v_{ijk} to approach $I(v_{pre,ij},k)$. Similarly, if the proportion difference $dP(V_{post},V)$ is less than the threshold, $T4_{short}$ takes a positive effect on v_{ijk} to approach $I(v_{post,ij},k)$.

E. Benchmark methods

The performance of FSSTSPM was examined by comparing with six SPM methods, including PSA, RBF, HNN, STPSA, spatio-temporal RBF (STRBF) [14] and STHNN [43]. Specifically, different from the original STPSA method proposed by Ling et al. [52] and Xu and Huang [42] that used only pre-date imagery, STPSA in the experiments in this paper incorporated both pre- and post-images for fair comparison. Moreover, to validate the benefit of considering the fast change-based temporal constraint, an incomplete version of FSSTSPM, namely, FSSTSPM(S) (i.e., FSSTSPM with only the slow change-based temporal constraint), was also compared, that is, FSSTSPM(S). SPM is essentially an ill-posed issue and is always implemented by exploiting various constraints on predictions. Therefore, the differences between spatial, proportion and temporal constraints for the seven benchmark methods and the proposed FSSTSPM method are listed in Table 1.

Table 1 The SPM methods used in the experiments

Constraint	Spatial		Proportion		Temporal	
Methods	Subpixel-to-subpixel	Subpixel-to-pixel	Strictly	Not strictly	Single image	Multiple images
PSA	√		√			
RBF		√	√			
HNN	√			√		
STPSA	√		√			√
STRBF		√	√			√
STHNN	√			√	√	
FSSTSPM(S)	√			√		√
FSSTSPM	√			√		√

1) Spatial constraint: PSA- and HNN- based methods (i.e., PSA, HNN, STPSA, STHNN and FSSTSPM) consider the spatial dependence between subpixels and subpixels. RBF-based methods (i.e., RBF and STRBF) are performed by considering the spatial dependence between the center subpixel and the adjacent coarse pixels.

2) Proportion constraint: For methods where the proportion constraint is strictly satisfied (i.e., PSA, RBF, STPSA and STRBF), the proportion is treated as ideal and the number of subpixels for each land cover class is fixed. In contrast, the proportions are not maintained perfectly during the optimization process in HNN-based methods including HNN, STHNN and FSSTSPM.

3) Temporal constraint: Spatio-temporal SPM methods (i.e., STPSA, STRBF, STHNN and FSSTSPM) are degraded to traditional SPM methods (i.e., PSA, RBF and HNN) when the temporal constraint is removed. In view of the spatio-temporal methods, one fine spatial resolution image is used in STHNN while both pre-date and post-date fine spatial resolution maps are used in STPSA, STRBF and FSSTSPM.

III. EXPERIMENTS

Experiments on four datasets were undertaken to examine the proposed FSSTSPM approach, including two synthetic datasets and two real datasets. In the experiments on the two synthetic datasets, a group of manually drawn images and a group of National Land Cover Database (NLCD) land cover maps were used. For the first real experiment, we used PlanetScope and Sentinel-2 images (PlanetScope-Sentinel-2) for a recent oil-spill incident. For the second real experiment, we used Landsat 5 Thematic Mapper images and MODIS images (Landsat-MODIS) for a deforestation case in Amazon rain forest. It should be noted that images from the same sensor were co-registered already in the published products. Meanwhile, images from different sensors (e.g., Landsat and MODIS images on the target date) were co-registered manually. Additionally, for the PlanetScope-Sentinel-2 dataset, the Sentinel-2 image was upscaled to 30 m to simulate a Landsat-like image, where a zoom factor of 10 for SPM was considered to produce a 3 m land cover map. As stated in Section II-C, seven SPM methods were considered for comparison, including PSA, RBF, HNN, STPSA, STRBF, STHNN and FSSTSPM(S). In the experiments, the parameters for all methods were set empirically or based on the suggestions of existing literature. Specifically, the weights in the energy functions for HNN, STHNN and FSSTSPM were equally

set to 1. The window size was set to 3×3 for all methods. Moreover, the threshold Δ was set to 20% empirically for the proposed FSSTSPM. It should be noted, however, Δ should be set case by case in practice, which can be affected greatly by the errors in proportions.

### A. Experiments utilizing synthetic datasets

1) Manually drawn images

In this experiment, three fine spatial resolution images (128 by 128 pixels) covering two objects (i.e., tree-like and road-like objects) were drawn manually, which represent the pre-image at T1, reference at T2 and post-image at T3, as displayed in Fig. 3. In the simulated time-series dataset, both objects remain unchanged in a short-time interval. Specifically, the left road-like object remains unchanged from T1 to T2 and the right tree-like object remains unchanged from T2 to T3. The input coarse proportions were simulated by degrading the image at T2 with two factors (i.e., s=4 and 8). To evaluate the influence of uncertainties in the coarse proportions on SPM, the proportions with various errors were simulated by adding a Gaussian distributed error to the ideal proportions [43]. The error of the simulated proportions was evaluated by the index of root mean square error (RMSE). Accordingly, three proportion images were produced, with RMSE=0.05, 0.10 and 0.15, respectively.

Fig.3 shows the SPM results of all methods for both s=4 and s=8, where proportions with RMSE=0 and 0.10 were considered. It is worth noting that when the RMSE is 0.10, the resultant PSA, RBF, STPSA and STRBF method maps contain a number of sporadic artifacts. The HNN-based methods (i.e., STHNN+T1, STHNN+T3, FSSTSPM(S) and FSSTSPM) produce cleaner results compared to the PSA- and RBF- based methods. This is because the PSA and RBF methods preserve strictly the class proportion in the SPM predictions, resulting in inaccurately labeled subpixels when proportion error exists.

Without temporal information, for PSA and RBF, the road-like features are aggregated incorrectly into patches, especially for a large zoom factor of s=8. Furthermore, the boundaries of the tree-like object are over-smoothed in the PSA and RBF results. For HNN, some elongated features disappeared because of the spatial smoothing effect. For the temporal-based

methods, the incorporation of fine spatial resolution maps can produce more satisfactory SPM results. Specifically, when using the T1 image for STHNN (i.e., STHNN+T1), the prediction of the road-like object that remained unchanged from T1 to T2 is closer to the reference than STHNN+T3. Inversely, for STHNN+T3, the reproduction of the tree-like object that remained unchanged from T2 to T3 is more accurate than STHNN+T1. This is because the STHNN method adopts a fast change-based temporal constraint, and if the coarse proportion is unchanged, this method outputs a comparable class distribution to that of the auxiliary fine spatial resolution map. Moreover, by utilizing slow change-based temporal constraints in the FSSTSPM(S) method, the unchanged objects from T1-T3 are reconstructed; however, land cover that is unchanged over a short interval is not well preserved. For the proposed FSSTSPM method, fast and slow change-based temporal constraints are considered simultaneously, and the available temporal information in both pre- and post-images is utilized. As a result, the proposed FSSTSPM method produces the most accurate results amongst all eight methods. The comparison between the FSSTSPM(S) and FSSTSPM methods reveals the importance of the fast change-based temporal constraint.

