The effect of the point spread function on sub-pixel mapping

Qunming Wang a, Peter M. Atkinson a,b,c,*

a Lancaster Environment Centre, Lancaster University, Lancaster LA1 4YQ, UK
b Geography and Environment, University of Southampton, Highfield, Southampton SO17 1BJ, UK
c School of Geography, Archaeology and Palaeoecology, Queen's University Belfast, BT7 1NN, Northern Ireland, UK

*Corresponding author
E-mail: pma@lancaster.ac.uk

Abstract: Sub-pixel mapping (SPM) is a process for predicting spatially the land cover classes within mixed pixels. In existing SPM methods, the effect of point spread function (PSF) has seldom been considered. In this paper, a generic SPM method is developed to consider the PSF effect in SPM and, thereby, to increase prediction accuracy. We first demonstrate that the spectral unmixing predictions (i.e., coarse land cover proportions used as input for SPM) are a convolution of not only sub-pixels within the coarse pixel, but also sub-pixels from neighboring coarse pixels. Based on this finding, a new SPM method based on optimization is developed which recognizes the optimal solution as the one that when convolved with the PSF, is the same as the input coarse land cover proportion. Experimental results on three separate datasets show that the SPM accuracy can be increased by considering the PSF effect.

Keywords: Land cover mapping, downscaling, sub-pixel mapping (SPM), super-resolution mapping, point spread function (PSF), Hopfield neural network (HNN).
1. Introduction

Mixed pixels are inevitable in remote sensing images and have brought great challenges in land cover mapping. The spectral unmixing technique has been studied for decades to estimate the proportions of land cover classes within mixed pixels (Bioucas-Dias et al., 2012; Heinz & Chang, 2001; Keshava & Mustard, 2002). The proportions are at the same spatial resolution as the input images and cannot inform the spatial distribution of classes within mixed pixels. To further estimate the spatial distribution of land cover, sub-pixel mapping (SPM) was developed as a post-processing analysis of spectral unmixing outputs. SPM divides mixed pixels to sub-pixels and predicts their class attributes under the coherence constraint from prior spectral unmixing predictions (i.e., coarse land cover class proportions). SPM transforms the conventional pixel-level classification to a finer spatial resolution hard classification (Atkinson, 1997), which can provide more explicit thematic information (e.g., the boundaries between land cover classes can be characterized by more pixels).

In recent decades, various SPM approaches have been developed. As a post-processing step of spectral unmixing, two main groups of SPM approaches can be identified. The first group considers the relation between sub-pixels and solutions are always produced based on defined objectives. Based on the assumption of spatial dependence, the objective can be determined empirically as maximizing the spatial attraction between sub-pixels (Makido & Shortridge, 2007), maximizing the Moran’s I (Makido et al., 2007) or minimizing the perimeter of the area for each class (Villa et al., 2011). Based on prior knowledge, the objective can also be matching prior patterns extracted from training images, such as characterized by the semivariogram (Tatem et al., 2002), two-point histogram (Atkinson, 2008) or landscape structure (Lin et al., 2011). The SPM solutions of this type of methods are achieved based on optimization, including the Hopfield neural network (HNN) (Ling et al., 2010; Muad & Foody, 2012; Nguyen et al., 2011; Tatem et al., 2001), pixel swapping algorithm (PSA) (Atkinson, 2005; Shen et al., 2009; Xu & Huang, 2014), maximum a posteriori method (Zhong et al., 2015), genetic algorithm (Li et al., 2015; Mertens et al., 2003; Tong et al., 2016), and particle swarm optimization (PSO) (Wang et al., 2012). Several iterations are involved for this group of SPM methods and,
thus, a relatively long computing time may be required. The second group of SPM methods considers the relation between sub-pixels and neighboring pixels. The coarse class proportions within each pixel are used directly to characterize the relation between it and sub-pixels and calculate the fine spatial resolution proportions for sub-pixels. Under the coherence constraint, the sub-pixel classes are determined by comparing the fine spatial resolution proportions. As the coarse proportions are fixed for a given pixel, iterations (as in the first method type) are not necessarily involved and SPM solutions can be produced more quickly. Methods falling into this type include sub-pixel/pixel spatial attraction model (SPSAM) (Mahmood et al., 2013; Mertens et al., 2006; Xu et al., 2014), back-propagation neural network-based algorithm (Gu et al., 2008; Zhang et al., 2008), learning-based algorithm (Zhang et al., 2014), kriging (Verhoeve & Wulf, 2002; Boucher & Kyriakidis, 2006) and radial basis function (RBF) (Wang et al., 2014a) interpolation. They can also be summarized as the soft-then-hard SPM (STHSPM) algorithms, a concept proposed in our previous work (Wang et al., 2014b; Chen et al., 2015). In addition, to reduce the uncertainty introduced by spectral unmixing, some SPM methods that do not rely absolutely on coarse proportions were developed, including spatial-spectral methods (Ardila et al., 2011; Kasetkasem et al., 2005; Li et al., 2014; Tolpekin & Stein, 2009), spatial regularization (Ling et al., 2014; Zhang et al., 2015) and contouring methods (Foody & Doan, 2007; Ge et al., 2014; Su et al., 2012).

