Sub-pixel Mapping with Point Constraints

Qunming Wang [a, *], Chengyuan Zhang [a], Peter M. Atkinson [b, c, d, e]

[a] College of Surveying and Geo-Informatics, Tongji University, 1239 Siping Road, Shanghai 200092, China

[b] Faculty of Science and Technology, Lancaster University, Lancaster LA1 4YR, UK

[c] School of Geography, Archaeology and Palaeoecology, Queen's University Belfast, BT7 1NN, Northern Ireland, UK

[d] Geography and Environment, University of Southampton, Highfield, Southampton SO17 1BJ, UK

[e] Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, Datun Road, Beijing 100101, China

*Corresponding author. E-mail: wqm11111@126.com

Abstract: Remote sensing images contain abundant land cover information. Due to the complex nature of land cover, however, mixed pixels exist widely in remote sensing images. Sub-pixel mapping (SPM) is a technique for predicting the spatial distribution of land cover classes within mixed pixels. As an ill-posed inverse problem, the uncertainty of prediction cannot be eliminated and hinders the production of accurate sub-pixel maps. In contrast to conventional methods that use continuous geospatial information (e.g., images) to enhance SPM, in this paper, a SPM method with point constraints into SPM is proposed. The method of fusing point constraints is implemented based on the pixel swapping algorithm (PSA) and utilizes the auxiliary point information to reduce the uncertainty in the SPM process and increase map accuracy. The point data are incorporated into both the initialization and optimization processes of PSA. Experiments were performed on three images to validate the proposed method. The influences of the performances were also investigated under different numbers of point data, different spatial characters of land cover and different zoom factors. The results show that by using the point data, the proposed SPM method can separate more small-sized targets from aggregated

artifacts and the accuracies are increased obviously. The proposed method is also more accurate than the advanced radial basis function interpolation-based method. The advantage of using point data is more evident when the point data size and scale factor are large and the spatial autocorrelation of the land cover is small. As the amount of point data increases, however, the increase in accuracy becomes less noticeable. Furthermore, the SPM accuracy can still be increased even if the point data and coarse proportions contain errors.

Keywords: Remote sensing images, sub-pixel mapping (SPM), super-resolution mapping, downscaling, pixel swapping algorithm (PSA), point constraints.

1. Introduction

Many significant applications require reliable land cover information extracted from remote sensing images and land cover mapping is an important topic in remote sensing. Conventional hard classification techniques allocate each pixel to a single land cover class (Atkinson et al., 1997). However, mixed pixels that contain more than one land cover class exist invariably in remote sensing images (Fisher, 1997). Mixed pixels are classified into two main types, namely: Low-resolution or L-resolution (where pixels are larger than the objects of interest) and High-resolution or H-resolution (where pixels are smaller than the objects of interest) (Atkinson, 2009). Irrespective of the kind of mixed pixels, one-class-per-pixel classification unavoidably results in a loss of land cover information. To deal with mixed pixels and obtain more reliable land cover information, soft classification, also termed spectral unmixing, has been developed to predict the proportion of each land cover class within pixels resulting in a set of land cover proportion images (Eastman and Laney, 2002). Sub-pixel class composition can be estimated by soft classification (Bioucas-Dias et al., 2012; Keshava and Mustard, 2002), but the spatial distribution of land cover classes within pixels is still unknown. To deal with this issue, sub-pixel mapping (SPM) has been developed, also termed super-resolution mapping in remote sensing. SPM

divides each pixel into sub-pixels and predicts their class labels, that is, the land cover classification is performed at a finer spatial resolution than that of the input images.

Generally, SPM is undertaken by utilizing the expectation that remotely sensed images are spatially dependent, that is, nearby observations (pixels or sub-pixels) are considered more likely to be of the same class than more distant ones. Various methods have been developed for SPM. According to the scheme used for characterizing spatial dependence, SPM can be divided into three main categories. The first concerns the spatial relation between each sub-pixel and its neighboring sub-pixels. In this set, SPM is an optimization problem, and the spatial distribution of sub-pixels is changed iteratively to maximize the quantified spatial dependence. Solutions commonly used in this group include pixel-swapping algorithm (PSA) (Atkinson, 2005), Hopfield neural network (HNN) (Foody and Doan, 2007; Tatem et al., 2001), genetic algorithms (Mertens et al., 2003), particle swarm optimization (Wang et al., 2012) and maximum *a posteriori* (MAP) (Zhong et al., 2015). The second set concerns the spatial relation between each sub-pixel and neighboring coarse pixels. For this group, it is critical to describe accurately the scale difference between the coarse and fine spatial resolution pixels. Methods belonging to this group include the sub-pixel/pixel spatial attraction model (Mertens et al., 2006), kriging (Jin et al., 2012; Verhoeye and De Wulf, 2002), radial basis function interpolation (RBF) (Wang et al., 2014a) and double-calculated spatial attraction model (DSAM) (Wu et al., 2018). The third set considers both of the spatial dependencies mentioned above. Examples include hybrid intra- and inter-pixel dependence (Ling et al., 2013) and class membership probabilities calculated by fusing a smoothness prior and the downscaled proportions (Ling et al., 2014). Chen et al. (2018a) characterized the spatial dependence at the object level, which was combined with the abovementioned two types of dependences (i.e., multiscale spatial dependence) in their later work (Chen et al., 2018b).

To circumvent the strong reliance on spectral unmixing of the above methods, spatial-spectral models have been developed that simultaneously consider the spatial goal of maximizing spatial dependence and the spectral constraints from the observed coarse data. Kasetkasem et al. (2005) and Tolpekin and Stein (2009) proposed Markov random field (MRF)-based SPM. Xu et al. (2018) introduced a spatial-spectral SPM model

that is performed directly for hyperspectral images. In addition, SPM can be realized through geostatistical solutions. As presented in Boucher et al. (2008), indicator kriging was used to predict the probabilities of land cover class occurrence at the sub-pixel level. Moreover, for the L-resolution case, the two-point histogram was applied to match the spatial structure of SPM prediction with the training image (Atkinson, 2004). Recently, Song et al. (2019) reformulated the SPM problem into a multi-objective optimization problem using a sparse representation method.

It has been recognized widely that SPM is an ill-posed problem (Ling et al., 2010). Specifically, for a given coarse spatial resolution image with a spatial size of $X \times Y$, a finer spatial resolution land cover map with a spatial size of $(X \times s) \times (Y \times s)$ needs to be predicted, where s is the zoom factor. There exists unavoidable uncertainty in SPM predictions where multiple solutions can lead equally to the same coherence constraint imposed by the coarse proportion or original coarse image. The uncertainty in predictions becomes greater when s is larger (Wu et al., 2018). As a result, aggregated artifacts exist widely in the predicted land cover maps and some features such as linear or elongated features cannot be satisfactorily restored. Such a problem can be alleviated by utilizing supplementary information.

Over the past decades, various sources of supplementary information, such as additional data, training images or *a priori* information, have been used to reduce prediction uncertainty and enhance SPM. Specifically, prior shape information was used to extract fine spatial resolution building objects from observed coarse images (Ling et al., 2012). The vectorial boundary information extracted from land-line digital vector data was used to refine the boundaries of objects after conventional SPM (Aplin and Atkinson, 2001). Ling et al. (2008) applied digital elevation model (DEM) to sub-pixel waterline mapping according to the physical features of the water flow. For global scale forest mapping based on MODIS data, the 250 m MODIS bands which have a finer spatial resolution than the input 500 m bands were incorporated into the SPM process to increase the accuracy of predicted 30 m forest maps (Zhang et al., 2017). Some geostatistical SPM methods extract spatial patterns of a prior structural model from training images (Boucher et al., 2008). Multiple shifted images were utilized by the HNN and indicator cokriging models respectively, to produce more reliable soft class attribute

of sub-pixels (Ling et al., 2010; Wang et al., 2017; Wang et al., 2014b). The panchromatic (PAN) image can also be applied to enhance SPM. Specifically, the PAN reflectance constraint was added into the HNN energy function (Nguyen et al., 2011) and the use of PAN image can increase the accuracy of tree crown detection (Ardila et al., 2011). Moreover, Light Detection And Ranging (LiDAR) elevation data were used to construct an additional height energy function in the HNN for more reliable building mapping (Nguyen et al., 2005). In the latest deep learning-based methods, the fine spatial resolution training images were also used to learn the relation between the coarse proportion images and fine land cover maps using a convolutional neural network, which can reproduce complex spatial structure in the SPM predictions (Jia et al., 2019; Ling and Foody, 2019).

In recent years, several SPM methods have been developed to borrow temporal information (i.e., temporally neighboring images covering the same area) and some spatio-temporal SPM models have been proposed. More precisely, for global scale monitoring at both fine spatial and temporal resolutions, SPM was performed on temporally dense, 500 m MODIS time-series data to predict land cover maps at 30 m (i.e., Landsat) spatial resolution, where several temporally sparse, 30 m Landsat images were available and the 30 m information was incorporated in the SPM process (Li et al., 2017; Wang et al., 2016; Zhang et al., 2017). In detail, Zhang et al. (2017) proposed a temporal-example learning-based method to predict class possibilities at the sub-pixel-scale. Li et al. (2017) utilized a series of coarse MODIS images for SPM and in the pre-spectral unmixing step. The changes in proportions in each coarse pixel across time were analyzed to extract more reliable endmembers for the post-unmixing and SPM steps. Wang et al. (2016) proposed a mathematical model to quantify spatio-temporal dependence by combining the spatial and temporal dependences linearly.

In some cases, for SPM of a specific region, there may be very few data available or they may be insufficiently reliable for enhancing SPM, such that it is necessary to seek other additional data. In practice, there exist various sources of data for obtaining the land cover class labels at the target fine spatial resolution, especially in the current era of "big data". In this paper, we refer to this type of available data as 'point data' and the corresponding sub-pixel as the 'informed sub-pixel'. Different from the abovementioned existing methods that use spatially continuous data for enhancing SPM, this paper uses spatially sparse (i.e., partly known) point

data as a new constraints in SPM to reduce the uncertainty and increase the prediction accuracy. The potential sources for such information can be summarized as follows:

1) Field surveys or manual interpretations. In conventional supervised pixel-level remote sensing image classification, one of the most common used ways to obtain training samples (i.e., informed pixels) is manual interpretations associated with field surveys (Bizzi et al., 2016). Similarly, informed pixels can be collected by field surveys or manual interpretation for SPM.
2) Social media data. Social media data have been used for pixel-level classification and have received increasing attention in recent years. Social media data can provide valuable information for land cover and land use mapping. For example, locations with more frequent check-in for social software such as Twitter, Facebook and Weibo, are more likely to be impervious surface (Frias-Martinez and Frias-Martinez, 2014; Fritz et al., 2012). Such data have been fused with fine spatial resolution remote sensing images for urban land use classification and the scheme was demonstrated to be promising (Liu et al., 2017). Compared to manual interpretation, social media data may be labor-saving but provide a broader source for land cover mapping.
3) Prior knowledge. The land cover information of some classes can be acquired empirically, such as by using temporal information and widely acknowledged change rules. For instance, if a pixel belongs to the urban class in the past few years, it will remain the same afterwards, as the change of most of the cities are generally assumed to be unidirectional, that is, expanding rather than degenerating into other classes such as vegetation (Jokar Arsanjani et al., 2013).

Such rich supplementary point information, although sparse in spatial distribution, is significant for classification of large regions. To the best of our knowledge, such point data have not been utilized in the existing SPM literature. To explore the effect of such widely available point data on enhancing SPM, this paper develops a new method and extends an existing SPM framework to incorporate the point data as SPM constraints. The main contributions of this paper are, thus, twofold.

1) Point data are proposed for enhancing SPM. Different from existing methods using supplementary

spatially continuous information, point data at the sub-pixel scale are spatially sparse.

2) A new SPM method is developed to incorporate the point constraints. In addition to conventional SPM methods performed based on a spatial goal and a coarse data coherence constraint, the proposed method accounts for the point constraints directly and simultaneously as part of the solution.

The rest of this paper is organized as follows. Section II first describes the benefit of using point data in SPM in Section II-A, followed by the details of the proposed SPM method incorporating point constraints in Section II-B. In Section III, experimental results for both the H- and L-resolution cases are provided. A discussion of the results is provided in section IV and section V concludes this paper.

2. Methods

2.1. The Benefit of Using Point Constraints

SPM is essentially an ill-posed problem. Finding available auxiliary information has become increasingly popular as a means of reducing the uncertainty in the SPM process. In this paper, point data, which refers to class-informed fine spatial resolution pixels (i.e., class-informed sub-pixels), are proposed for enhancing SPM. The method of incorporating point constraints is different to existing methods that use spatially continuous information (i.e., data covering the entire study area), such as learning -based methods (Ling et al., 2016). The benefit of using point information is analyzed by an explicit example below.

With auxiliary point data, additional point constraints aside from spatial correlation are provided to reduce the uncertainty in the allocation of sub-pixel classes. Fig. 1 shows an example to describe the practical significance of point data (i.e., class-informed sub-pixels) in the process of SPM. Two classes (white background and gray target) were considered in a region covering 3×3 coarse pixels. Suppose the proportion of the target in the central mixed pixel is 1/3 and the zoom factor s=3. Amongst the nine sub-pixels, three sub-pixels should be allocated to the gray class.

Based on the assumption of maximizing spatial correlation, multiple distributions of the gray sub-pixels in the central coarse pixel are listed in Fig. 1(a). There are no other constraints to determine which solution is optimal.

If there is some supplemental information at the sub-pixel level, such as the red-marked sub-pixel at (1, 1) that is assigned to the gray class in Fig. 1(b), all the sub-pixels in the red shaded part (the neighboring window of this point) are attracted by the informed sub-pixel. The closer the to-be-determined sub-pixel is to it, the larger the probability that it belongs to be the same gray class. Therefore, solutions that do not satisfy the constraint imposed by this point are excluded (i.e., the second and the fourth allocations in which a red fork is drawn are excluded, as shown in Fig. 1(b)).