Fig. 4 presents the quantitative assessment results of the overall accuracy (OA) for the eight methods for s=4 and 8 under various RMSEs. It turns out that the inaccurate proportions impact SPM negatively and the OA values of all the methods decrease with an increase in the proportion error. Furthermore, as the RMSE increases, the OA values of the HNN-based methods decrease much more slowly than for the PSA- and RBF-based methods, because the HNN-based methods release the proportion constraints and are, therefore, less affected by the proportion error. Moreover, the FSSTSPM method is more accurate than the STHNN+T1 and STHNN+T3 methods. However, the FSSTSPM(S) method produces a smaller OA than the STHNN method, suggesting that its advantage in considering both fast and slow change-based temporal constraints in SPM. In summary, the accuracy of FSSTSPM is consistently greater than that of the other methods in the various cases.
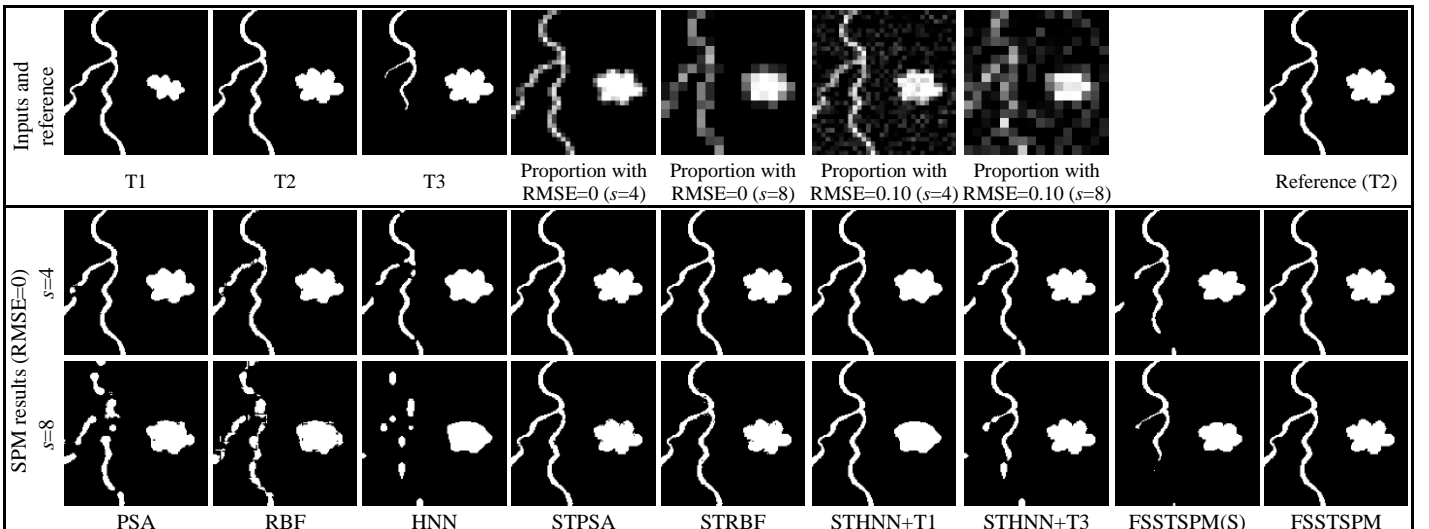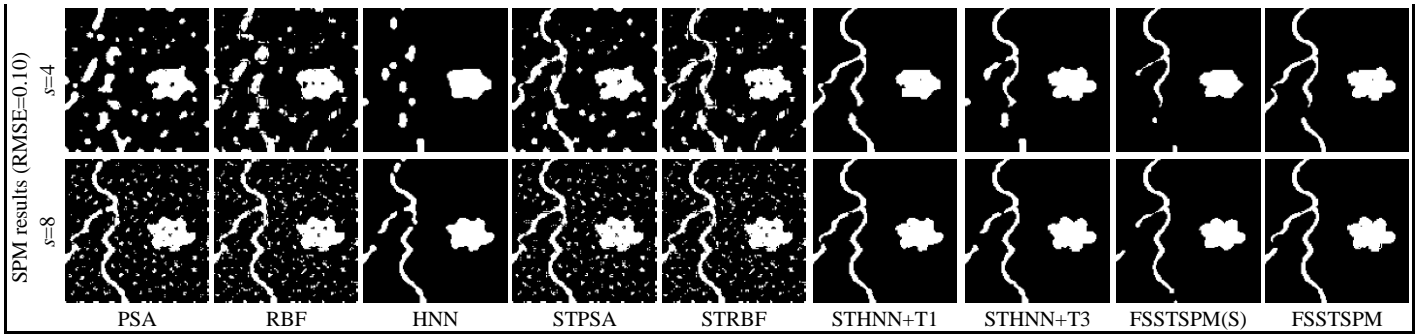
Fig. 3. SPM results of different methods for the manually drawn dataset (*s*=8).
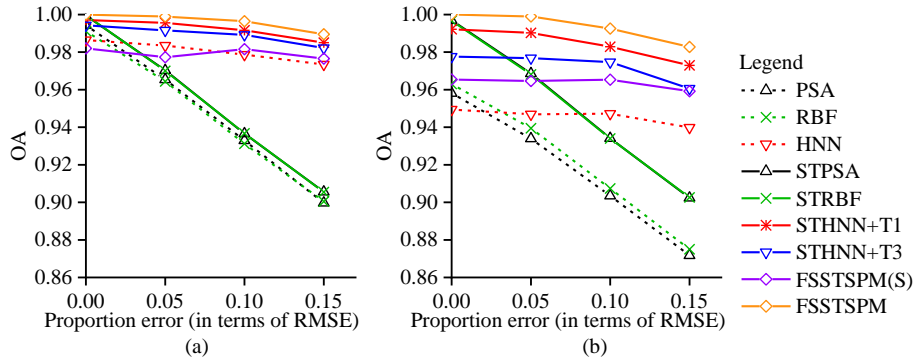

Fig. 4. OA of different SPM methods for the manually drawn images with different proportion errors. (a) *s*=4. (b) *s*=8.

2) NLCD land cover maps

In this experiment, the pre-date, reference, and post-date fine spatial resolution land cover maps were acquired from the 2001, 2006 and 2011 NLCD, respectively. The NLCD is a 30 m raster-based land-cover classification database covering the 50 U.S. states and Puerto Rico. As exhibited in Fig. 5, the study area is located in South Carolina, and each map covers 320×320 pixels, presenting four classes: urban, agriculture, water, and forest. The goal was to reproduce the 30 m NLCD map in 2006, where the 2001 and 2011 NLCD maps were used as pre- and post-images in SPM, respectively. The spectral unmixing results, (i.e., proportion images) in 2006 were produced by degrading the binary map of each class in the 2006 NLCD map. Moreover, proportions errors with RMSE=0.05, 0.10 and 0.15 were considered in the simulation process. Two zoom factors (*s*=4 and 8) were utilized to simulate 120 m and 240 m coarse proportion images, respectively.

The resultant maps of all the methods utilizing simulated proportion images for *s*=8 and RMSE=0 and 0.10 are presented in Fig. 5. STHNN+2001 and STHNN+2011 denote the STHNN method utilizing the 2001 and 2011 NLCD maps for temporal constraints, respectively. Without utilizing the land cover spatial information in the pre-date and post-date maps, the resultant PSA, RBF and HNN method maps are significantly different from the reference. Conversely, the temporal-based SPM methods, based on spatial and temporal dependences, produce more detailed and visually more compelling boundaries.

Subareas of the resultant maps are enlarged in lines 4 and 6 of Fig. 5 to provide a clearer difference between these methods. The results of the six temporal-based methods (i.e., STPSA, STRBF, STHNN+2001, STHNN+2011, FSSTSPM(S), and FSSTSPM) vary with the RMSE. When the RMSE is 0, the STPSA, STRBF, and FSSTSPM method SPM predictions are visually close to the reference. For the resultant STHNN+2001,

STHNN+2011, and FSSTSPM(S) method maps, the urban class boundary is more compact than the reference, and some linear-shaped objects are merged into patches. When the RMSE is 0.10, apparent speckle artifacts are present in the STPSA and STRBF results. Although the proportion error is spatially eliminated in the STHNN+2001 and STHNN+2011 methods, some patches are aggregated due to the proportion errors in the classes. For example, the urban class is contaminated by other classes, and some agricultural patches are reduced in size. Notably, even with a proportion error, the proposed FSSTSPM method produces visually satisfactory predictions, with most features analogous to those in the reference. Furthermore, the FSSTSPM method restores more details by utilizing the fast change-based temporal constraint than the FSSTSPM(S) method.

Figs. 6 and 7 show the producer's accuracy (PA) and OA for all eight methods, respectively. Based on the quantitative evaluation results, three observations can be made, which are described below.

First, consistent with the visual evaluation from Fig. 5, the accuracies of the STPSA, STRBF, STHNN, and FSSTSPM methods are greater than those of the PSA, RBF and HNN methods, indicating that the resultant SPM map is more accurate when temporal constraints are considered. Moreover, the proposed FSSTSPM method exhibited the greatest accuracy of all eight methods, with PA and OA values exceeding 0.96 and 0.93 for *s*=4 and 8, respectively.