In remote sensing images, the point spread function (PSF) effect exists ubiquitously. It means that the signal for a given pixel is a weighted combination of contributions from within the pixel and also contributions from neighboring pixels (Townshend et al., 2000; Van der Meer, 2012). The PSF can brighten dark objects and darken bright objects observed from the surface (Huang et al., 2002). It results in a fundamental limit on the amount of information that remote sensing images can contain (Manslow & Nixon, 2002). The PSF is a two-dimensional function accounting for both the across-track and along-track directions (Campagnolo & Montano, 2014; Radoux et al., 2016). The PSF effect is caused mainly by the optics of the instrument, the detector and electronics, atmospheric effects, and image resampling (Huang et al., 2002; Schowengerdt, 1997).

The PSF effect may not be an important issue for homogeneous regions, but it is crucial for heterogeneous landscapes dominated by mixed pixels. To the best of our knowledge, very few SPM methods have considered
the PSF effect in downscaling. For example, in most of the existing SPM methods, the coherence constraint from class proportions is satisfied simply by fixing the number of sub-pixels for each class within a single coarse pixel (i.e., the ideal square wave PSF is considered). The number of sub-pixels to be allocated to a class within a coarse pixel is calculated as the product of the coarse class proportion within the coarse pixel and the square of the zoom factor. Due to the PSF effect, however, the coarse proportions estimated by spectral unmixing are actually a function, in part, of the neighboring coarse pixels. The uncertainty in coarse proportions is propagated to the post-SPM process where the coarse proportions contaminated by neighboring coarse pixels are used as the coherence constraint. There is, therefore, a great need for an approach accounting for the PSF effect in SPM to increase the prediction accuracy.

There are two plausible solutions to cope with the PSF effect in SPM. One is to consider the PSF effect in the pre-spectral unmixing process and to estimate more reliable coarse proportions from observed multispectral images. Based on the more reliable predictions, the PSF need not be considered in SPM (i.e., the ideal square wave PSF can be considered in SPM, as in existing SPM approaches). However, spectral unmixing is an ill-posed inverse problem. It is more complicated when part of the neighboring coarse pixels (i.e., neighboring sub-pixels) are involved as this technique is generally performed at the pixel resolution. Currently, it is challenging to account for the PSF in spectral unmixing and obtain reliable proportions. The alternative solution, considered here, is to model the PSF effect in the SPM process, based on the proportions contaminated by neighboring coarse pixels. This strategy is more feasible as SPM is conducted at the sub-pixel scale and contributions from neighboring sub-pixels in PSF can be straightforwardly modeled.