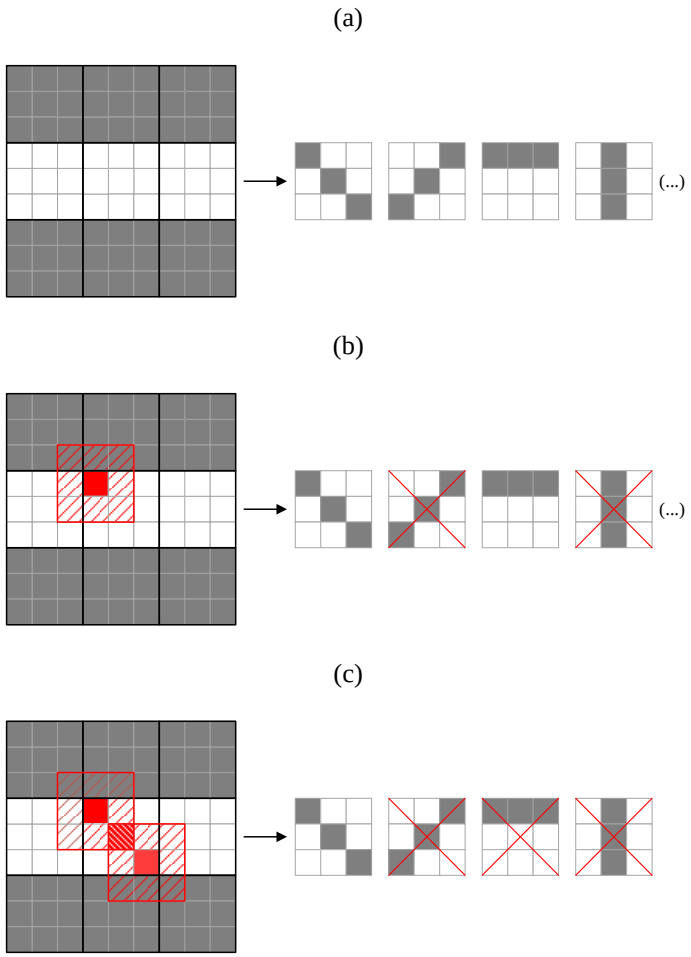


Fig. 1. An example of incorporating point constraints in SPM and possible sub-pixel class distributions in the central mixed pixel. (a) No point constraint. (b) Possible distributions with a single informed sub-pixel. (c) Possible distributions with two informed sub-pixels.

When another red-marked sub-pixel at (3, 3) is also assigned to the gray class, the common neighboring sub-pixel of the two constraints at (2, 2) can be determined with little uncertainty, as the spatial attraction exerted on it strong. Thus, the most likely distribution, that is the first allocation in Fig. 1(c), can be selected.

2.2. SPM with Point Constraints

The known point data exist at the sub-pixel resolution. Thus, it is a natural choice to consider SPM methods that characterize the spatial dependence at the sub-pixel level (i.e., spatial relation between sub-pixels). This type of sub-pixel to sub-pixel methods depends on optimization solutions based on a general model composed of a goal and constraint as:

$$E = k_1 G + k_2 C_{\text{pixel}} \tag{1}$$

where k_1 and k_2 are weights, the term G is the spatial goal of maximizing spatial dependence and C_{pixel} is the pixel proportion constraint at the coarse spatial resolution.

In existing optimization-based SPM methods, only the upper bound (pixel proportion at coarse spatial resolution) is considered as a constraint. That is, the number of sub-pixels for each class should be consistent with the class proportion at the coarse pixel resolution.

In the proposed SPM with point constraints method, however, the constraint term is extended by considering the sub-pixel resolution hard data, and the abovementioned model is extended to:

$$E = k_1 G + k_2 C_{\text{pixel}} + k_3 C_{\text{sub-pixel}} \tag{2}$$

where $C_{\text{sub-pixel}}$ is the point constraint term at the sub-pixel level.

In this way, we have constraints at both the lower bound (sub-pixel) and upper bound (coarse proportion), rather than only the upper bound in existing optimization-based SPM methods. The point data provide more detailed information at the target fine spatial resolution, which can impose more reliable constraints for sub-pixel class prediction, especially for the neighbors of the point data. Thus, some inappropriate SPM solutions can be excluded (as illustrated in Fig. 1) and the solution space can be reduced.

2.3. PSA-based SPM

Common choices for sub-pixel to sub-pixel methods are PSA and HNN. The HNN method involves a large number of iterations and thus, is computationally expensive. Moreover, the HNN cannot preserve linear and elongated features as well as small objects satisfactorily and the predictions tend to be overly smooth. Alternatively, the PSA method proposed in Atkinson (2005) is employed to account for the proposed point constraints in this paper.

In this section, the principle of PSA is briefly introduced. The implementation of PSA consists of two main steps: initialization of sub-pixel classes according to the coarse proportion constraint and optimization based on the goal of maximizing spatial correlation through pixel swapping.

Suppose p_i is a sub-pixel in coarse pixel P and sub-pixels $p_j (j=1,2,...,N)$ are its neighbors, s is the zoom factor (i.e., each coarse pixel is divided into $s \times s$ sub-pixels), and F_k is the coarse proportion of the kth ($k=1,2,\ldots,K$, K is the number of land cover classes) class in pixel P. The implementation of PSA is described briefly as follows:

Step 1) Initialization. For each coarse pixel, according to the scale factor s and its coarse proportions F_k, a fixed number of sub-pixels are allocated to class k. An initial sub-pixel map is generated after pixels for each class are exhausted. Since all sub-pixels are allocated randomly in this step, the uncertainty is obvious.

Step 2) Optimization. Two sub-pixels within pixel P are selected randomly and sub-pixel swapping is allowed if it increases the sum of attractiveness of all sub-pixels in pixel P. Otherwise, the original sub-pixel classes are kept. The procedure is repeated until convergence. The attractiveness $A(p_i)$ of sub-pixel p_i is predicted as a distance-weighted function of its N neighbors:

$$A(p_i)=\sum_{j=1}^{N} \lambda_{i,j} \delta(p_i, p_j) \tag{3}$$

where $\lambda_{i,j}$ is a weighting coefficient which is depicted as:

$$\lambda_{i,j} = \exp\left(\frac{-d(p_i, p_j)}{a}\right) \tag{4}$$

in which $d(p_i, p_j)$ is the Euclidean distance between p_i and p_j and a is a nonlinear parameter in the exponential model. In Eq. (3), $\delta(p_i, p_j)$ is a binary indicator:

$$\delta(p_i, p_j) = \begin{cases} 1, \text{ if } p_i \text{ and } p_j \text{ belongs to the same class} \\ 0, \text{ otherwise.} \end{cases} \quad (5)$$

The PSA model considers only how to maximize the spatial correlation when optimizing the spatial distribution of sub-pixels. In reality, however, the distribution of land cover may not fully accord with this maximum spatial correlation hypothesis. Auxiliary information is necessary for increasing the accuracy of SPM in this situation.

2.4. The Proposed Method for Incorporating Point Constraints

Based on the above analysis of using point data and the uncertainty in PSA, the SPM method with point constraints is proposed. It makes the following extensions to the original PSA method by incorporating sub-pixel level supplemental information.

1) Initialization Stage: In the original PSA method, for a mixed pixel, the number of sub-pixels belonging to class k (k=1,2,…,K) is determined as:

$$N_k = \text{round}(F_k \bullet s^2) \quad (6)$$

where round($\bullet$) function takes the integer nearest to $\bullet$.

Under the coarse proportion constraints in Eq. (6), the number of possible initialization results for any coarse pixel is quantified as an arrangement:

$$I_1 = \prod_{k=1}^{K} C_{s^2 - \sum_{k=1}^{K} N_{k-1}}^{N_k} \quad (7)$$

where $C_m^{N_k}$ is a combination that calculates the total possible distributions of N_k sub-pixels assigned to class k and m represents the number of remaining unassigned sub-pixels. Particularly, N_k=0 if k=0.

In the proposed SPM with point constraints method, some of the sub-pixels are informed and the class labels are free of uncertainty. Let $N_{point(k)}$ be the number of informed sub-pixels belonging to class k in mixed pixel P,

and the sum of $N_{point(k)}$ for all classes is N_{point}. The number of remaining sub-pixels with uncertainty about the assignment to class k can be computed as:

$$L_k = N_k - N_{point(k)} \tag{8}$$

Then the number of possible initialization results for the remaining sub-pixels is quantified as:

$$I_2 = \prod_{k=1}^{K} C_{s^2 - \sum_{k=1}^{K} N_{k-1} + N_{point(k)}}^{L_k} \tag{9}$$

Therefore, by accounting for the point constraints, the decrease in the solution space is calculated by subtracting Eq. (9) from Eq. (7). A simple example is given in Fig. 2 to illustrate the decrease in uncertainty. Suppose the scale factor s=3 and the proportions of the two coarse pixels for classes A, B and C are all 1/3. The red sub-pixel represents the informed sub-pixel (i.e., known point data). The gray sub-pixel represents the to-be-allocated sub-pixel.

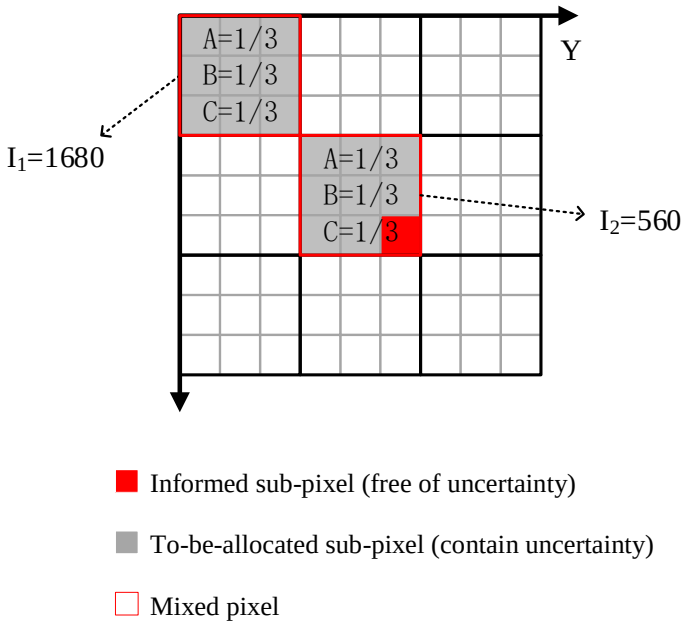


Fig. 2. Two mixed pixel under initialization in PSA. One has no informed sub-pixel and the other has one informed sub-pixel. The numbers of possible distributions are calculated.

As shown in the upper-left corner of Fig. 2, at the initialization stage in the original PSA model, all nine sub-pixels contain uncertainty and there are $I_1=C_9^3 \cdot C_6^3 \cdot C_3^3=1680$ possible initialization results. When one sub-pixel labeled in red in the central pixel is determined through point data, the remaining 8 sub-pixels need to be initialized. In this case, the number of possible initializations is $I_2=C_8^2 \cdot C_6^3 \cdot C_3^3=560$ and the uncertainty is reduced by 67% compared to PSA without supplemental point data.

2) Calculation of Attractiveness: Based on the initialized sub-pixel map, the final SPM result is obtained through an iterative pixel swapping process so that the spatial correlation approaches a maximum gradually. During each iteration, each swap is directly influenced by the calculation of spatial correlation in terms of spatial attractiveness.

As depicted in Fig. 3, after integrating the point constraint, the new attractiveness of sub-pixel p_i consists of two parts, that is, the attractiveness from the informed sub-pixels (i.e., colored sub-pixels in Fig. 3) without uncertainty and the attractiveness from the to-be-determined sub-pixels (i.e., white sub- pixels in Fig. 3).

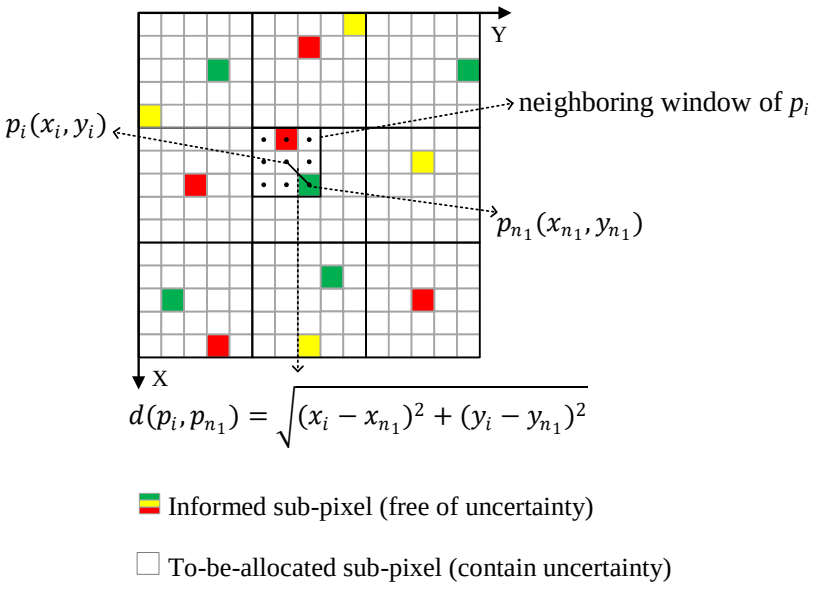


Fig. 3. Calculation of attractiveness with point constraints.