Second, as shown in Fig. 7, the OA values for all the methods decrease gradually with increasing RMSE, and the OA value of FSSTSPM is the largest amongst all the methods for each RMSE value. For the methods incorporating temporal constraints, the OA decreases much faster in the STPSA and STRBF methods than the FSSTSPM method, indicating that the FSSTSPM method can mitigate the negative effect of the proportion error
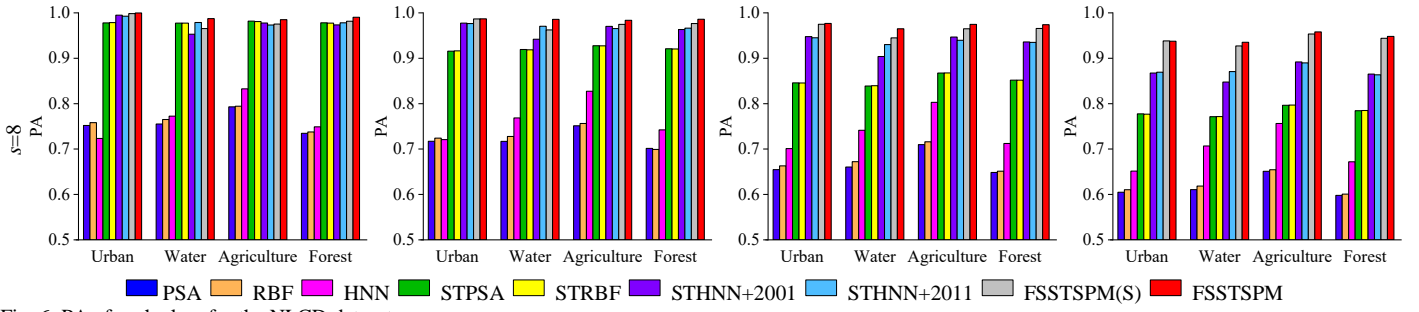
more effectively than the other methods. For example, when the RMSE increases from 0 to 0.15 for s=8, the OA of the FSSTSPM method decreases from 0.99 to 0.95, while those of the STPSA and STRBF methods decrease from 0.98 to 0.79.

Third, the advantage of utilizing fast change-based temporal constraints is obvious for the proposed FSSTSPM method compared with the FSSTSPM(S) method. For example, when the RMSE is 0 for s=4, the PA values of the FSSTSPM method are approximately 0.01, 0.02, 0.01, and 0.01 larger than those of the FSSTSPM(S) method for the urban, agriculture, water and forest classes, respectively.

In summary, the qualitative and quantitative evaluations demonstrate that FSSTSPM produces the most accurate SPM predictions amongst the eight SPM methods. The comparison between FSSTSPM(S) and FSSTSPM suggests that the fast change-based temporal constraint effectively enhances the SPM results. Moreover, the advantage of FSSTSPM is more obvious with increasing RMSE in the proportions compared with other methods, suggesting that the proposed FSSTSPM method is less sensitive to spectral unmixing accuracy.

Fig. 5. SPM results of different methods for the NLCD dataset (s=8).

Fig. 6. PA of each class for the NLCD dataset.
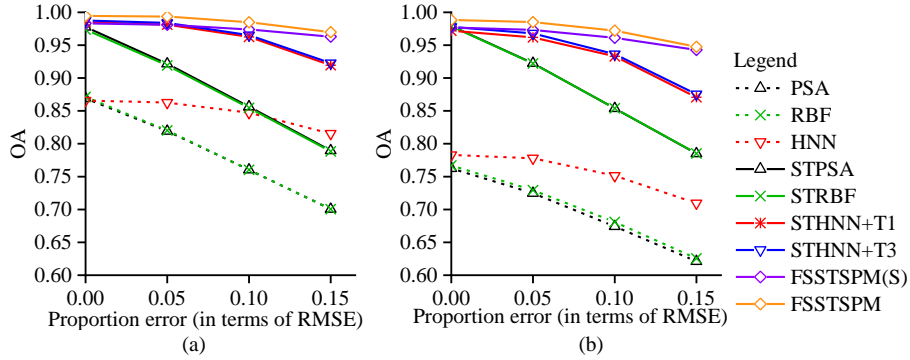
Fig. 7. OA of different SPM methods for the NLCD dataset with different proportion errors. (a) *s*=4. (b) *s*=8.

B. Experiments utilizing real datasets

1) PlanetScope-Sentinel-2 oil-spill event experiment

The PlanetScope-Sentinel-2 dataset was utilized to validate the proposed FSSTSPM method on an oil-spill event in the Ambarnaya River in Norilsk, Russia. The event occurred on May 29, 2020 and was cleaned immediately afterwards, leading to dynamic changes in distribution of oil-spill, especially in the first several weeks. However, it should be noted that the total clean up would take several years. Three 3 m PlanetScope images were acquired on May 30 and 31 and June 4, 2020 (T1-T3, respectively). Land cover maps with oil and non-oil classes were drawn manually from the images T1-T3, and utilized as pre-image, reference, and post-image, respectively. Additionally, one 10 m Sentinel-2 image acquired on May 31, 2020 was utilized as the coarse image. The coarse proportion images were produced via the linear spectral unmixing model (LSMM) based on the 10 m Sentinel-2 image with 210×210 pixels, which were utilized as the input for the SPM process. The performance of FSSTSPM was further evaluated utilizing a larger zoom factor of 10 (e.g., scale difference between the PlanetScope and widely utilized Landsat images). Specifically, the 10 m Sentinel-2 image was upscaled to 30 m (the spatial resolution of the Landsat images) to simulate a coarse image with 70×70 pixels, and the resultant proportion image was input into the SPM process. The 3 m PlanetScope images, 3 m land cover maps, and 30 m coarse proportion image are presented in line 1 of Fig. 8. The 3 m land cover map was restored utilizing different SPM methods and zoom factors of 3 and 10, as presented in lines 3 and 5, respectively, of Fig. 8. Subareas of the maps in lines 3 and 5 (Fig. 8) are enlarged to provide an enhanced visual comparison and are presented in lines 4 and 6, respectively, of Fig. 8.

The SPM methods produce more detailed oil boundaries than the coarse images, and many sporadic artifacts were included in the PSA, RBF, STPSA and STRBF method results. Notably, although fine spatial resolution map pairs were utilized in the STPSA and STRBF methods, the resultant maps are not noticeably different from those of the PSA and RBF methods because the PSA and RBF methods are slavish to the coarse proportions and, thus, the proportion uncertainties are identified as noisy pixels. Moreover, smoother boundaries are produced via the HNN-based methods (HNN, STHNN+T1, STHNN+T3, and FSSTSPM), especially in non-oil areas. However, the oil class boundaries in the HNN, STHNN+T1, and STHNN+T3 predictions are wavy and differ significantly from the reference, as shown in line 6 of Fig. 8. The boundaries of the FSSTSPM prediction are more continuous and more accurate than those of the other methods, with a few speckles presented in the predictions for *s*=3 and 10.

Fig. 9 shows the error maps of the SPM predictions presented in Fig. 8. Many errors can be observed, including many oil pixels predicted as non-oil pixels and a few non-oil pixels predicted as oil pixels. The main reason is that the errors in the coarse proportions reduce the accuracy of the pure non-oil pixels. Generally, there are fewer pixel errors in the resultant FSSTSPM maps, suggesting that this method can provide a more reliable class boundary restoration than the other methods.