In this paper, to increase the SPM accuracy, the PSF effect is considered directly in the SPM process. Most SPM methods need to first calculate the number of sub-pixels for each class within each coarse pixel. Based on these fixed numbers, the sub-pixel classes are then predicted. This is not a problem for the ideal square wave PSF, as mentioned earlier. When considering the non-ideal PSF, however, the coarse proportions are a convolution of the sub-pixel class values in a larger local window, rather than the single coarse pixel in the ideal square-wave PSF. In this case, the number of sub-pixels for each class in each coarse pixel cannot be
determined using only the single coarse proportion (i.e., product of the coarse proportion and the square of the zoom factor, as in existing SPM methods), and it actually cannot be calculated explicitly. In this case, SPM methods such as the STHSPM algorithms are not suitable choices. A plausible solution to this issue is to convolve the fine spatial resolution SPM realization with the PSF and compare the estimated proportion with the actual coarse proportion, and use the error to guide further updating of the current realization. The iteration-based HNN is a method of this type. Therefore, in this paper, the HNN is used to reduce the uncertainty in SPM introduced by the PSF effect.

The remainder of this paper is organized into four sections. Section 2 first introduces the mechanism of the PSF effect in SPM and then the details of the proposed strategy for considering the PSF in SPM. The experimental results for three groups of datasets are provided in Section 3 for validation of the proposed method. Section 4 further discusses the proposed SPM method, followed by a conclusion in Section 5.

2. Methods

2.1. The PSF effect in SPM

This section will illustrate the PSF effect in SPM and demonstrate that the coarse proportions in SPM are a convolution of the sub-pixel class values in a local window centered at the coarse pixel. Let \( S_v \) be the spectrum of coarse pixel \( V \), \( R_k \) be the spectrum of class endmember \( k \) \((k=1, 2, ..., K, \) where \( K \) is the number of land cover classes), and \( F_k(V) \) be the proportion of class \( k \) in pixel \( V \). Based on the classical linear spectral mixture model (Bioucas-Dias et al., 2012; Heinz & Chang, 2001; Keshava & Mustard, 2002), the spectrum of each coarse pixel is a linear combination of the spectrum of endmembers, where the weights are determined as the class proportions within the coarse pixel. That is

\[
S_v = \sum_{k=1}^{K} R_k F_k(V). \tag{1}
\]
Due to the PSF effect in remote sensing images, Eq. (2) holds

$$S_v = S_v \ast h_v$$  \hspace{1cm} (2)$$

where $S_v$ is the spectrum of sub-pixel $v$, $h_v$ is the PSF and $\ast$ is the convolution operator. For sub-pixel $v$, its spectrum $S_v$ can be characterized as

$$S_v = \sum_{k=1}^{K} R_k I_k(v)$$  \hspace{1cm} (3)$$

in which $I_k(v)$ is a class indicator as follows

$$I_k(v) = \begin{cases} 1, & \text{if sub-pixel } v \text{ belongs to class } k, \\ 0, & \text{otherwise} \end{cases} \hspace{1cm} (4)$$

By substituting Eq. (3) into Eq. (2), we have

$$S_v = \left[ \sum_{k=1}^{K} R_k I_k(v) \right] \ast h_v = \sum_{k=1}^{K} R_k \left[ I_k(v) \ast h_v \right].$$  \hspace{1cm} (5)$$

The comparison between Eq. (1) and Eq. (5) leads to

$$F_k(V) = I_k(v) \ast h_v.$$

Let $z$ be the zoom factor, that is, each coarse pixel is divided into $z$ by $z$ sub-pixels. As shown in the one dimensional illustration in Fig. 1, when the PSF takes the ideal square wave filter in Eq. (7)

$$h_v(i, j) = \begin{cases} 1, & \text{if } (i, j) \in V(i, j) \\ 0, & \text{otherwise} \end{cases} \hspace{1cm} (7)$$