To alleviate the uncertainty in the optimization process, this paper uses informed sub-pixels in the computation of attractiveness. Suppose $N_{point}^{(i)}$ is the number of informed sub-pixels in the local window centered at sub-pixel p_i. The new attractiveness $A'(p_i)$ of sub-pixel p_i is predicted as a distance-weighted function of its $N_{point}^{(i)}$ informed neighbors and $N - N_{point}^{(i)}$ remaining neighbors:

$$A'(p_i) = \sum_{n_1=1}^{N_{point}^{(i)}} \lambda_{i,n_1} \delta(p_i, p_{n_1}) + \sum_{n_2=1}^{N-N_{point}^{(i)}} \lambda_{i,n_2} \delta(p_i, p_{n_2}) \tag{10}$$

where $\delta(p_i, p_{n_1})$ and $\delta(p_i, p_{n_2})$ are binary indicators as defined in Eq. (5), indicating whether the two sub-pixels belong to the same class. λ_{i,n_1} and λ_{i,n_2} are weight coefficients as defined in Eq. (4), and the calculation of Euclidean distance between p_i and p_{n_1} is provided in Fig. 3.

By comparing Eqs. (3) and (10), it can be found that the first item for attractiveness in Eq. (10) does not contain uncertainty due to the informed sub-pixels (point data). Undoubtedly, by considering the point data, the reliability of sub-pixel class prediction increases. For a sub-pixel p_i, the increased reliability for class k can be quantified as:

$$R_k = \frac{1}{Z_k} \sum_{n=1}^{N_{point(k)}^{(i)}} \lambda_{i,n} \delta(p_i, p_n) \tag{11}$$

where $Z_k = \sum_{n=1}^{N_{point}^{(i)}} \lambda_{i,n} \delta(p_i, p_n)$ is a normalization constant, and $N_{point(k)}^{(i)}$ is the number of informed sub-pixels belonging to class k in the local window centered at sub-pixel p_i. That is, the closer the target sub-pixel p_i to p_n that is informed as class k, the larger the likelihood that p_i belongs to the same class as p_n.

The implementation of the proposed SPM with point constraints method can be summarized briefly in three steps:

1) Initialization. Different to original PSA, this step is performed not only under the proportion constraint, but also under the point constraint (i.e., the informed sub-pixels are allocated first).
2) Sub-pixel swapping. The key is the calculation of the total attractiveness of all sub-pixels before and after

swapping. Informed sub-pixels are used to increase the reliability of the attractiveness calculation.

3) Iteration. Step 2 is repeated until convergence or the pre-determined maximum number of iterations is reached.

3. Experiments

Three datasets were used to demonstrate the effectiveness of the proposed SPM with point constraints method, including two for the H-resolution case and one for the L-resolution case. The SPM results of the proposed point-constraints-based method were compared to the original PSA method without point constraints, under the conditions of different volume of points, different scale factors and different spatial structure. To explore the feasibility of the proposed method with uncertainties in the point data, the influence of the attribute error (i.e., the error in the class labels of points) and geo-locational error (i.e. the mismatch between points and the proportion images) were fully considered. Furthermore, the radial basis function (RBF) interpolation-based method (Wang et al., 2014a), which has a strong nonlinear fitting ability and which has been shown to be a competitive SPM method, was used as a benchmark method in this paper. Considering the heavy reliance of the PSA method on the accuracy of spectral unmixing results, the impact of proportion errors was also investigated.

3.1. Experiment on the synthetic dataset

In the first experiment (H-resolution case), two images were used to examine the performance of the proposed SPM method accounting for point constraints. The first is a QuickBird image with a spatial resolution of 0.61 m acquired in August 2005. It covers an area in Xuzhou City, China. The multispectral image was classified into a land cover map using a support vector machine. As shown in Fig. 4, the land cover map contains seven classes: shadow, water, road, tree, grass, roof and bare soil. The second dataset is a 30 m land cover map from the National Land Cover Database 2001 (NCLD 2001, a raster-based classification dataset

with over 50 U.S. states and Puerto Rico). It covers an area in South Carolina and contains four classes, namely, water, urban, agriculture and forest (Fig. 5). Both land cover maps have a size of 512×512 pixels.

(a) (b)

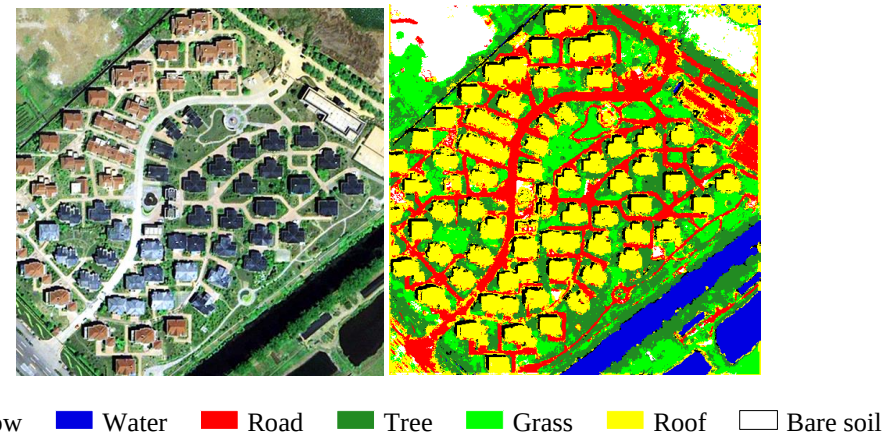


Fig. 4. The 0.61 m Quick Bird images of Xuzhou (512×512 pixels). (a) Original image. (b) Classified land cover map.

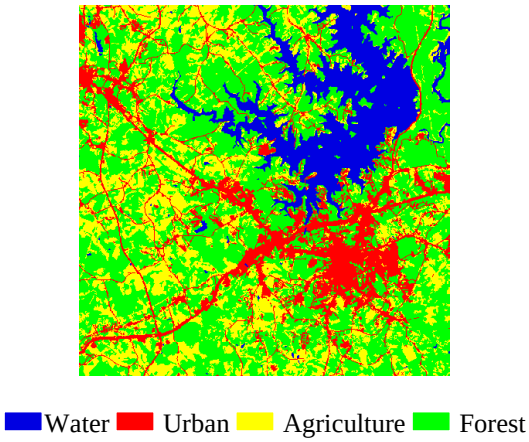


Fig. 5. The 30 m NCLD land cover map of South Carolina (512×512 pixels).

(a) (b)

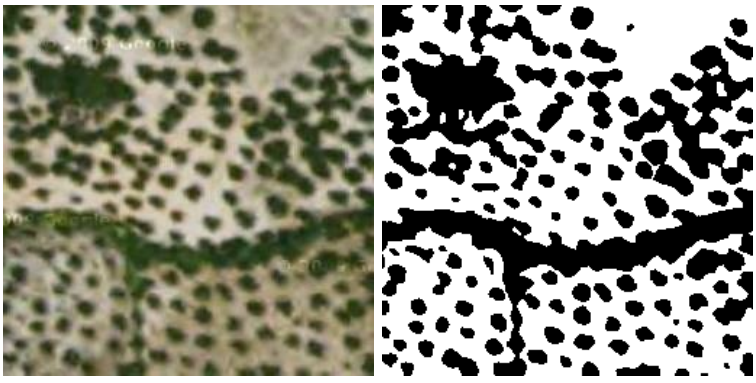


Fig. 6. The SPOT images of Granada, Spain (256×256 pixels). (a) Original image. (b) Classified land cover map.

The experiment for the L-resolution case was implemented using a SPOT image contained in Google Earth covering an area in Granada, Spain (Muad and Foody, 2012). The image was acquired on 1 October 2004. The land cover map containing vegetation and background was derived using a support vector machine. Fig. 6 shows the original map and the classified land cover map.

The workflow of the experiments is shown in Fig. 7. To avoid the uncertainty in spectral unmixing and concentrate solely on the performance of the new SPM method itself, coarse proportions synthesized from fine spatial resolution land cover maps were used. For a map with K land cover classes, the K proportion images were simulated by degrading the corresponding fine spatial resolution binary land cover map with a factor of s (i.e., each $s \times s$ pixels were aggregated to a coarse pixel). α% sub-pixels were then extracted randomly from the fine spatial resolution land cover map. By this strategy of synthesizing data, the reference map is known perfectly for objective evaluation. Moreover, the uncertainty in acquiring point data can also be avoided and, thus, we can focus solely on the performance of the proposed method. The accuracy of SPM was evaluated quantitatively in the light of the percentage of correctly classified sub-pixels (PCC). Notably, the informed sub-pixels were not included in the accuracy statistics since they will increase the accuracy without providing any useful information on the performance of the proposed method accounting for point constraints.

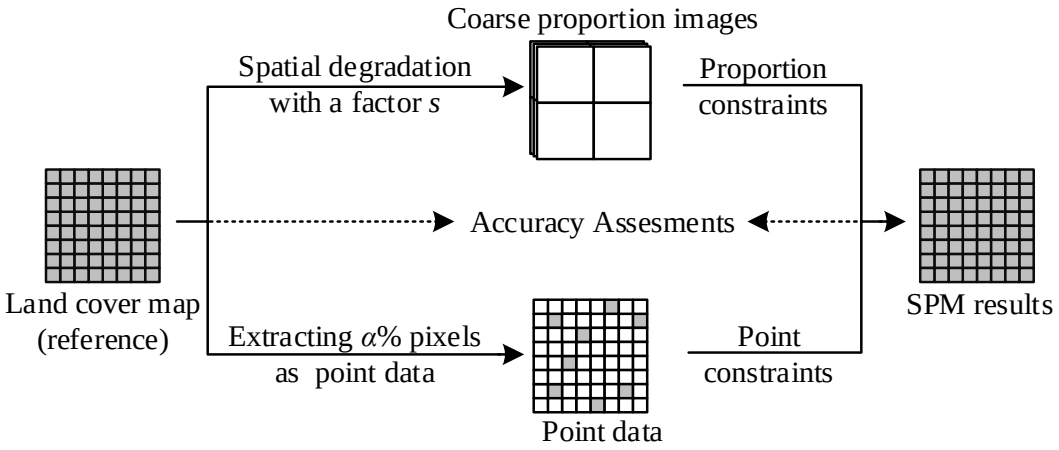


Fig. 7. Experimental approach for validation of the SPM with point constraints method.

3.2. Experiments for H-resolution Case

1) Impact of the number of point data: For the Xuzhou and South Carolina maps, four scale factors (i.e., s=4, 6, 8, and 10) were considered and four groups of coarse proportion images were produced as the inputs of SPM. Fig. 8 shows the proportion images for the seven classes in the Xuzhou map which were produced using a scale factor of s=10. Point data were sampled by ratios α= 5%, 10%, 15%, 20%, 25%, and 30%, respectively.

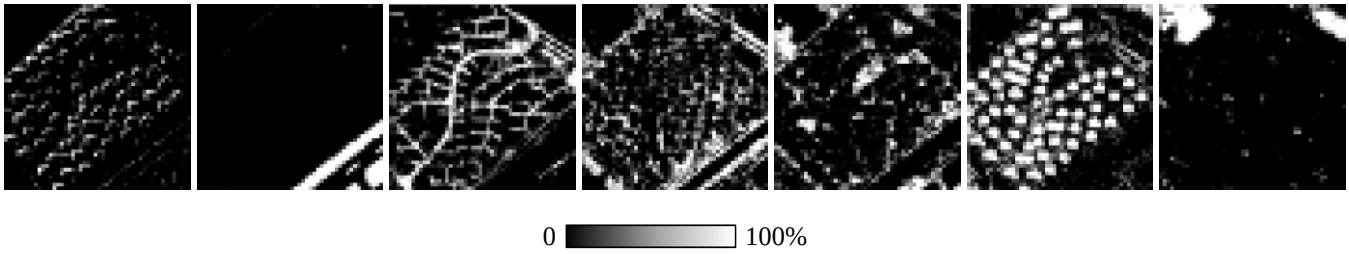


Fig. 8. Coarse proportion images of the Xuzhou area (50×50 pixels, created by degrading the land cover map with a zoom factor of 10). From left to right: shadow, water, road, tree, grass, roof, and bare soil.

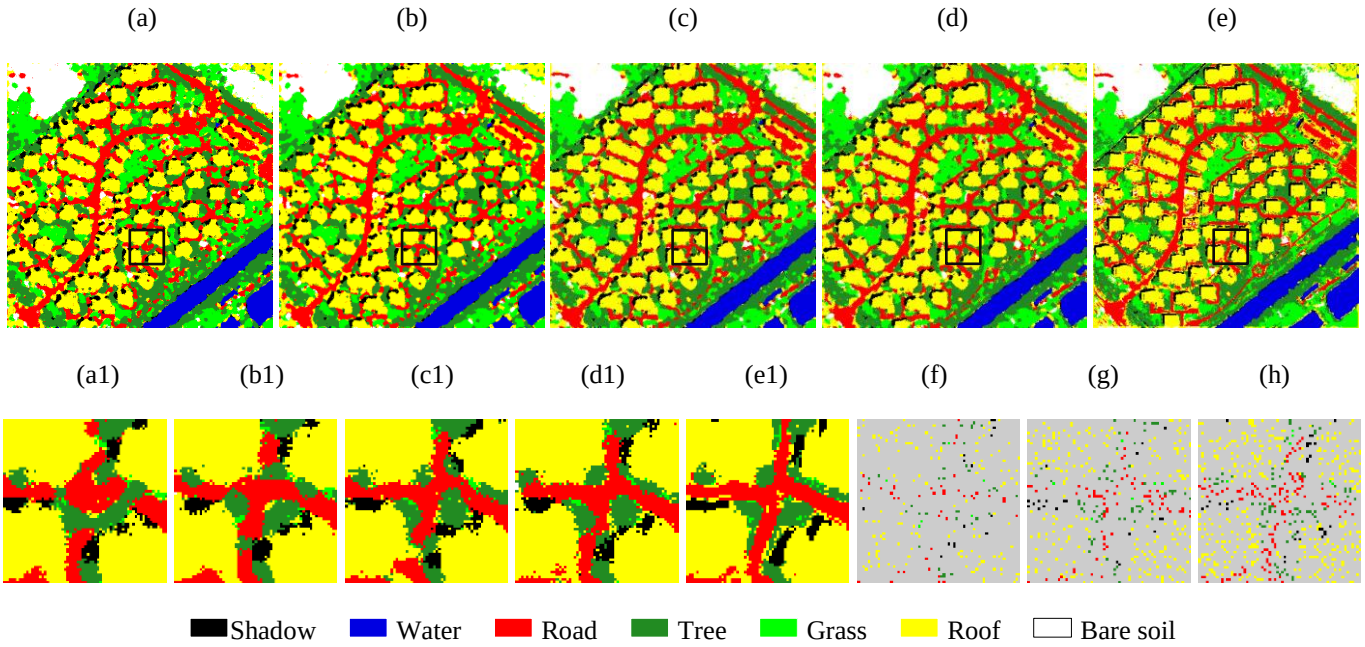


Fig. 9. SPM results of the Xuzhou map (s=10) and corresponding point data maps. (a) Original PSA. (b) PSA with 5% point data. (c) PSA with 10% point data. (d) PSA with 15% point data. (e) Reference map. (a1)-(e1) Zoomed subareas of (a)-(e). (f) 5% point data map. (g) 10% point data map. (h) 15% point data map.