The quantitative evaluation results (i.e., OA and omission and commission errors) for all the methods are listed in Table 2. Although the STHNN+T1 and STHNN+T3 methods produce larger OAs than the other methods, the proposed FSSTSPM method, with OA values of 98.42% and 98.85% for *s*=3 and 10, respectively, is the most accurate method, attributed to its ability to consider the fast and slow change-based temporal constraints simultaneously. In conclusion, based on the visual and statistical accuracy assessments, the proposed FSSTSPM method produces the most accurate SPM predictions.
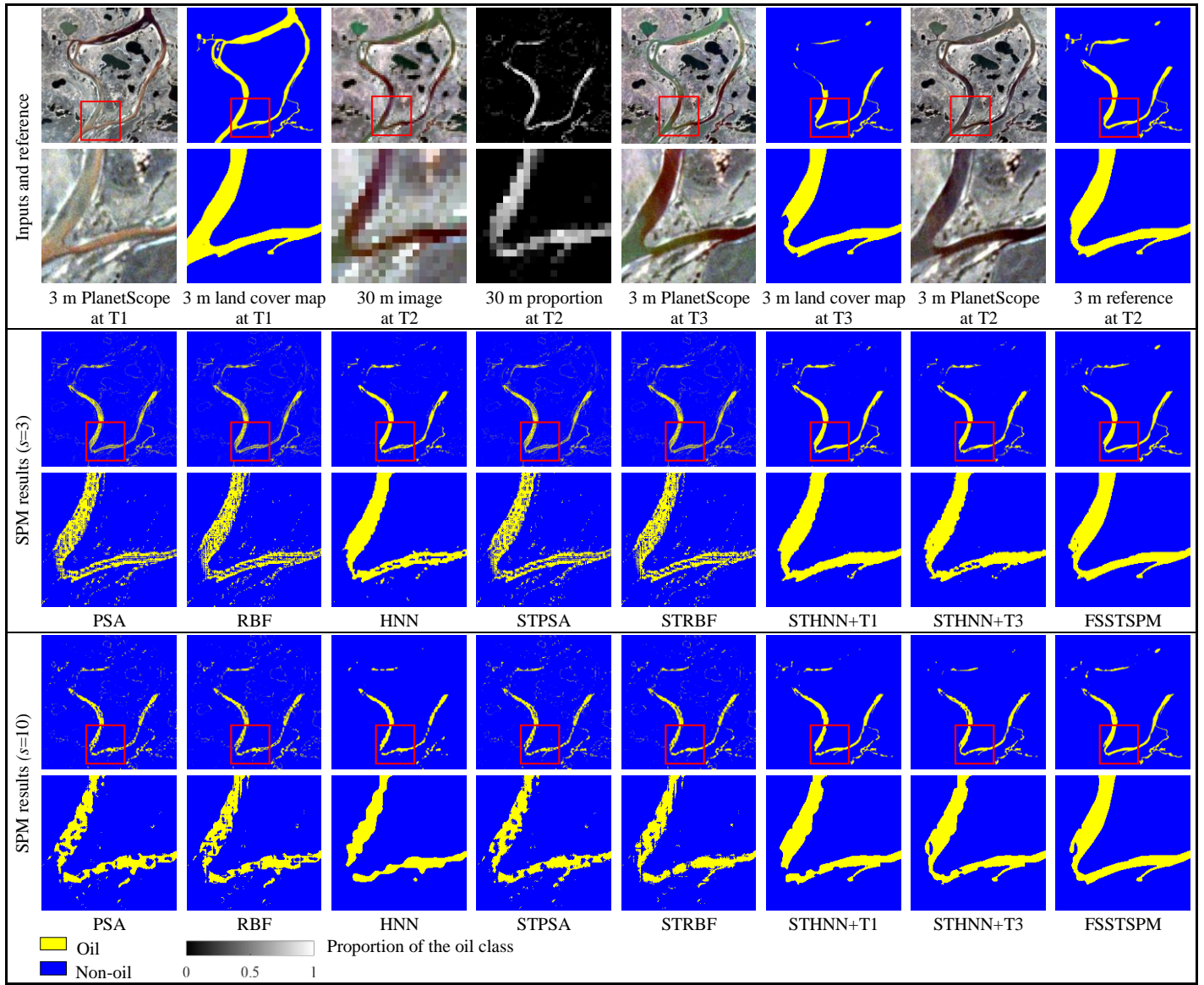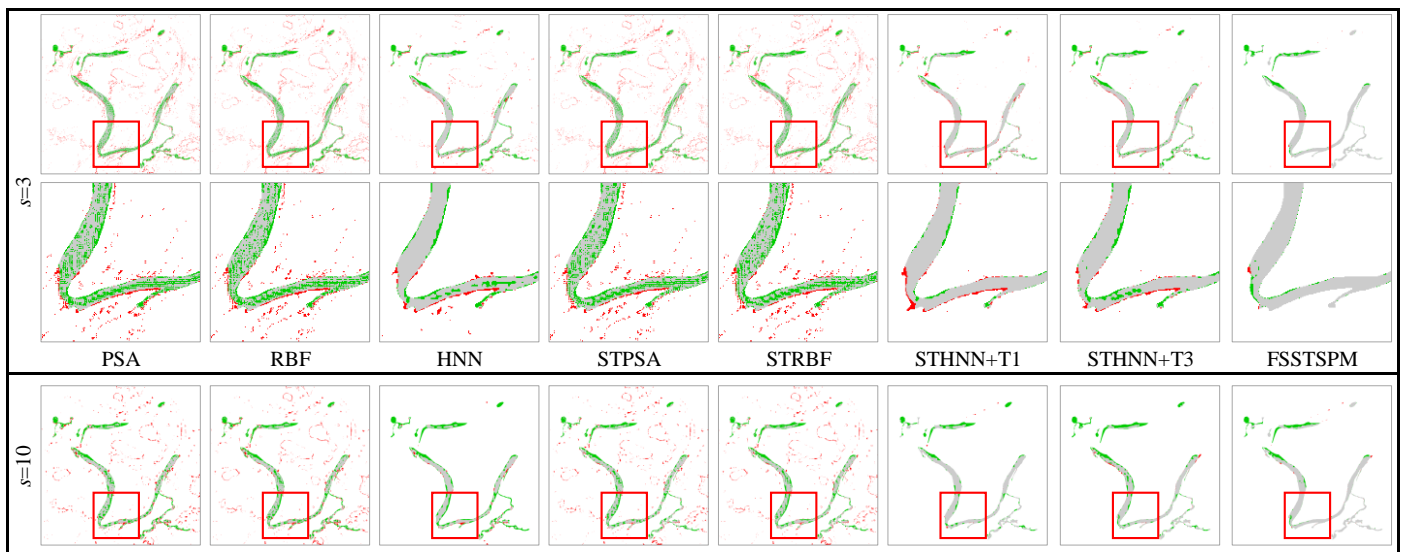
Fig. 8. SPM results of different methods for the PlanetScope-Sentinel-2 dataset (*s*=3 and 10).
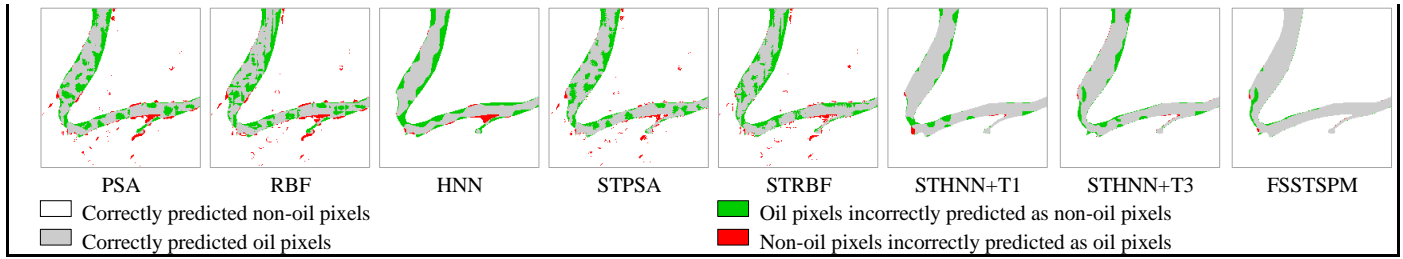
Fig. 9. Error maps for the PlanetScope-Sentinel-2 dataset (*s*=3 and 10).