where $(i, j)$ is the spatial location of the sub-pixel and $V(i, j)$ is the spatial extent of the coarse pixel $V$ containing the sub-pixel at $(i, j)$. Further, the convolution in Eq. (6) can be simplified as

$$F_k(V) = \frac{1}{z^2} \sum_{i,j} I_k(v_{ij}).$$  \hspace{1cm} (8)$$

In Eq. (8), all sub-pixels $v_{ij}$ are located within coarse pixel $V$. This means that based on the assumption of an ideal square wave PSF, the coarse class proportion estimated from the linear spectral mixture model in Eq. (1)
is viewed as the average of all sub-pixel class indicators within the pixel. This is the common case considered in most SPM methods.

![Ideal square PSF](image)

**Fig. 1** A one dimensional illustration of the PSF in SPM. Based on the real PSF, the coarse pixel value (or coarse class proportion) should be a convolution of the sub-pixel values (or sub-pixel class indicators) located within the local window centered at the coarse pixel, rather than only within the single coarse pixel (as assumed in the ideal square wave PSF).

In reality, the PSF is different to the ideal square wave filter. The spatial coverage of $h_v$ is generally larger than a coarse pixel size, but is finite, as shown in Fig. 1. For example, the PSF can be the Gaussian filter, which is used widely in remote sensing (Campagnolo & Montano, 2014; Huang et al., 2002; Manslow & Nixon, 2002; Townshend et al., 2000; Van der Meer, 2012; Wenny et al., 2015)

$$h_v(i, j) = \begin{cases} 
\frac{1}{2\pi\sigma^2} \exp \left[-\frac{(i^2 + j^2)}{2\sigma^2}\right], & \text{if } (i, j) \in V'(i, j) \\
0, & \text{otherwise} 
\end{cases}$$  

(9)

where $\sigma$ is the standard deviation (width of the Gaussian PSF) and $V'(i, j)$ is the spatial extent of the local window centered at coarse pixel $V$ ($V$ covers the sub-pixel at $(i, j)$), see Fig. 1. Thus, any sub-pixel $v$ involved in the calculations in Eq. (2) and Eq. (6) falls within a wider spatial coverage $V'(i, j)$, rather than only within $V(i, j)$. This means that based on the linear spectral mixture model in Eq. (1), the estimated coarse class proportions within each coarse pixel are actually a convolution of the sub-pixels in the local window centered at the coarse pixel, rather than only the sub-pixels within the coarse pixel (i.e., not the case in Eq. (8)).
2.2. Enhancing SPM by considering the PSF effect

In conventional SPM, the number of sub-pixels for each class needs to be determined in advance, which is used as a coherence constraint in the mapping process. Let \( N_k(V) \) be the number of sub-pixels for class \( k \) in pixel \( V \), which is usually calculated as

\[
N_k(V) = \text{round} \left( F_k(V) \times z^2 \right).
\]  

(10)

From Section 2.1, it is concluded that in reality the estimated coarse proportions are contaminated by their neighboring pixels due to the PSF effect (see Eq. (6)). Therefore, it is generally incorrect to use Eq. (10) for the coherence constraint in SPM, unless the correct land cover proportions (i.e., not contaminated by neighbors) can be produced. As can be seen from Eq. (5), even when the endmember set \( \mathbf{R} \) and PSF \( h_v \) are known perfectly, it is not trivial to determine the correct proportions. This is because the coarse proportions can sometimes be functions of neighboring sub-pixels, as can be seen from Eq. (6) and Fig. 1. In this case where sub-pixels are involved, spectral unmixing becomes challenging, as unmixing is carried out at the pixel resolution and cannot account for proportions at the sub-pixel resolution. Townshend et al. (2000) and Huang et al. (2002) proposed an interesting deconvolution method to reduce the impact of the PSF in coarse proportion predictions. This method was developed at the pixel resolution, which quantifies the contributions from neighbors in units of coarse pixels. That is, it treats all sub-pixels in a surrounding coarse pixel equally and assumes a uniform contribution to the center coarse pixel. According to Fig. 1, however, sub-pixels in the surrounding coarse pixel have a different spatial distance to the center coarse pixel and thus, should have different contributions to the center coarse proportion (e.g., some sub-pixels even have no contribution).