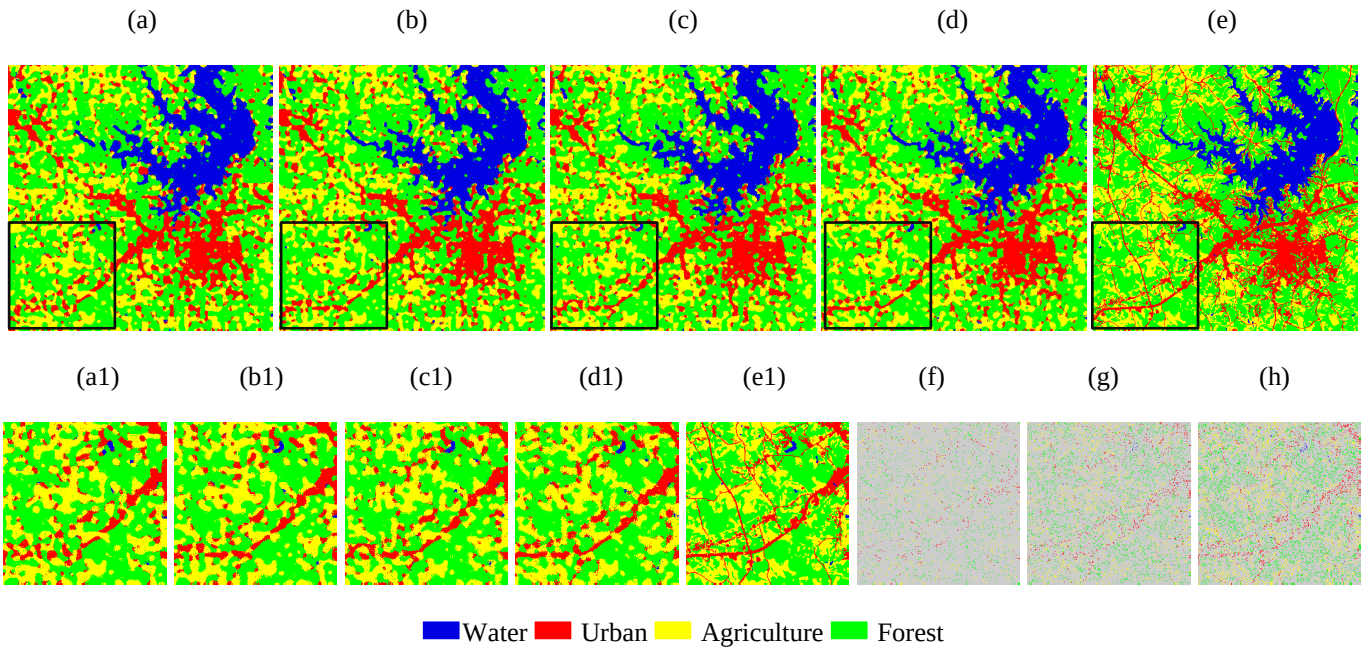


Fig. 10. SPM results of the South Carolina map (s=10) and corresponding point data maps. (a) Original PSA. (b) PSA with 5% point data. (c) PSA with 10% point data. (d) PSA with 15% point data. (e) Reference map. (a1)-(e1) Zoomed subareas of (a)-(e). (f) 5% point data map. (g) 10% point data map. (h) 15% point data map.

Figs. 9 and 10 shows the SPM results of the original PSA and PSA with 5%, 10%, and 15% point data for the case of s=10 for the Xuzhou and South Carolina maps, respectively. It is seen that no matter whether point data are used, the SPM results contain more spatial details than the coarse proportion images where the image is pixelated and the boundaries of classes are rough. For each resulting map in both Figs. 9 and 10, the water class is satisfactorily restored because the distribution of the water is spatially aggregated in nature, which is consistent with the spatial correlation hypothesis in SPM. For the road and shadow classes in Xuzhou and the urban class in South Carolina, which mainly have a linear feature in the reference map, we can find that they are incorrectly divided into several scattered patches in the SPM result of original PSA. As shown in Fig. 9(c), when incorporating 5% point data, the contour of road and shadow looks smoother and continuous, which is closer to the reference map. When the number of point data increases to 15%, the restoration of the road has been obviously improved. This is because the original PSA method tends to produce aggregated artifacts to increase the overall spatial correlation. The same phenomenon can be observed when checking the spatial

structure of the urban class in South Carolina, which becomes more continuous with the increase in the number of point data, see Fig.10(c)-(e). By using point data, however, some prior structure information is provided, especially for elongated features. Thus, the continuity of the elongated objects can be reproduced to some extent by the proposed SPM method with point constraints.

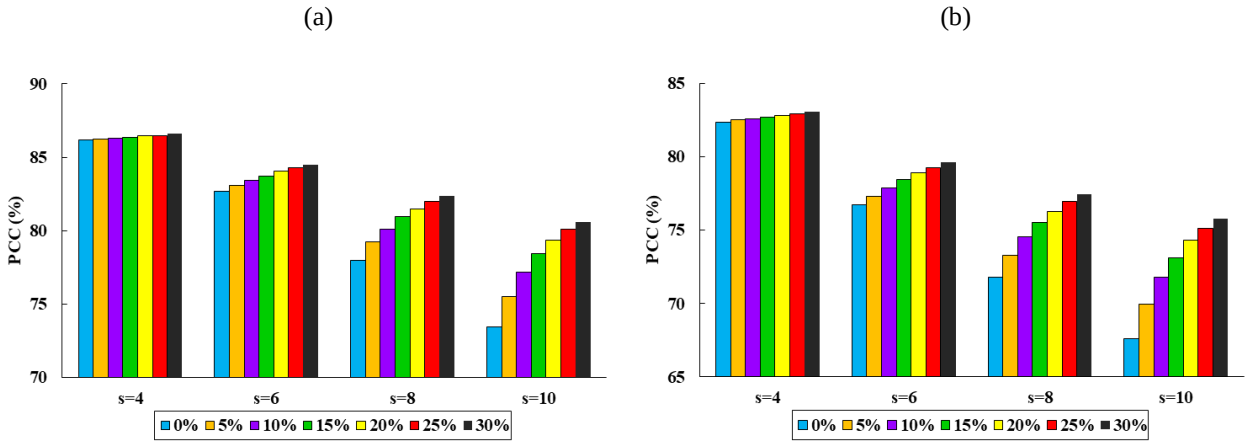


Fig. 11. PCC (in percentage) of the proposed PSA method (with 5%, 10%, 15%, 20%, 25%, and 30% point data, respectively) and the original PSA methods for four zoom factors: 4, 6, 8, and 10. (a) Xuzhou map. (b) South Carolina map.

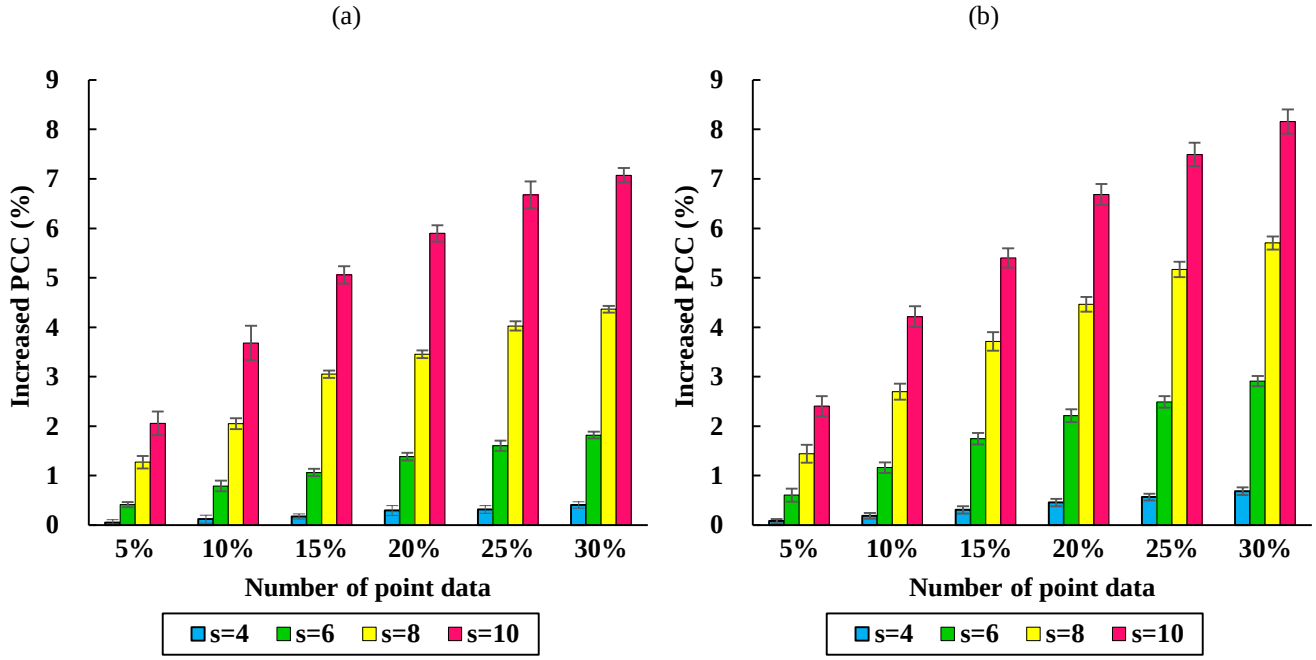


Fig. 12. Difference in PCC (in percentage) between the proposed PSA method (with 5%, 10%, 15% 20%, 25%, and 30% point data, respectively) and the original PSA methods for four zoom factors: 4, 6, 8, and 10 (Error lines represent the standard deviation). (a) Xuzhou map. (b) South Carolina map.

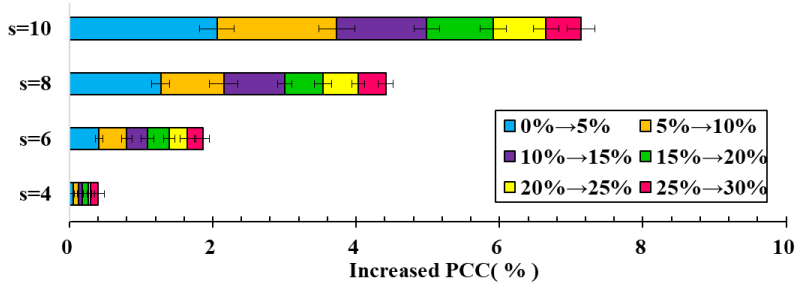
For clearer visual comparison between the SPM results produced with different schemes, two subareas of the two maps are also shown in the second line of Figs. 9 and 10. With the increase in the number of point data, the restoration of the land cover becomes increasingly accurate. For example, the road class in Fig. 9 and urban class in Fig. 10 are predicted as several isolated patches when no point data are considered, but more and more road pixels are correctly classified as the volume of point data increases. As a result, when 15% of the point data is used, the distribution of the road class in the SPM prediction is visually very close to the reference, where most of the continuous features have been reproduced. From the point data maps in Fig. 9(f)-(h) and Fig. 10(f)-(h), it can be seen that although the point data are relatively dispersed, the enhanced performances reveal that a small number of point data can also provide valuable information for SPM.

Fig. 11 displays the quantitative evaluation results in terms of the PCC for the original PSA (i.e., PSA with 0% point constraint) and the proposed method with 5% to 30% point constraints under four zoom factors. Obviously, as the number of points increases, the PCC increases gradually. To observe clearly the increase in SPM accuracy after including point constraints, the difference in PCC between the original PSA method and the proposed PSA method for each class is shown in Fig.12. The y-axis indicates the increase in PCC of the proposed method when compared to the original PSA method. For the two datasets, the increased PCC is larger than zero in each case and becomes larger as more point data are considered.

The bar charts in Fig. 13 show the increase of PCC for each additional 5% point data. For example, the length of the blue rectangle represents the increased PCC of PSA by adding 5% point data compared to PSA adding 0% (i.e., original PSA), while red represents the increased PCC when the point data increase from 25% to 30%. Undoubtedly, for a given zoom factor s, incorporating more point data as constraints is more helpful.