Table 2 The omission error (%), commission error (%) and OA (%) of different methods for the PlanetScope-Sentinel-2 dataset

			PSA	RBF	HNN	STPSA	STRBF	STHNN+T1	STHNN+T3	FSSTSPM
s=3	Oil	Commission error	23.74	23.71	17.52	30.32	30.23	13.33	8.99	**1.06**
		Omission error	8.38	8.36	7.15	12.88	12.82	2.78	3.25	**1.86**
	Non-oil	Commission error	13.87	13.84	12.30	21.30	21.20	4.68	5.68	**3.37**
		Omission error	15.03	15.01	10.44	19.20	19.14	8.22	5.22	**0.58**
		OA	85.39	85.41	88.89	80.05	80.12	93.05	94.61	**98.42**
s=10	Oil	Commission error	2.44	2.45	2.46	2.21	2.20	1.37	1.82	**1.04**
		Omission error	1.52	1.53	0.24	1.28	1.27	0.04	0.12	**0.05**
	Non-oil	Commission error	31.37	31.53	6.96	26.41	26.31	0.89	2.98	**1.09**
		Omission error	42.62	42.76	43.47	38.48	38.39	23.92	31.96	**18.21**
		OA	96.23	96.22	97.39	96.69	96.70	98.65	98.14	**98.95**

*2) Landsat-MODIS deforestation event experiment*

An experiment was performed for the deforestation event in the Amazon utilizing three 30 m Landsat 5 TM images acquired on July 21, 2003 (T1), August 8, 2004 (T2) and August 11, 2005 (T3), respectively. These images were classified into two classes (i.e., forest and non-forest) via a support vector machine and employed as the pre-date, reference, and post-date fine spatial resolution land cover maps. Additionally, a single eight-day, 467 m MODIS product (MOD09A1) acquired on August 4, 2004 was pre-projected and resampled to 480 m. Then, a coarse proportion image was produced via LSMM for utilization as the SPM input. As shown in Fig. 10, the study area exceeds 5800 km^2, and includes 2880×2240 fine spatial resolution Landsat pixels and 180×140 coarse MODIS pixels. The SPM methods were compared visually based on the resultant and error maps presented in Fig. 10.

Based on the results, it turns out that SPM produces a more detailed land cover map in comparison with the coarse proportion image. As shown in the enlarged subareas in line 4 of Fig. 10, many small, jagged artifacts exist in the non-forest patches of the PSA, RBF, STPSA and STRBF maps. By exploiting temporal information from the pre-date or post-date fine spatial resolution maps, the FSSTSPM+T1 and FSSTSPM+T3 methods produce clearer results than the previous four methods, but fail to restore the boundaries of the

non-forest class smoothly. The prediction of FSSTSPM is visually the most similar to the reference. By comparing the error maps of the different methods, FSSTSPM is found to produce fewer error pixels than the other methods. However, error pixels (marked in red in Fig. 10) occur in all the results because significant changes occurred from T1 to T3. Thus, sufficient boundary information in the temporal constraints is unavailable for the SPM process. Furthermore, the non-forest class spectrum varies greatly, causing large proportion errors.

As displayed in Table 3, FSSTSPM produces the smallest omission and commission errors for both classes. For example, the commission errors are 4.82% and 4.12% for the forest and non-forest classes, respectively. Furthermore, the OA of the proposed FSSTSPM method is the largest amongst the eight methods.

The uncertainties in classification of fine spatial resolution images at T1 and T3 were also considered in this experiment. Specifically, 5%, 10% and 15% incorrectly classified pixels were simulated in both land cover maps at T1 and T3. As presented in Fig. 11, all STSPM methods (i.e., STPSA, STRBF, STHNN+T1, STHNN+T3 and FSSTSPM) produce smaller OA with an increase of classification error. Under the same classification error, however, the FSSSTSPM method produces the largest OA amongst all methods.
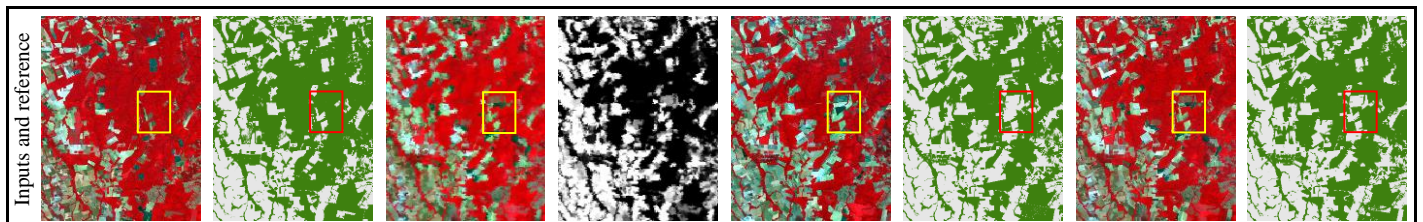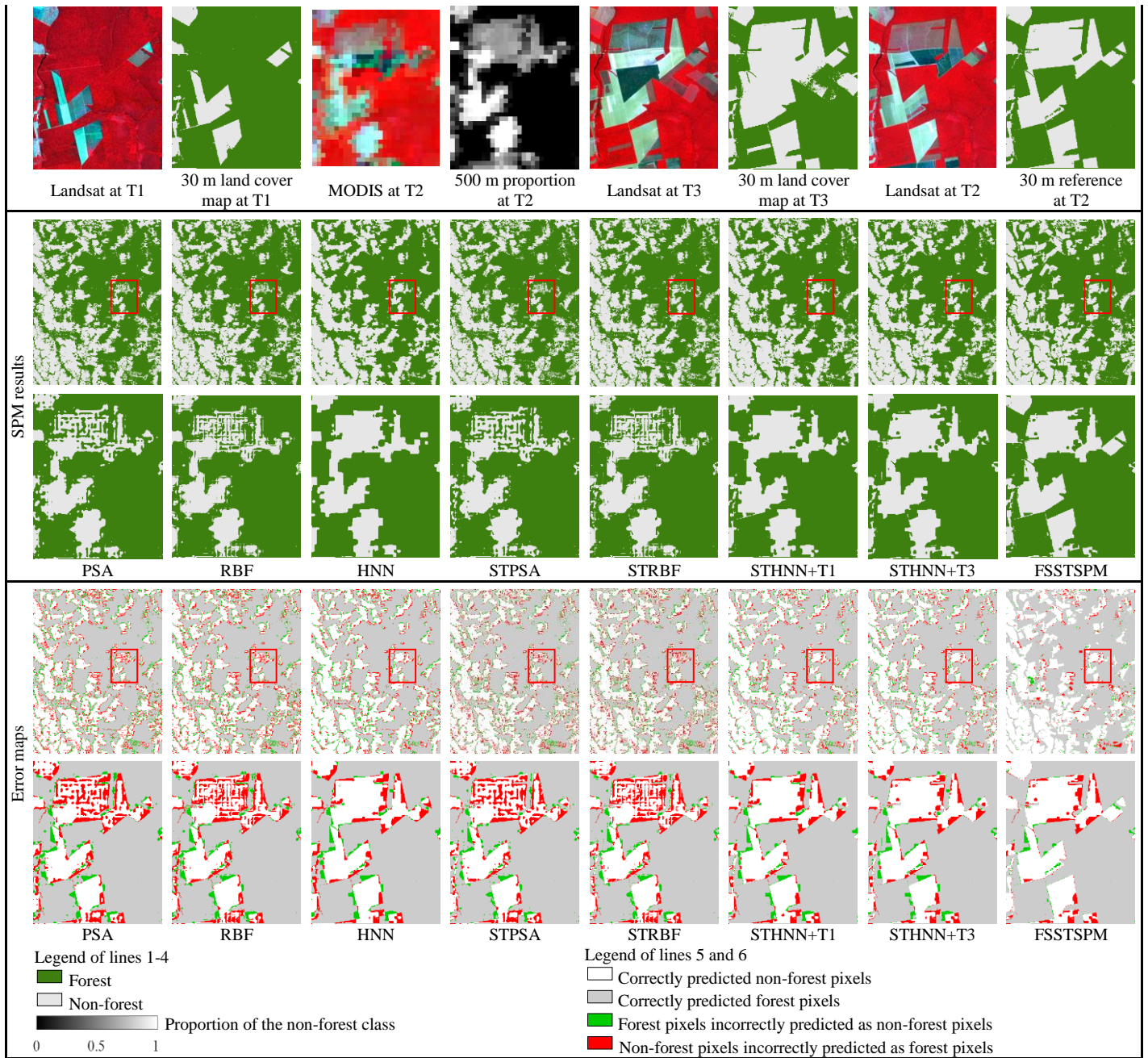
Fig. 10. SPM results of different methods for the Landsat-MODIS dataset (*s*=16).