It is worth noting that, as the point data volume continues to increase, the length of rectangles for both datasets becomes smaller, which refers to a slowdown in the positive effect of the same number of data. More precisely, for each 5% additional data of s=10, the corresponding increased PCCs are 2.06%, 1.67%, 1.26%, 0.93%, 0.74%, and 0.49% for the Xuzhou map, and 2.4%, 1.79%, 1.36%, 1.19%, 0.83%, and 0.65% for the South Carolina map.

(a)



(b)

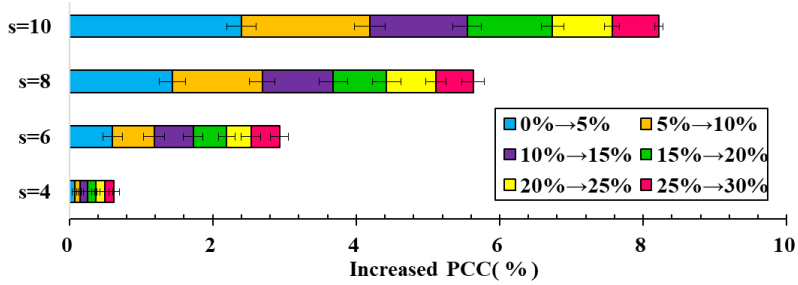


Fig. 13. Difference in PCC (in percentage) between PSA with α% points and PSA with (α-5) % points at four scales: 4, 6, 8, and 10 (0% means PSA without point data, namely, original PSA). (a) Xuzhou map. (b) South Carolina map.

2) Impact of the scale factors s: For the same volume of point data, it can be found in Fig. 12 that the increased PCC is larger for a larger scale factor *s*. That is, as *s* increases, although the accuracy of SPM is smaller in general, the advantage of the proposed PSA with point constraints method is more apparent. For example, when 15% point data are available for the Xuzhou map, the PCC is increased by around 1% for *s*=6, but by above 5% for *s*=10. For the South Carolina map, when the volume of the point data is 30%, the increased PCC is around 3% and 8% for *s*=6 and 10, respectively. The error lines in Figs. 12 and 13 represent the standard deviation of the accuracies, while the bar represents the average of the accuracies. It shows that the standard deviation is within the range of increased PCC. Moreover, for the same *s*, when more point data are available, the standard deviation is smaller, suggesting the performance is more stable.

3) Impact of spatial structure: As seen from the results in Figs. 12 and 13, the increased PCC for the South Carolina map is generally larger than for the Xuzhou map under the same scale factor *s*. To explain this phenomenon and explore the relationship between the effect of point constraints and the land cover class itself,

the increased accuracy of each class for *s*=8 is shown in Fig. 14. For the seven classes in the Xuzhou map, the gains of PCC become smaller in the order of shadow, tree, grass, road, roof, bare soil and water. For the four classes in the South Carolina map, the corresponding order is agriculture, urban, forest, and water. Checking the original reference maps of Xuzhou and South Carolina, it is obvious that the spatial distribution of the shadow and agriculture classes is the sparsest in the Xuzhou and South Carolina maps, respectively. On the contrary, the distribution of water is the most aggregated in both maps.

The Moran's *I* has been used widely as a measure of spatial autocorrelation in spatial statistics (Anselin, 2019; Overmars et al., 2003). Specifically, a lager Moran's *I* value means a larger spatial correlation and more aggregated spatial distribution. As listed in Table 1, the most aggregated water class in both maps has the largest Moran's *I*, which is consistent with visual inspection. The most dispersed class, that is, shadow in the Xuzhou map or agriculture in the South Carolina map, has the smallest Moran's *I*. Checking the order of Moran's *I* in Table I and the order of increased PCC in Fig. 14, we can see the interesting phenomenon that the class with the larger Moran's *I* always has the smaller increase in PCC. This is because the hypothesis of maximizing spatial correlation is more suitable for restoring aggregated features (i.e., land cover classes with greater autocorrelation) and, thus, the space for enhancing SPM of such features is limited. For features such as elongated and small-size patches whose spatial autocorrelation is small, however, there exists great uncertainty in SPM (i.e., the PCC is relatively smaller) when only the spatial dependence assumption is applied. In this case, the additional point constraints will provide more valuable information for reproduction of these features and the increased accuracy will be more obvious.

Inter-comparison between the two groups of Moran's *I* values for the two maps reveals that the values of the South Carolina map are generally smaller than whose for the Xuzhou map, suggesting the overall spatial autocorrelation is smaller for the former. Thus, the space for enhancing SPM is greater for the South Carolina map, as seen in Figs. 12 and 13.

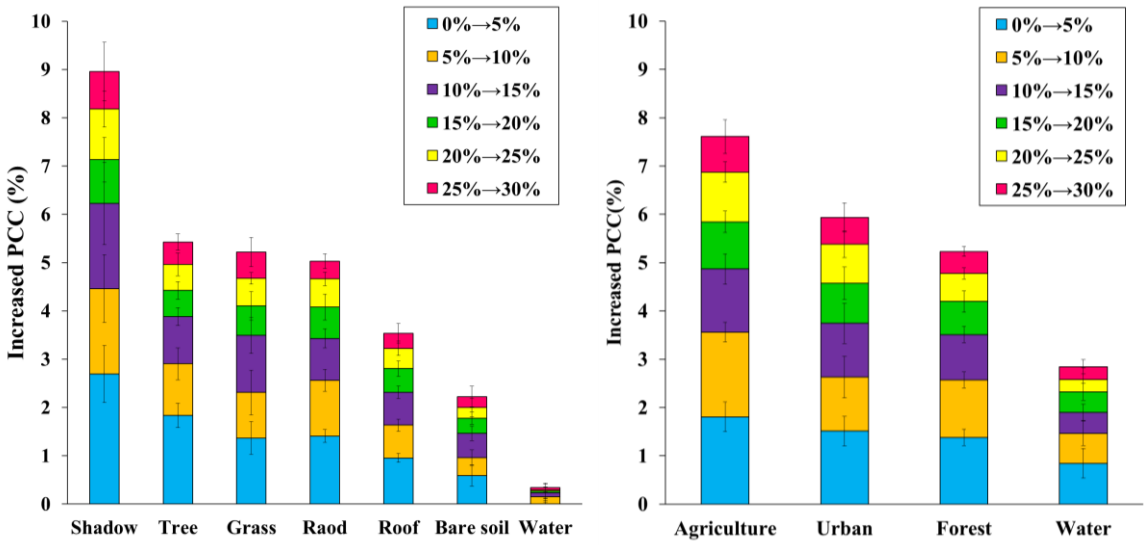


Fig. 14. Difference in PCC for each land cover class between PSA with α% points and PSA with (α-5) % points (s=8, 0% means PSA without point data, that is, original PSA). (a) Xuzhou map. (b) South Carolina map.

Table 1 Moran's I of each land cover class in the two original fine spatial resolution classification maps

Xuzhou map	Class	Shadow	Tree	Grass	Road	Roof	Bare soil	Water
	Moran's *I*	0.6573	0.7226	0.7384	0.7832	0.7503	0.8836	0.9556
	Rank	1	2	3	5	4	6	7
South Carolina map	Class	Agriculture	Urban	Forest	Water			
	Moran's *I*	0.6282	0.6644	0.6896	0.9096			
	Rank	1	2	3	4			

4) Impact of attribute errors in point data: In the above experiments, it was assumed that the point data do not contain any uncertainty and the class labels are completely accurate for all informed pixels. In reality, however, errors exist commonly in such point data due to data source contamination (e.g., false registration data in social software) and interpretation processes (e.g., the point to be interpreted may contain more than one class in reality). To explore the impact of uncertainty of point data on SPM results, simulated incorrect data were used in this experiment. Specifically, part of the set of informed pixels was simulated incorrectly by artificially changing the original class labels to others. As illustrated in previous experiments, the maximum amount of point data accounts for 30% of the total number of sub-pixels. Thus, in this section, the error rate was increased from 5% to 30%, with a step of 5%. When the error rate is 30%, it means the input data are completely wrong.

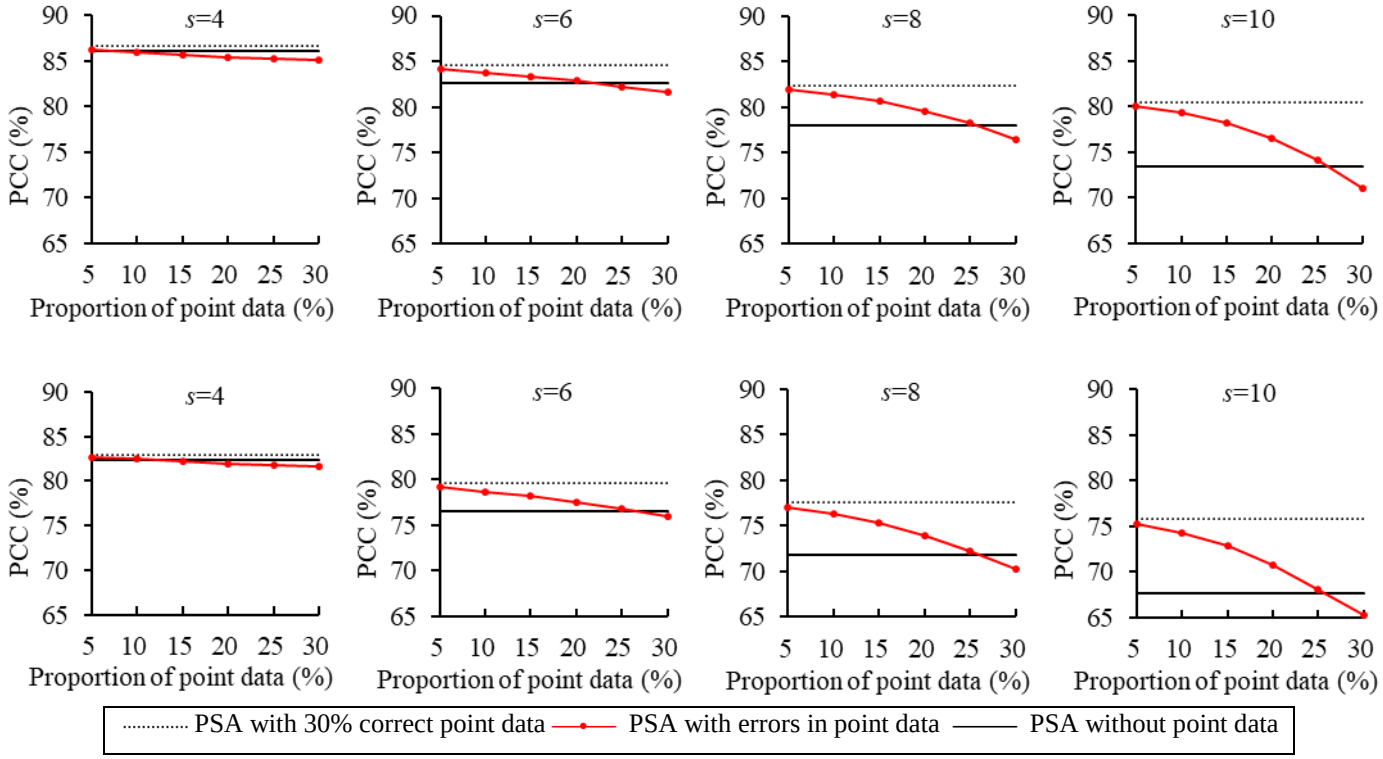


Fig. 15. PCC of SPM results with errors in point data (excluding point data for PCC calculation). (Line 1) Xuzhou. (Line 2) South Carolina.

The PCC of the proposed SPM method that considers point data containing errors is shown in Fig. 15. For clear comparison, the PCCs of the original PSA and PSA with completely correct point data are also provided. From this figure, we can observe clearly that as the error increases, the PCCs for both datasets decrease. Undoubtedly, the erroneous point data have a negative impact and the increased slope indicates a faster decrease in accuracy. Interestingly, the accuracy of PSA with errors is still greater than the original PSA method until the proportion of errors increases to a certain value, see the intersection of the red and black lines in Fig. 15. For example, for the Xuzhou results in line 1 of Fig. 15, when the proportion of error reaches 10% and 20% for s=8, the PCCs are 3.39% and 1.63% larger than for the original PSA method. The PCC is decreased to be smaller than original PSA when the proportion of error exceeds 25%. Therefore, although there exist unavoidable errors in the point data, the use of point data is still advantageous for SPM if the proportion of data in error is not very large.

(a) (b)

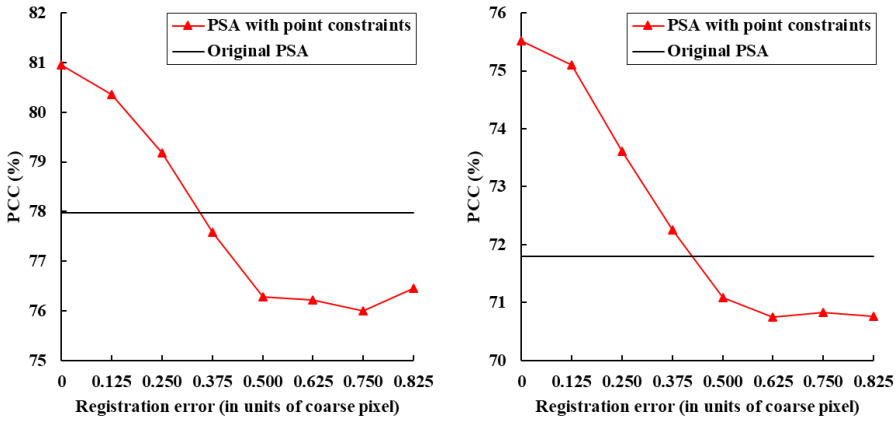


Fig. 16. PCC (in percentage) of the proposed PSA method with 15% point data for various registration errors (s=8). (a) Xuzhou map. (b) South Carolina map.

5) Impact of geo-locational errors in point data: The experiments above assumed that the point data have a completely reliable spatial co-registration to the image. In reality, however, there always exists a geo-locational mismatch between data acquired from different sources (e.g., mismatch between remote sensing images and data from social software). To analyze the influence of point data with geo-locational errors, we simulated points that were mis-registered with the coarse proportion images. Specifically, the original real locations of points were changed by shifting n/s (n=0, 1…, n-1) coarse pixel both horizontally and vertically. In this section, the proposed PSA method was performed with 15% point data at s=8 for both Xuzhou and South Carolina datasets. The PCCs of SPM using points with geo-locational errors (i.e., registration errors) and the original PSA method are given in Fig. 16. It is worth mentioning that similarly to the previous experiments, the known points were excluded from the accuracy assessment. Obviously, the registration error imposes a negative effect on the performance of the proposed method (in both cases, the PCC for points with registration error is smaller than that without registration error), as incorrect spatial information is used. It should be noted that, however, when the error is not very large (e.g., below 0.5 coarse pixel in the South Carolina case), the PCC of PSA using point constraints is still larger than that of the original PSA method. Thus, the use of point data is helpful for SPM even when geo-locational errors exist in the point data in practice, under the condition

that the error is not large or necessary pre-processing has been performed to reduce the mismatch to a certain value (e.g., <0.5 coarse pixel in this case study).

6) Comparison with other SPM methods: In this section, the proposed SPM with point constraints method was compared with the advanced RBF method (Wang et al., 2014a). The RBF method consists of the fine spatial resolution soft class attribute value and hard class label estimation. The same point data were used at the hard class label estimation stage for the RBF method. The PCCs of different methods for two datasets (s=8) are given in Fig.17. It can be seen that the PCCs of the RBF are larger than that of the original PSA, but much smaller than that of the proposed method. As the number of point data increases, the advantage of the proposed method is more evident. More explicitly, using 5%-30% (with an interval of 5%) point data, the PCC gains of the proposed method over RBF are 1.17%, 2.07%, 2.8%, 3.32%, 3.73%, and 4.08% for the Xuzhou map, and 0.06%, 1.18%, 2.13%, 2.84%, 3.47%, and 3.92% for the South Carolina map. This is because in the RBF method, the spatial correlation is considered only at the coarse pixel scale, and the point data were not well utilized. As for the proposed PSA-based method, the spatial correlation is considered at the sub-pixel scale and the known points provide valuable spatial information for their surrounding sub-pixels during the spatial optimization process.

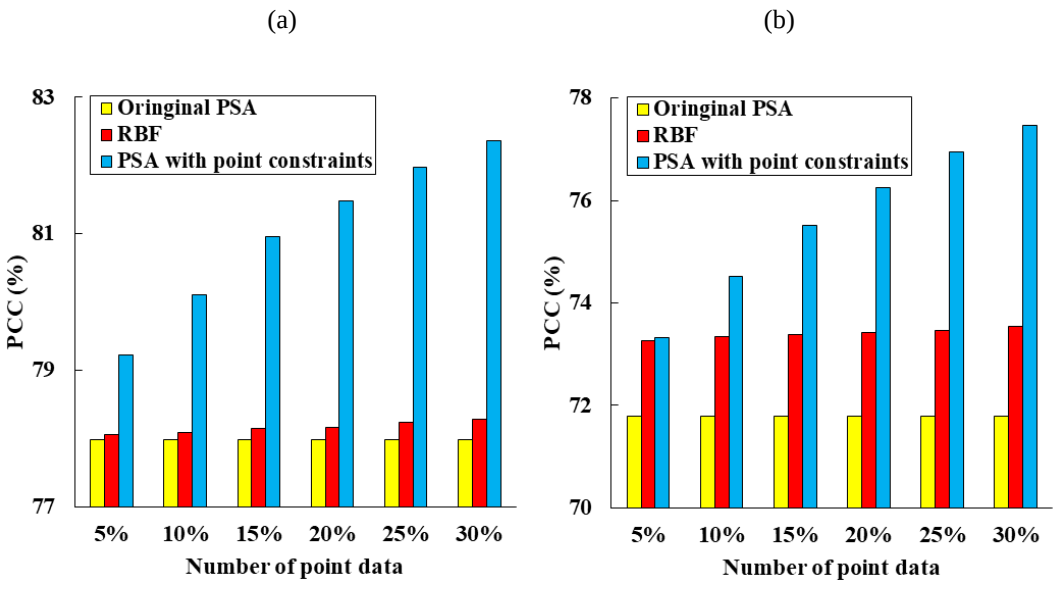


Fig. 17. PCC (in percentage) of the original PSA method, the RBF method and the PSA method incorporating point data (with 5%, 10%, 15% 20%, 25%, and 30% point data, respectively) (s=8). (a) Xuzhou map. (b) South Carolina map.

7) Impact of errors in coarse proportions: In the experiments, it is assumed that coarse proportions (i.e., input of SPM) are estimated without error. However, spectral unmixing is still an open problem. Given that the quality of the estimated class proportions plays an important role in SPM, especially for methods such as PSA that fix the number of sub-pixels for each class according to the input coarse proportions, in this experiment, the performance of the proposed method was explored when errors exist in coarse proportions. Specifically, proportion images with different errors were simulated and the case of zoom factor s=8 and 25% point data was considered. Errors with a Gaussian distribution were added to the ideal proportions to generate the proportion images with errors. The accuracy of the simulated proportions was quantified by the root mean square error (RMSE) between the simulated proportions and the reference proportions. The proportion errors with RMSE= 0.025, 0.05, 0.075 and 0.1 were simulated, and details of the simulation process can be found in Li et al. (2014). Fig. 18 shows that proportion errors cause a decrease in SPM accuracy for both the original PSA and the proposed method. However, the PCCs of the proposed method are still larger than that of the original PSA for both Xuzhou and South Carolina maps. Thus, the benefits of using point constraints are still noticeable when errors exist in the coarse proportions.

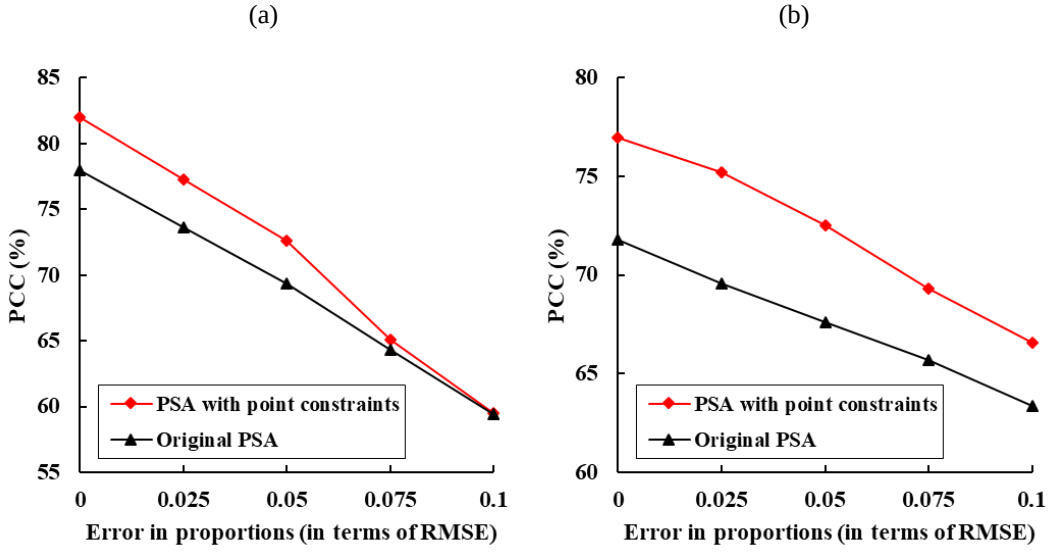


Fig. 18. PCC (in percentage) of the proposed PSA method with 25% point data and the original PSA method under proportion errors (s=8). (a) Xuzhou map. (b) South Carolina map.

3.3. Experiment for L-resolution Case

The L-resolution case refers to pixels that are larger than the objects of interest (Atkinson, 2009). In the experiment, the Granada dataset was used. The coarse proportion images were synthesized by degrading the Granada land cover map with a factor of 9. Fig. 19(b) shows the coarse proportion image for the vegetation class, where the size of most of the vegetation patches is smaller than the pixels, that is, the vegetation falls within the mixed pixels. Visually, the object (i.e., vegetation) can hardly be identified accurately at such a coarse spatial resolution. The task of this experiment is to restore the land cover map that has the same spatial resolution as Fig. 19(a), based on the input of the image in Fig. 19(b) and the available point data.

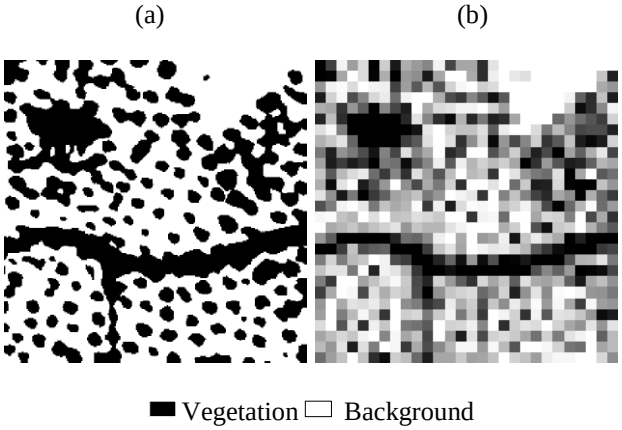


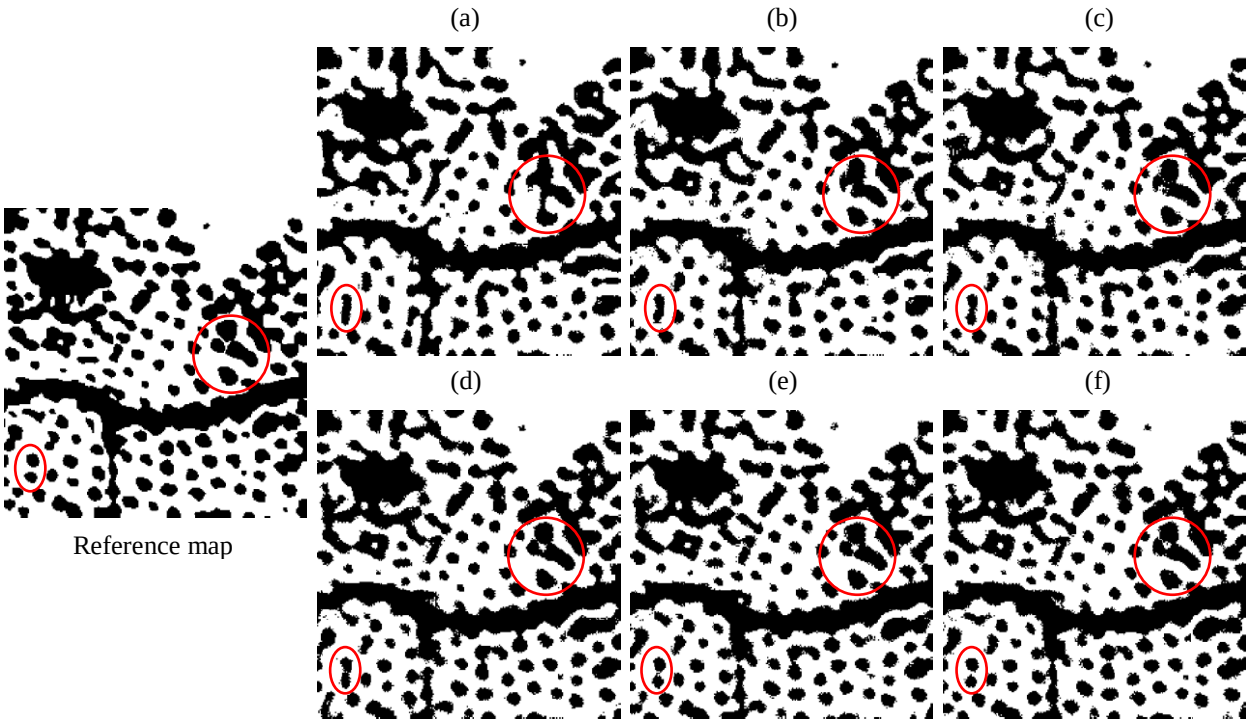
Fig. 19. Granada land cover map and proportion image. (a) Reference map. (b) Proportion image of vegetation created by degrading (a) with a scale factor of 9 (28 ×28pixels).

The SPM results produced with different number of point data are shown in Fig. 20. As demonstrated visually in Fig. 20, the original isolated patches of vegetation are incorrectly merged by the original PSA method: see the examples marked in red circles. This is because in L-resolution case, the targets smaller than pixels are connected as a whole based only on spatial dependence. When point data are available, the prediction is enhanced noticeably. For example, the results in Fig. 20(b) show that PSA with 6% point data can restore the vegetation class more accurately, where some small-sized vegetation patches can be separated. As the number of point data increases, the prediction becomes increasingly accurate, as shown in Fig. 20(c)-(f). When 30% point data were used, the prediction is visually highly similar to the reference. This is because the point data,

618 especially for those located at the boundaries of small-sized objects, can provide important guidance in
619 sub-pixel class determination for their sub-pixel neighbors. Thus, small objects (i.e., smaller than a coarse pixel)
620 can be correctly separated when sufficient point data are available.