Table 3 The omission error (%), commission error (%) and OA (%) of different methods for the Landsat-MODIS dataset

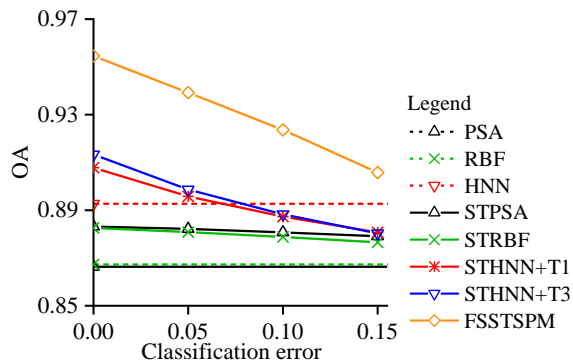|			PSA	RBF	HNN	STPSA	STRBF	STHNN+T1	STHNN+T3	FSSTSPM
s=16	Forest	Commission error	12.57	12.48	9.42	11.17	11.22	8.35	7.80	**4.82**
		Omission error	9.79	9.70	8.77	8.35	8.40	7.25	6.88	**2.80**
	Non-forest	Commission error	14.61	14.47	12.62	12.46	12.53	10.50	9.94	**4.12**
		Omission error	18.47	18.34	13.51	16.42	16.49	12.03	11.22	**7.01**
		OA	86.63	86.73	89.27	88.32	88.26	90.78	91.33	**95.46**

Fig. 11. OA of different SPM methods for the Landsat-MODIS dataset with various classification errors in the 30 m land cover maps at T1 and T3.

IV. DISCUSSION

A. Differences between spatio-temporal fusion and the proposed FSSTSPM method

The three main differences between the proposed FSSTSPM and existing spatio-temporal fusion methods are discussed below.

First, conceptually, the proposed FSSTSPM method aims to produce land cover classification maps instead of multispectral images. Therefore, both coarse image downscaling and land cover mapping are performed in FSSTSPM. Conversely, spatio-temporal fusion is designed to produce temporally continuous multispectral images at fine spatial resolution, although the output can be used for further land cover mapping.

Second, although coarse and fine spatial resolution images are required for both methods, they have different functions. For the proposed FSSTSPM method, only one coarse multispectral image is needed to acquire the class proportions based on spectral unmixing, and the pre-date and post-date fine spatial resolution land cover maps are utilized to construct temporal constraints in SPM. In contrast, spatio-temporal fusion generally requires at least one coarse-fine image pair pre- or post- the date of the coarse image to predict the temporal spectral change of each band. Consequently, consistent spectral ranges for each band between the fine and coarse images are essential for spatio-temporal fusion, whereas the proposed FSSTSPM method is able to utilize fine spatial resolution images with different spectral ranges. For example, temporally neighboring, fine spatial resolution hyperspectral images can also be utilized to obtain land cover maps at target fine spatial resolution for the FSSTSPM method.

Third, as an SPM method, FSSTSPM is developed to predict more detailed land cover maps than the original coarse image, and it can be applied to detect subpixel resolution land cover changes across time. That is, it is suitable for dynamic monitoring of LCLU changes. In contrast, spatio-temporal fusion is generally undertaken based on the assumption of no LCLU changes between images on different dates, and the prediction is more suitable for dynamic monitoring of phenological changes of vegetation (e.g., crop growth).

B. Uncertainty in the fast change-based temporal constraint

The experimental results indicate that a fast change-based constraint can facilitate spatio-temporal SPM by exploiting temporal change information in a short-time interval. It should

be noted that the performance of fast and slow change-based constraint is affected essentially by the speed of land cover change, that is, not necessarily the length of time interval between images. For example, land cover may change fast within several days (e.g., inundation) or change slow within up to one year (e.g., urbanization in highly developed cities). Generally, for the same area, the land cover change tends to be smaller if the temporal distance between two images is also smaller. Thus, for implementation of FSSTSPM, we normally need to search for the fine spatial resolution images that are temporally closest to the target coarse image. However, uncertainty exists inevitably in this constraint. In the proposed FSSTSPM method, the short interval temporal change is detected based on the difference between the coarse proportions at two times (one is the estimated coarse proportion based on spectral unmixing of coarse images at the prediction time, while the other is the coarse proportion degraded from the pre-date or post-date fine spatial resolution land cover maps). Generally, spectral unmixing assumes that the coarse proportions in the coarse pixel are the average of the subpixel indicators within the pixel. However, the actual proportion is influenced by various factors, such as noise and the point spread function. The latter means that the signal of a pixel is influenced by its adjacent pixels. Therefore, uncertainty exists in the proportion predictions. To account for the uncertainty in the short interval constraint, an adaptive model may be considered by including an adaptive weight for each coarse pixel based on the proportion difference at the two times. This is also our ongoing research.

C. Generalization ability of the proposed FSSTSPM method

In this study, the proposed FSSTSPM method was applied in the PlanetScope-Sentinel-2 and Landsat-MODIS dataset experiments and was demonstrated to be more accurate than the benchmark methods. As mentioned in Section IV-A, the proposed FSSTSPM method can utilize images with different spectral ranges. The proposed FSSTSPM method can be applied to scenes with temporal land cover changes. More importantly, it is a method suitable for various datasets. For example, it can be applied for SPM of 300 m Sentinel-3 time-series, with the aid of 10 m Sentinel-2 images on the pre-date and post-date, to reproduce temporally continuous, 10 m fine spatial resolution land cover maps.

D. Limitation of the proposed FSSTSPM

The proposed FSSTSPM method is developed based on the availability of two fine spatial resolution land cover maps that pre-date and post-date the target coarse spatial resolution image. The slow change-based temporal constraint is constructed using pre- and post-images directly. Thus, FSSTSPM is suitable for handling historic time-series images, rather than real-time images (i.e., the post-fine land cover map is unknown). It would be an interesting avenue to develop methods applicable to SPM of real-time data.

Moreover, to extract reliable unchanged information in the long and short periods, the fine spatial resolution land cover maps closer to the target date are more favorable for FSSTSPM. Nevertheless, due to revisit frequency and cloud contamination, sometimes the fine spatial resolution data may be temporally far from the target time, where land cover may experience great changes. Facing with this issue, it may be worthwhile to perform

SPM of coarse images closest to the fine spatial resolution data first and the prediction is then used as auxiliary data for SPM of the next coarse image. The process is repeated until the target time is reached. However, errors would be propagated and accumulated to the final predictions on the target date. The benefits of this potential scheme would depend heavily on the speed of land cover changes across time, and the scheme seems more suitable for gradual changes.

E. The common problem of intra-class spectral variation

In the Landsat-MODIS dataset experiment, the forest and non-forest class spectra were extracted for utilization in the simple, widely employed LSMM model, and the same estimated coarse proportion input was utilized in all the methods. Intra-class spectral variation is one of the most challenging issues in spectral unmixing, and the accuracy of spectral unmixing tends to decrease with an increase in intra-class spectral variation. It would be interesting to develop more reliable solutions to deal with the unavoidable intra-class spectral variation to reduce the uncertainties in the coarse proportions, and further, increase the SPM accuracy. In our previous research [53], a spatio-temporal spectral unmixing (STSU) approach was developed to explore fully the available temporal information in the multi-scale time-series images, which does not require endmembers and can cope with intra-class spectral variation. Specifically, supervised information (i.e., spectra of coarse pixel coupled with the corresponding land cover proportions) at the prediction time is extracted by change detection, which is used to fit a learning model for subsequent proportion prediction. STSU, however, needs coarse images at both the pre-date and post-date for change detection. More importantly, STSU is carried out assuming that the land cover changes are unidirectional during the investigated period.