621 Quantitative assessments by PCC and increased PCC for the L-resolution case are provided in Fig. 21.
622 Clearly, as the number of point data increases, the SPM accuracy increases, which is consistent with the visual
623 results in Fig. 20. Furthermore, it is seen from Fig. 21(b) that when the volume of point data increases from 6%
624 to 30% with a step of 6%, the increase in PCC is 7.98%, 13.79%, 16.26%, 18.05%, and 18.95%, respectively.
625 Moreover, as the number of point data increases, the increased in PCC tends to decrease, but is still larger than
626 zero. That is, the space for accuracy increase is limited when sufficient point data are available, which is
627 consistent with the conclusion in the previous experiments for the H-resolution case.

628



629

630 Fig. 20. SPM results of Granada map (s=9). (a) Original PSA. (b) PSA with 6% point data. (c) PSA with 12% point data. (d) PSA with
631 18% point data. (e) PSA with 24% point data. (f) PSA with 30% point data.

632

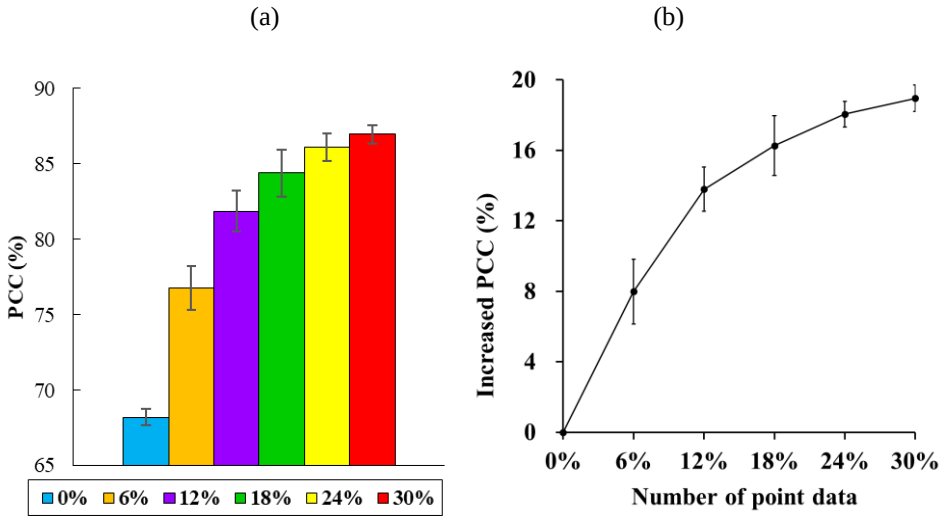


Fig. 21. Accuracy of the SPM methods for the L-resolution case (s= 9) (Error lines represent the standard deviation). (a) PCC (in percentage) of the proposed method. (b) Difference in PCC between the proposed method (with 6%, 12% 18%, 24%, and 30% point data, respectively) and the original PSA method.

4. Discussion

This paper proposed to utilize point data to enhance SPM. The goal was realized based on the extension of PSA. The proposed method provides a new option for increasing the accuracy of SPM by using auxiliary information. This section opens discussions on both the data and algorithm.

4.1. Alternatives to PSA

Similar to many SPM methods, PSA is slavish to the predictions (i.e., coarse proportions) of the pre-spectral unmixing, which means the errors in proportions adversely affect the final SPM predictions directly (see the results in Fig. 18). This is a limitation of many existing methods including PSA. As mentioned earlier, it is a natural choice to consider SPM methods that characterize the spatial dependence between sub-pixels, as the known point data are at sub-pixel resolution. Methods falling into this type include PSA, HNN and MAP. Thus, it seems worthwhile to develop alternatives based on HNN and MAP. Different to the PSA method, MAP and HNN consider the spatial goal of maximizing spatial dependence and the coarse pixel proportion constraint

simultaneously in their models, see Eq. (1). Thus, the contribution of the proportion constraint term can be tuned to reduce the influence of proportion error theoretically. On the other hand, when the proportion constraint is less emphasized, MAP and HNN methods generally produce over-smooth predictions where the small patches and elongated features cannot be restored satisfactorily and even disappear in the predicted land cover map. This is not the case for PSA, where the small patches and elongated features can be retained to some extent. For potential solutions to incorporating point constraints in SPM by MAP and HNN, the key potential solutions to this issue may be to develop schemes that adaptively control the contributions of coarse proportion constraints. For example, for local areas with weak autocorrelation, greater proportion constraints should be considered to guarantee the reproduction of pixels for small patches and elongated features (i.e., in the general model in Eq. (1), the weight k_2 should be set to a larger value). How to balance the contributions of point and coarse proportion constraints would be a critical issue (i.e., weights k_2 and k_3 in Eq. (2) should be set to appropriate values).

It is also worth developing related methods to incorporate point data for another main type of SPM methods, namely the one that characterizes the spatial dependence between sub-pixel and neighboring coarse pixels. This type of method does not need an iteration-based optimization process and can realize SPM quickly. The challenge is that the coarse neighbors and point data are at different scales. In this paper, a preliminary attempt was made with RBF in the experiment, but the point data were not sufficiently integrated into the SPM process. It would be a key issue to develop models to characterize the spatial scale difference of data. The kriging-based SPM method (Verhoeye and De Wulf, 2002) considers the spatial relation between sub-pixel and pixels. In recent years, it has been extended to area-to-point kriging (Wang et al., 2015) to take the scale difference (also termed change of support in geostatistics) into account. It would be of great interest to develop area-to-point kriging-based models to incorporate point data in SPM, such as a co-kriging model. In this case, however, the kriging equation needs to be constructed locally for each pixel (rather than a single equation in area-to-point kriging-based image downscaling), as the spatial distribution of point data is not fixed and always varies in each local window. Moreover, since some locations may lack point data, for example, due to difficulties in data

acquisition, the size of local window needs to be determined adaptively to incorporate sufficient point data. Thus, the complexity and computational burden would probably increase when point data are considered.

Multiple point statistics (MPS) (Mariethoz et al., 2010) is a relatively new branch of geostatistics. It considers the relation (both spatial distance and direction) between multiple points based on a pre-defined template, rather than the vector separation between only two points, as in kriging. Thus, MPS is particularly advantageous for restoration of spatially continuous patterns such as linear features, and it has been applied widely in geology (e.g., ground water flow mapping (Mariethoz et al., 2010)). MPS is implemented based on the availability of spatially sparse point data (also termed hard data) and fills in the data at other locations with the aid of available point data, which fits well with the case of this paper. However, MPS requires a training image at the spatial resolution of the point data to search for data events and find the appropriate data for filling, which hampers its application in SPM. Therefore, the proposed SPM method accounting for point data in this paper is advantageous in terms of implementation as it does not need any training image. Once such training data are available, however, it would be favorable to develop MPS models to incorporate point data in SPM.

For the H- and L- resolution cases in the experiments, it needs to be clarified that the definition of H- and L-resolution was not in an absolute sense. More precisely, they are defined based on a relative flexible assumption according to the proportion of H- and L- resolution pixels in the entire image. In the H-resolution cases, in the coarse Xuzhou and South Carolina images, most (but not all) of the land cover patches occupy over one coarse pixel. In the L- resolution case, most of the land cover patches (i.e., vegetation of the Granada dataset in Fig. 19) fall within one coarse pixel.

4.2. Uncertainties in Point Information

To eliminate the influence of uncertainty in point information and concentrate on the performance of the proposed model itself, the point data were simulated by sampling randomly from the reference land cover map in the experiments. In real applications, there are several considerations about the uncertainties in point information, which are discussed as follows.

703 *1) Uncertainties in data source:* As mentioned earlier, there are various sources for point data and
704 uncertainties always exist in the source data, that is, the issue of reliability of spatial data. For example, there
705 may be false registration in social media data such as Weibo, Instagram, Twitter, etc.. Moreover, time
706 inconsistencies may exist between the observed image for SPM and the acquired point data, where land cover
707 changes may occur. The point data acquired from different sources may vary greatly. Pre-processing may be
708 undertaken to harmonize them and blend them into a single SPM framework. Potential solutions may be to
709 construct models (e.g., linear or non-linear regression such as machine learning-based models) to obtain the
710 relation between the points of the same location but from different sources, which can be used to update the
711 point data at other locations.

712 *2) Uncertainties in data sampling process:* The spatial locations of simulated point data in the experiments
713 are distributed randomly. However, in actual field collection, there may be some requirements and conditions
714 when sampling the data. For example, equal interval sampling may be required in some applications. On the
715 other hand, in some remote areas such as mountains, it may be more difficult to obtain sufficient data, while in
716 areas with vehicular transport, such as cities, the acquisition of point data is likely to be simpler. That is, the
717 number of point data in some areas may be large but in other areas the points may be spatially very sparse. It
718 would be important to design appropriate sampling strategies. The impact of the spatial distribution of the
719 obtained point data is also worthy of further study.

720 *3) Uncertainties in spatial support:* In the experiments, the spatial coverage of the input points is assumed to
721 be in agreement with the target sub-pixels. In reality, however, the point may not be strictly coincident with one
722 sub-pixel spatially, that is, uncertainties exist in the spatial support of the point data itself. In an extreme case,
723 suppose the sampling point covers a 1 m^2 area and contains 0.5 m^2 bare soil and 0.5 m^2 grass. The sub-pixel
724 size at the target fine spatial resolution is larger than the 1 m^2 point. The question is which class should be
725 assigned to this point? Aiming at the mismatch of spatial support between point data and target sub-pixel, a
726 potential solution may be to upscale the point data to match the support of the sub-pixel. Another problem is the
727 inconsistency of support size from different data sources. Specifically, the support size of social media data is

728 always difficult to define (assumed to be a small point in general), and the spatial coverage of field data can be
729 extensive. How to match the size of points from different sources would be an interesting issue to address.

730 *4) Uncertainties in class attribute of point data:* Influenced by the complex distribution of land cover class
731 and the size of spatial support as discussed above in 3), it is difficult to assign the point to a single class as it
732 may cover multiple classes. That is, the hard class attribute of point data always contains uncertainties.
733 Confronted with this problem, it would be interesting to develop models that can take the uncertainty of the
734 hard attribute into account to obtain more reliable SPM results. For example, the class attribute of point data
735 can be expressed mathematically by a vector, with each element representing the proportion of a class. The
736 effective model needs to consider the soft attributes (rather than the hard attributes or class indicators in this
737 paper) together.

738 *5) Usage rate of point information:* In practical applications, not all of the acquired point information may be
739 valid or can be fully utilized. Point data located in pure pixels (such as the point data for the roof class in Fig.
740 9(f)-(h)), conveys no useful information because the SPM process for pure pixels is very simple and does not
741 require any additional information: all sub-pixels in pure pixels are allocated to the same class as the pixel.
742 Nevertheless, point data located within mixed pixels are much more informative and are considered as
743 effective points. Most of these points are located in boundary areas. To investigate the effective usage rate of
744 point information in the experiments, the proportion of effective points to 20% known points is calculated in
745 Fig. 22. Obviously, most points are effective and for each class, the number of effective points increases with
746 the increase in zoom factor. On the other hand, since there may be multiple sources for point data, it is
747 necessary to avoid repeated acquisition of point data in the same location. Moreover, if that occurs, there is a
748 need to explore how to select the most appropriate sources.

749

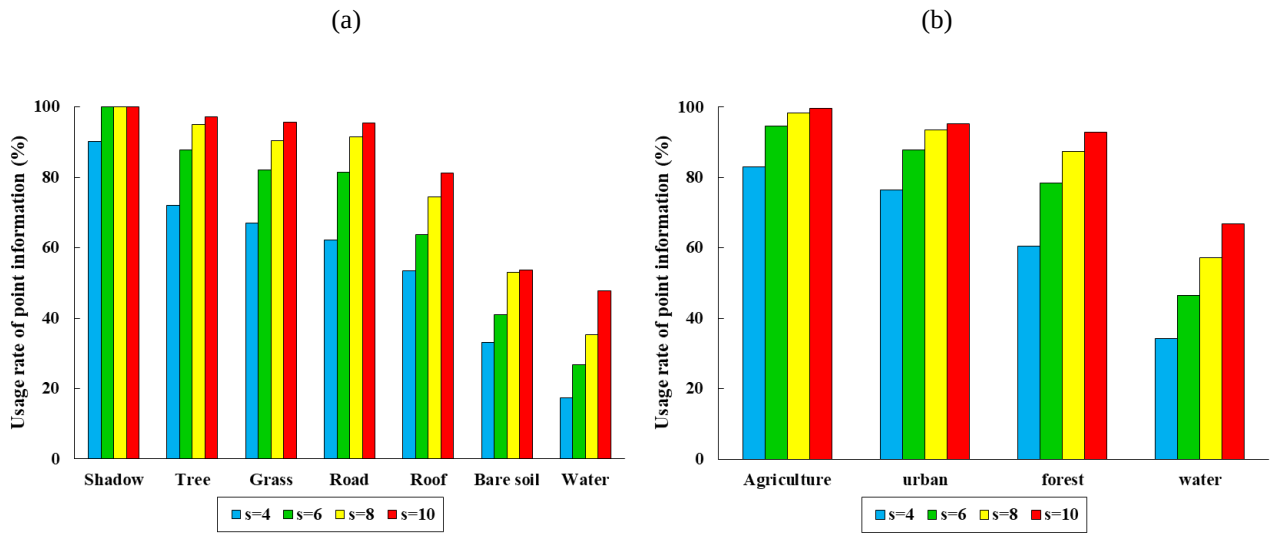


Fig. 22. The proportion of effective points to 20% known points for each land cover class at four scales. (a)Xuzhou map. (b)South Carolina map.

5. Conclusion

SPM is an ill-posed problem and it is necessary to explore useful information from additional data for possible enhancement. In this paper, a novel type of supplementary data, namely, point data, was proposed for increasing the accuracy of SPM. The point data are incorporated in SPM by the PSA method, where the data can exert additional sub-pixel level constraints to reduce the uncertainty in SPM. The proposed method was tested with three datasets and the conclusions are summarized as follows.