F. Application to oil-spill mapping

In the oil-spill experiment, oil-spill mapping was utilized to detect sudden oil-spill events in Russia. This experiment revealed the noticeable capability of SPM to identify more detailed oil-polluted areas than traditional classification methods. Moreover, despite the general utilization of synthetic aperture radar (SAR) imagery, optical images have been demonstrated to be more effective for monitoring oil presence, type, and quality [54]. Consequently, SPM has great potential for monitoring oil-spill areas experiencing rapid land cover changes, which can be undertaken without the need for images with both fine spatial and temporal resolution. For example, Rajendran et al.[55] used historical images with relatively coarse spatial resolution from May 23 to June 8 for analyzing the oil-spill event on the Ambarnaya river. By applying spatio-temporal SPM (e.g., FSSTSPM developed in this paper) to those multi-temporal coarse images, oil velocity and direction can be further analyzed at a finer spatial resolution, which may contribute more to oil cleanup and underwater biological conservation. Furthermore, the oil-spill pollution primarily exists above the water surface. Therefore, the water mask can be distilled to confine the oil mapping boundary (i.e., pixels outside the water mask can be directly assigned to the non-oil class) such that commission errors resulting from the land surface can be avoided.

V. Conclusion

SPM has been acknowledged widely as an ill-posed issue. The use of temporal information is an important solution to overcome this limitation. Existing spatio-temporal SPM methods generally assume that LCLU change is a continuous process, and only the slow change-based temporal constraint is considered. In some cases, however, the LCLU change can be considered as a break-point process with fast change in a short period, rather than change in the entire period. In existing methods, the valuable information in the short period is always ignored. In this research, the fast change-based temporal constraint was defined to fully explore the information in the time-series, and further, both slow and fast change-based temporal constraints were considered in SPM to increase the prediction accuracy. The proposed FSSTSPM method was compared to the PSA, RBF, HNN, STPSA, STRBF and STHNN methods based on the experiments on four datasets, including two synthetic and two real datasets (for oil-spill and deforestation mapping using real datasets). The main findings are summarized briefly as below.

1) Visual and quantitative assessments revealed that FSSTSPM can provide more accurate SPM prediction than the benchmark methods.
2) In FSSTSPM, the use of the fast change-based temporal constraints is advantageous for enhancing SPM.
3) As the error (in terms of RMSE) in the proportions increases, the SPM accuracy decreases for all methods. However, the proposed FSSTSPM method remains the most accurate in each case, being less sensitive to the proportion error and decreasing more slowly with proportion error than the other methods.
4) As concluded from the PlanetScope-Sentinel-2 and Landsat-MODIS experiments, the proposed FSSTSPM method can produce satisfactory subpixel maps for the key applications of oil-spill monitoring and deforestation monitoring with fine spatio-temporal resolution.

References

[1] C. Homer, J. Dewitz, S. Jin et al., "Conterminous united states land cover change patterns 2001-2016 from the 2016 national land cover database," *ISPRS Journal of Photogrammetry and Remote Sensing,* vol. 162, pp. 184-199, 2020.

[2] C. Zhang, P. A. Harrison, X. Pan et al., "Scale sequence joint deep learning (ss-jdl) for land use and land cover classification," *Remote Sensing of Environment,* vol. 237, pp. 111593, 2020.

[3] D. R. Richards, B. S. Thompson, and L. Wijedasa, "Quantifying net loss of global mangrove carbon stocks from 20 years of land cover change," *Nature Communications,* vol. 11, no. 1, pp. 4260, 2020.

[4] L. Li, Y. Chen, X. Yu et al., "Sub-pixel flood inundation mapping from multispectral remotely sensed images based on discrete particle swarm optimization," *ISPRS Journal of Photogrammetry and Remote Sensing,* vol. 101, pp. 10-21, 2015.

[5] A. Lin, H. Wu, G. Liang et al., "A big data-driven dynamic estimation model of relief supplies demand in urban flood disaster," *International Journal of Disaster Risk Reduction,* vol. 49, pp. 101682, 2020.

[6] H. Wu, A. Lin, K. C. Clarke et al., "A comprehensive quality assessment framework for linear features from volunteered geographic information," *International Journal of Geographical Information Science*, pp. 1-22, 2020.

[7] X. Zhu, F. Cai, J. Tian et al., "Spatiotemporal fusion of multisource remote sensing data:Literature survey, taxonomy, principles, applications, and future directions," *Remote Sensing,* vol. 10, no. 4, pp. 527, 2018.

[8] Q. Wang, Y. Tang, X. Tong et al., "Virtual image pair-based spatio-temporal fusion," *Remote Sensing of Environment,* vol. 249, pp. 112009, 2020.

[9] M. H. R. Sales, S. de Bruin, C. Souza *et al.*, "Land use and land cover area estimates from class membership probability of a random forest classification," *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1-11, 2021.

[10] X. Luo, X. Tong, and H. Pan, "Integrating multiresolution and multitemporal sentinel-2 imagery for land-cover mapping in the xiongan new area, china," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 59, no. 2, pp. 1029-1040, 2021.

[11] L. Zhang, M. Lan, J. Zhang *et al.*, "Stagewise unsupervised domain adaptation with adversarial self-training for road segmentation of remote-sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1-13, 2021.

[12] P. M. Atkinson, M. E. J. Cutler, and H. Lewis, "Mapping sub-pixel proportional land cover with avhrr imagery," *International Journal of Remote Sensing,* vol. 18, no. 4, pp. 917-935, 1997.

[13] G. M. Foody, "Status of land cover classification accuracy assessment," *Remote Sensing of Environment,* vol. 80, no. 1, pp. 185-201, 2002.

[14] Q. Wang, W. Shi, and P. M. Atkinson, "Spatiotemporal subpixel mapping of time-series images," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 54, no. 9, pp. 5397-5411, 2016.

[15] Y. F. Su, G. M. Foody, A. M. Muad *et al.*, "Combining pixel swapping and contouring methods to enhance super-resolution mapping," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 5, no. 5, pp. 1428-1437, 2012.

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

[17] A. J. Tatem, H. G. Lewis, P. M. Atkinson *et al.*, "Super-resolution mapping of multiple-scale land cover features using a hopfield neural network." pp. 3200-3202 vol.7.

[18] Y. Zhong, Y. Wu, X. Xu *et al.*, "An adaptive sub-pixel 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.

[19] 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, pp. 1-15, 2014.

[20] K. C. Mertens, B. de Baets, L. P. C. Verbeke *et al.*, "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.

[21] 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.

[22] K. Mertens, "Sub-pixel mapping and sub-pixel sharpening using neural network predicted wavelet coefficients," *Remote Sensing of Environment,* vol. 91, no. 2, pp. 225-236, 2004.

[23] 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.

[24] S. Wu, J. Ren, Z. Chen *et al.*, "Influence of reconstruction scale, spatial resolution and pixel spatial relationships on the sub-pixel mapping accuracy of a double-calculated spatial attraction model," *Remote Sensing of Environment,* vol. 210, pp. 345-361, 2018.

[25] Y. Chen, Y. Ge, Y. Chen *et al.*, "Subpixel land cover mapping using multiscale spatial dependence," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 56, no. 9, pp. 5097-5106, 2018.

[26] P. Wang, L. Wang, H. Leung *et al.*, "Super-resolution mapping based on spatial-spectral correlation for spectral imagery," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 59, no. 3, pp. 2256-2268, 2021.

[27] 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.

[28] H. Wu, Z. Li, K. C. Clarke *et al.*, "Examining the sensitivity of spatial scale in cellular automata markov chain simulation of land use change," *International Journal of Geographical Information Science,* vol. 33, no. 5, pp. 1040-1061, 2019.

[29] Y. Ge, Y. Jin, A. Stein *et al.*, "Principles and methods of scaling geospatial earth science data," *Earth-Science Reviews,* vol. 197, pp. 102897, 2019.

[30] Y. Chen, Y. Ge, and Y. Jia, "Integrating object boundary in super-resolution land-cover mapping," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 10, no. 1, pp. 219-230, 2017.

[31] Z. Mahmood, M. A. Akhter, G. Thoonen *et al.*, "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.

[32] X. Li, F. Ling, Y. Du *et al.*, "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.

[33] Y. Jia, Y. Ge, Y. Chen *et al.*, "Super-resolution land cover mapping based on the convolutional neural network," *Remote Sensing,* vol. 11, no. 15, pp. 1815, 2019.

[34] F. Ling, Y. Du, F. Xiao *et al.*, "Super-resolution land-cover mapping using multiple sub-pixel shifted remotely sensed images," *International Journal of Remote Sensing,* vol. 31, no. 19, pp. 5023-5040, 2010.

[35] M. Song, Y. Zhong, A. Ma *et al.*, "Multiobjective subpixel mapping with multiple shifted hyperspectral images," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 58, no. 11, pp. 8176-8191, 2020.

[36] Q. Wang, C. Zhang, and P. M. Atkinson, "Sub-pixel mapping with point constraints," *Remote Sensing of Environment,* vol. 244, pp. 111817, 2020.

[37] Y. Zhang, Y. Du, F. Ling *et al.*, "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.

[38] F. Ling, and G. M. Foody, "Super-resolution land cover mapping by deep learning," *Remote Sensing Letters,* vol. 10, no. 6, pp. 598-606, 2019.

[39] Y. Zhang, F. Ling, X. Li *et al.*, "Super-resolution land cover mapping using multiscale self-similarity redundancy," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 8, no. 11, pp. 5130-5145, 2015.

[40] D. He, Q. Shi, X. Liu *et al.*, "Deep subpixel mapping based on semantic information modulated network for urban land use mapping," *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1-19, 2021.

[41] F. Ling, W. Li, Y. Du *et al.*, "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.

[42] 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.

[43] X. Li, F. Ling, Y. Du *et al.*, "A spatial-temporal hopfield neural network approach for super-resolution land cover mapping with multi-temporal different resolution remotely sensed images," *ISPRS Journal of Photogrammetry and Remote Sensing,* vol. 93, pp. 76-87, 2014.

[44] D. He, Y. Zhong, and L. Zhang, "Spectral-spatial-temporal map-based sub-pixel mapping for land-cover change detection," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 58, no. 3, pp. 1696-1717, 2020.

[45] X. Li, F. Ling, G. M. Foody *et al.*, "Generating a series of fine spatial and temporal resolution land cover maps by fusing coarse spatial resolution remotely sensed images and fine spatial resolution land cover maps," *Remote Sensing of Environment,* vol. 196, pp. 293-311, 2017.

[46] Y. Zhang, G. M. Foody, F. Ling *et al.*, "Spatial-temporal fraction map fusion with multi-scale remotely sensed images," *Remote Sensing of Environment,* vol. 213, pp. 162-181, 2018.

[47] Y. Shendryk, Y. Rist, C. Ticehurst *et al.*, "Deep learning for multi-modal classification of cloud, shadow and land cover scenes in planetscope and sentinel-2 imagery," *ISPRS Journal of Photogrammetry and Remote Sensing,* vol. 157, pp. 124-136, 2019.

[48] H. Kimm, K. Guan, C. Jiang *et al.*, "Deriving high-spatiotemporal-resolution leaf area index for agroecosystems in the u.S. Corn belt using planet labs cubesat and stair fusion data," *Remote Sensing of Environment,* vol. 239, pp. 111615, 2020.

[49] H. Huang, and D. P. Roy, "Characterization of planetscope-0 planetscope-1 surface reflectance and normalized difference vegetation index continuity," *Science of Remote Sensing*, pp. 100014, 2021.

[50] S. Wu, J. Wang, Z. Yan *et al.*, "Monitoring tree-crown scale autumn leaf phenology in a temperate forest with an integration of planetscope and drone remote sensing observations," *ISPRS Journal of Photogrammetry and Remote Sensing,* vol. 171, pp. 36-48, 2021.

[51] A. J. Tatem, H. G. Lewis, P. M. Atkinson *et al.*, "Super-resolution target identification from remotely sensed images using a hopfield neural network," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 39, no. 4, pp. 781-796, 2001.

[52] F. Ling, X. Li, Y. Du *et al.*, "Super-resolution land cover mapping with spatial-temporal dependence by integrating a former fine resolution map," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 7, no. 5, pp. 1816-1825, 2014.

[53] Q. Wang, X. Ding, X. Tong *et al.*, "Spatio-temporal spectral unmixing of time-series images," *Remote Sensing of Environment,* vol. 259, pp. 112407, 2021.

[54] C. Hu, Y. Lu, S. Sun *et al.*, "Optical remote sensing of oil spills in the ocean: What is really possible?," *Journal of Remote Sensing,* vol. 2021, pp. 1-13, 2021.

[55] S. Rajendran, F. N. Sadooni, H. A. Al-Kuwari *et al.*, "Monitoring oil spill in norilsk, russia using satellite data," *Sci Rep,* vol. 11, no. 1, pp. 3817, 2021.
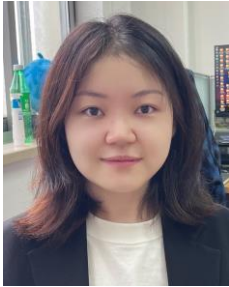
Chengyuan Zhang received the B.S. degree from the Central China Normal University, Wu han, China, in 2019. She is currently working toward the Ph.D. degree at Tongji University, Shanghai, China.

Her research interests include downscaling and mixed pixel analysis (subpixel mapping and spectral unmixing).
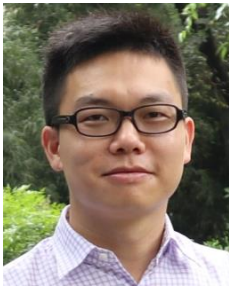
Qunming Wang received the Ph.D. degree from the Hong Kong Polytechnic University, Hong Kong, in 2015.

He is currently a Professor with the College of Surveying and Geo-Informatics, Tongji University, Shanghai, China. He was a Lecturer (Assistant Professor) with Lancaster Environment Centre, Lancaster University, Lancaster, U.K., from 2017 to 2018. His 3-year Ph.D. study was supported by the hypercompetitive Hong Kong Ph.D. Fellowship and his Ph.D. thesis was awarded as the Outstanding Thesis in the Faculty. He has authored or coauthored 70 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 research interests include remote sensing, image processing and geostatistics.

Dr. Wang serves as Associate Editor for *Science of Remote Sensing* (sister journal of *Remote Sensing of Environment*) and *Photogrammetric Engineering & Remote Sensing*, and was Associate Editor for *Computers and Geosciences* (2017−2020).

Ping Lu received the M.S. degree in Earth Observations from Katholieke Universiteit Leuven, Belgium, in 2007, and the PhD degree in Earth Sciences from University of Firenze, Italy, in 2011.

He is currently an associate professor with College of Surveying and Geo-Informatics, Tongji University, China. His research interests include remote sensing techniques applied to natural hazards.

Yong Ge obtained the Bachelor and Master degree of Engineering in surveying and mapping from Wuhan University in 1995 and 1998, respectively, and the Ph.D. degree in cartography and GIS from Chinese Academy of Sciences in 2001.

Since 2004, she has been working at the State Key Laboratory of Resources and Environmental Information System (LREIS). Now she is a full professor and her research field is the information extraction from spatial data. She has authored more than 200 papers and is the associate editor of the *Computers and Geosciences* and the editor board member of *Spatial Statistics*, which places it at the Q1 of statistical journals. She was also the guest editor of various special issues of journals.

Peter M. Atkinson received the Ph.D. 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 Distinguished Professor of Spatial Data Science and Dean of the Faculty of Science and Technology at Lancaster University, UK. He was previously Professor of Geography at the University Southampton, where he is currently Visiting Professor. He is also Visiting Professor at the Chinese Academy of Sciences, Beijing. He previously held the Belle van Zuylen Chair at Utrecht University, the Netherlands, is a recipient of the Peter Burrough Award of the International Spatial Accuracy Research Association and is a Fellow of the Learned Society of Wales. The main focus of his research is in remote sensing, geographical information science and spatial (and space-time) statistics applied to a range of environmental science and socio-economic problems. He has published over 300 peer-reviewed articles in international scientific journals and around 50 refereed book chapters. He has also edited nine journal special issues and eight books.

Professor Atkinson is Editor-in-Chief of *Science of Remote Sensing*, a sister journal of *Remote Sensing of Environment*. He also sits on the editorial boards of several further journals including *Geographical Analysis*, *Spatial Statistics*, *International Journal of Applied Earth Observation and Geoinformation*, and *Environmental Informatics*.