1) The proposed point constraints-based SPM method provides a satisfactory solution for enhancing SPM. From both visual and quantitative assessments, it was found that the point information can help to produce more accurate sub-pixel maps. The accuracy is also greater than the advanced RBF-based SPM method.
2) As the number of point data increases, the accuracy increases correspondingly, but the accuracy gain becomes smaller.
3) The advantages of using point data are more obvious for the case involving a large zoom factor.
4) The point data are more beneficial for land cover classes with small spatial autocorrelation, and for

classes with large autocorrelation, the increase in accuracy is relatively smaller.

5) Even when point data contain attribute or geo-locational errors, they can still be useful for enhancing SPM, but the proportion of errors cannot be very large.
6) The proposed method accounting for point constraints is a general method suitable for both the H- and L-resolution cases.
7) When coarse proportions contain errors, the SPM accuracy is reduced, but the use of point constraints is still beneficial.

Acknowledgment

This work was supported by National Natural Science Foundation of China under Grant 41971297, Fundamental Research Funds for the Central Universities under Grant 02502150021 and Tongji University under Grants 02502350047.

References

Anselin, L. (2019). A local indicator of multivariate spatial association: extending Geary's C. Geographical Analysis, 51, 133-150

Aplin, P., & Atkinson, P.M. (2001). Sub-pixel land cover mapping for per-field classification. International Journal of Remote Sensing, 22, 2853-2858

Anselin, L. (2019). A local indicator of multivariate spatial association: Extending geary's c. Geographical Analysis, 51, 133-150.

Aplin, P., & Atkinson, P.M. (2001). Sub-pixel land cover mapping for per-field classification. International Journal of Remote Sensing, 22, 2853-2858.

Ardila, J.P., Tolpekin, V.A., Bijker, W., & Stein, A. (2011). Markov-random-field-based super-resolution mapping for identification of urban trees in vhr images. ISPRS Journal of Photogrammetry and Remote Sensing, 66, 762-775.

Atkinson, P.M., Cutler, M.E.J., & Lewis, H. (1997). Mapping sub-pixel proportional land cover with avhrr imagery. International Journal of Remote Sensing, 18, 917-935.

796 Atkinson, P.M. (2004). Super-resolution land cover classification using the two-point histogram. GeoENV IV-Geostatistics for
797 Environmental Applications, 13, 15.
798 Atkinson, P.M. (2005). Sub-pixel target mapping from soft-classified, remotely sensed imagery. Photogrammetric Engineering &
799 Remote Sensing, 71, 839-846.
800 Atkinson, P.M. (2009). Issues of uncertainty in super-resolution mapping and their implications for the design of an inter-comparison
801 study. International Journal of Remote Sensing, 30, 5293-5308.
802 Bioucas-Dias, J.M., Plaza, A., Dobigeon, N., Parente, M., Du, Q., Gader, P., & Chanussot, J. (2012). Hyperspectral unmixing
803 overview: Geometrical, statistical, and sparse regression-based spproaches. IEEE Journal of Selected Topics in Applied Earth
804 Observations and Remote Sensing, 5, 354-379.
805 Bizzi, S., Demarchi, L., Grabowski, R.C., Weissteiner, C.J., & Van de Bund, W. (2016). The use of remote sensing to characterise
806 hydromorphological properties of european rivers. Aquatic Sciences, 78, 57-70.
807 Boucher, A., Kyriakidis, P.C., & Cronkite-Ratcliff, C. (2008). Geostatistical solutions for super-resolution land cover mapping. IEEE
808 Transactions on Geoscience and Remote Sensing, 46, 272-283.
809 Chen, Y., Ge, Y., Heuvelink, G.B.M., An, R., & Chen, Y. (2018a). Object-based superresolution land-cover mapping from remotely
810 sensed imagery. IEEE Transactions on Geoscience and Remote Sensing, 56, 328-340.
811 Chen, Y., Ge, Y., Chen, Y., Jin, Y., & An, R. (2018b). Sub-pixel land cover mapping using multiscale spatial dependence. IEEE
812 Transactions on Geoscience and Remote Sensing, 56, 5097-5106.
813 Eastman, J.R., & Laney, R.M. (2002). Bayesian soft classification for sub-pixel analysis: A critical evaluation. Photogrammetric
814 Engineering & Remote Sensing, 68, págs. 1149-1154.
815 Fisher, P. (1997). The pixel: A snare and a delusion. International Journal of Remote Sensing, 18, 679-685.
816 Foody, G.M., & Doan, H.T.X. (2007). Variability in soft classification prediction and its implications for sub-pixel scale change
817 detection and super resolution mapping. Photogrammetric Engineering & Remote Sensing, 73, 923-933.
818 Frias-Martinez, V., & Frias-Martinez, E. (2014). Spectral clustering for sensing urban land use using twitter activity. Engineering
819 Applications of Artificial Intelligence, 35, 237-245.
820 Fritz, S., McCallum, I., Schill, C., Perger, C., See, L., Schepaschenko, D., van der Velde, M., Kraxner, F., & Obersteiner, M. (2012).
821 Geo-wiki: An online platform for improving global land cover. Environmental Modelling & Software, 31, 110-123.
822 Jia, Y., Ge, Y., Chen, Y., Li, S., Heuvelink, B.M.G., & Ling, F. (2019). Super-resolution land cover mapping based on the
823 convolutional neural network. Remote Sensing, 11, 1815.
824 Jin, H., Mountrakis, G., & Li, P. (2012). A super-resolution mapping method using local indicator variograms. International Journal
825 of Remote Sensing, 33, 7747-7773.

826 Jokar Arsanjani, J., Helbich, M., Kainz, W., & Darvishi Boloorani, A. (2013). Integration of logistic regression, markov chain and
827 cellular automata models to simulate urban expansion. International Journal of Applied Earth Observation and Geoinformation,
828 21, 265-275.
829 Kasetkasem, T., Arora, M.K., & Varshney, P.K. (2005). Super-resolution land cover mapping using a markov random field based
830 approach. Remote Sensing of Environment, 96, 302-314.
831 Keshava, N., & Mustard, J.F. (2002). Spectral unmixing. IEEE Signal Processing Magazine, 19, 44-57.
832 Li, X., Ling, F., Du, Y., Feng, Q., & Zhang, Y. (2014). A spatial–temporal hopfield neural network approach for super-resolution
833 land cover mapping with multi-temporal different resolution remotely sensed images. ISPRS Journal of Photogrammetry and
834 Remote Sensing, 93, 76-87.
835 Li, X., Ling, F., Foody, G.M., Ge, Y., Zhang, Y., & Du, Y. (2017). Generating a series of fine spatial and temporal resolution land
836 cover maps by fusing coarse spatial resolution remotely sensed images and fine spatial resolution land cover maps. Remote
837 Sensing of Environment, 196, 293-311.
838 Ling, F., Xiao, F., Du, Y., Xue, H.P., & Ren, X.Y. (2008). Waterline mapping at the sub-pixel scale from remote sensing imagery
839 with high‐resolution digital elevation models. International Journal of Remote Sensing, 29, 1809-1815.
840 Ling, F., Du, Y., Xiao, F., Xue, H., & Wu, S. (2010). Super-resolution land-cover mapping using multiple sub-pixel shifted remotely
841 sensed images. International Journal of Remote Sensing, 31, 5023-5040.
842 Ling, F., Li, X., Xiao, F., Fang, S., & Du, Y. (2012). Object-based sub-pixel mapping of buildings incorporating the prior shape
843 information from remotely sensed imagery. International Journal of Applied Earth Observation and Geoinformation, 18,
844 283-292.
845 Ling, F., Li, X., Du, Y., & Xiao, F. (2013). Sub-pixel mapping of remotely sensed imagery with hybrid intra- and inter-pixel
846 dependence. International Journal of Remote Sensing, 34, 341-357.
847 Ling, F., Du, Y., Li, X., Zhang, Y., Xiao, F., Fang, S., & Li, W. (2014). Superresolution land cover mapping with multiscale
848 information by fusing local smoothness prior and downscaled coarse fractions. IEEE Transactions on Geoscience and Remote
849 Sensing, 52, 5677-5692.
850 Ling, F., Zhang, Y., Foody, G.M., Li, X., Zhang, X., Fang, S., Li, W., & Du, Y. (2016). Learning-based superresolution land cover
851 mapping. IEEE Transactions on Geoscience and Remote Sensing, 54, 3794-3810.
852 Ling, F., & Foody, G.M. (2019). Super-resolution land cover mapping by deep learning. Remote Sensing Letters, 10, 598-606.
853 Liu, X., He, J., Yao, Y., Zhang, J., Liang, H., Wang, H., & Hong, Y. (2017). Classifying urban land use by integrating remote sensing
854 and social media data. International Journal of Geographical Information Science, 31, 1675-1696.

855 Mariethoz, G., Renard, P., & Straubhaar, J. (2010). The direct sampling method to perform multiple-point geostatistical simulations.
856 Water Resources Research, 46, 1-14.
857 Mertens, K.C., Verbeke, L.P.C., Ducheyne, E.I., & De Wulf, R.R. (2003). Using genetic algorithms in sub-pixel mapping.
858 International Journal of Remote Sensing, 24, 4241-4247.
859 Mertens, K.C., de Baets, B., Verbeke, L.P.C., & De Wulf, R.R. (2006). A sub-pixel mapping algorithm based on sub-pixel/pixel
860 spatial attraction models. International Journal of Remote Sensing, 27, 3293-3310.
861 Muad, A.M., & Foody, G.M. (2012). Impact of land cover patch size on the accuracy of patch area representation in hnn-based super
862 resolution mapping. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 5, 1418-1427.
863 Nguyen, Q.M., Atkinson, P.M., & Lewis, H.G. (2005). Superresolution mapping using a hopfield neural network with lidar data.
864 IEEE Geoscience and Remote Sensing Letters, 2, 366-370.
865 Nguyen, Q.M., Atkinson, P.M., & Lewis, H.G. (2011). Super-resolution mapping using hopfield neural network with panchromatic
866 imagery. International Journal of Remote Sensing, 32, 6149-6176.
867 Overmars, K.P., De Koning, G., & Veldkamp, A. (2003). Spatial autocorrelation in multi-scale land use models. Ecological
868 modelling, 164, 257-270.
869 Song, M., Zhong, Y., Ma, A., & Feng, R. (2019). Multiobjective sparse sub-pixel mapping for remote sensing imagery. IEEE
870 Transactions on Geoscience and Remote Sensing, 57, 4490-4508.
871 Tatem, A.J., Lewis, H.G., Atkinson, P.M., & Nixon, M.S. (2001). Super-resolution target identification from remotely sensed images
872 using a hopfield neural network. IEEE Transactions on Geoscience and Remote Sensing, 39, 781-796.
873 Tolpekin, V.A., & Stein, A. (2009). Quantification of the effects of land-cover-class spectral separability on the accuracy of
874 markov-random-field-based superresolution mapping. IEEE Transactions on Geoscience and Remote Sensing, 47, 3283-3297.
875 Verhoeye, J., & De Wulf, R.R. (2002). Land cover mapping at sub-pixel scales using linear optimization techniques. Remote Sensing
876 of Environment, 79, 96-104.
877 Wang, P., Wang, L., Mura, M.D., & Chanussot, J. (2017). Using multiple sub-pixel shifted images with spatial–spectral information
878 in soft-then-hard sub-pixel mapping. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 10,
879 2950-2959.
880 Wang, Q., Wang, L., & Liu, D. (2012). Integration of spatial attractions between and within pixels for sub-pixel mapping. Journal of
881 Systems Engineering and Electronics, 23, 293-303.
882 Wang, Q., Shi, W., & Atkinson, P.M. (2014a). Sub-pixel mapping of remote sensing images based on radial basis function
883 interpolation. ISPRS Journal of Photogrammetry and Remote Sensing, 92, 1-15.

884 Wang, Q., Shi, W., & Wang, L. (2014b). Indicator cokriging-based sub-pixel land cover mapping with shifted images. IEEE Journal
885 of Selected Topics in Applied Earth Observations and Remote Sensing, 7, 327-339.
886 Wang, Q., Shi, W., Atkinson, P.M., & Zhao, Y. (2015). Downscaling modis images with area-to-point regression kriging. Remote
887 Sensing of Environment, 166, 191-204.
888 Wang, Q., Shi, W., & Atkinson, P.M. (2016). Spatiotemporal sub-pixel mapping of time-series images. IEEE Transactions on
889 Geoscience and Remote Sensing, 54, 5397-5411.
890 Wu, S., Ren, J., Chen, Z., Jin, W., Liu, X., Li, H., Pan, H., & Guo, W. (2018). Influence of reconstruction scale, spatial resolution and
891 pixel spatial relationships on the sub-pixel mapping accuracy of a double-calculated spatial attraction model. Remote Sensing of
892 Environment, 210, 345-361.
893 Xu, X., Tong, X., Plaza, A., Li, J., Zhong, Y., Xie, H., & Zhang, L. (2018). A new spectral-spatial sub-pixel mapping model for
894 remotely sensed hyperspectral imagery. IEEE Transactions on Geoscience and Remote Sensing, 56, 6763-6778.
895 Zhang, Y., Atkinson, P.M., Li, X., Ling, F., Wang, Q., & Du, Y. (2017). Learning-based spatial–temporal superresolution mapping of
896 forest cover with modis images. IEEE Transactions on Geoscience and Remote Sensing, 55, 600-614.
897 Zhong, Y., Wu, Y., Xu, X., & Zhang, L. (2015). An adaptive sub-pixel mapping method based on map model and class determination
898 strategy for hyperspectral remote sensing imagery. IEEE Transactions on Geoscience and Remote Sensing, 53, 1411-1426.
899