In this paper, the coarse proportions contaminated by pixel neighbors are directly considered in SPM. As \( N_k(V) \) cannot be calculated explicitly, conventional SPM methods such as the STHSPM algorithms cannot be used. Guided by the mechanism in Eq. (6), the ideal SPM solution should be the one that when convolved with the PSF, is the same as the coarse proportion. Alternatively, an optimization-based SPM method is employed in this paper, where the convolution of the current SPM realization is compared with the contaminated coarse
proportions for updating SPM predictions iteratively. The HNN is a method suitable for this task and is used for accounting for the PSF in SPM. The HNN has been used widely in SPM, appreciating its highly satisfactory performance (Ling et al., 2010; Muad & Foody, 2012; Nguyen et al., 2011; Tatem et al., 2001). It considers each sub-pixel as a neuron and works by minimizing an energy function composed of a goal and a coherence constraint term. The HNN accounting for the PSF effect is introduced below.

Suppose \((k, i, j)\) is a neuron for sub-pixel at location \((i, j)\) and is on the network layer representing land cover class \(k\), and \(q_{kij}\) is its output. If \(q_{kij} = 1\), it means that the sub-pixel at \((i, j)\) belongs to class \(k\), and if \(q_{kij} = 0\), the sub-pixel does not belong to class \(k\). Using the HNN, the outputs are update iteratively for each sub-pixel. The output \(q_{kij}\) is a function of the input signal \(u_{kij}\)

\[
q_{kij} = \frac{1}{2} \left[ 1 + \tanh(\lambda u_{kij}) \right]
\]

where \(\lambda\) determines the steepness of the function. The input \(u_{kij}\) for the \(t\)-th iteration is updated by

\[
u_{kij}(t + dt) = u_{kij} + \frac{du_{kij}(t)}{dt} dt
\]

in which \(dt\) is a time step. The second term on the right hand side of Eq. (12) is the energy change of the neuron and is described as

\[
\frac{du_{kij}(t)}{dt} = -\frac{dE_{kij}}{dq_{kij}}.
\]

In Eq. (13), \(E\) represents the network energy function and is described as

\[
E = \sum_k \sum_i \sum_j \left( w_1 G_{1kij} + w_2 G_{2kij} + w_3 P_{kij} + w_4 M_{kij} \right)
\]

where \(w_1, w_2, w_3\) and \(w_4\) are four weights, \(G1\) and \(G2\) are two spatial clustering functions characterizing spatial dependence, \(P\) is the proportion constraint and \(M\) is the multi-class constraint (Ling et al., 2010; Nguyen et al., 2011). Accordingly, the energy change of the neuron for sub-pixel at \((i, j)\) in Eq. (13) is determined by

\[
\frac{dE_{kij}}{dq_{kij}} = w_1 \frac{dG_{1kij}}{dq_{kij}} + w_2 \frac{dG_{2kij}}{dq_{kij}} + w_3 \frac{dP_{kij}}{dq_{kij}} + w_4 \frac{dM_{kij}}{dq_{kij}}.
\]
The first two terms of the right hand side of Eq. (15) are calculated as

$$\frac{dG_1\\{kij}}{dq\\{kij}} = \frac{1}{2} \left[ 1 + \tanh \left( \frac{1}{8} \sum_{b=i-1}^{i+1} \sum_{c=j-1}^{j+1} q_{b, c} - q_{kij} \right) - 0.5 \right] \lambda \left( q_{kij} - 1 \right)$$  \hspace{1cm} (16)$$

$$\frac{dG_2\\{kij}}{dq\\{kij}} = \frac{1}{2} \left[ 1 - \tanh \left( \frac{1}{8} \sum_{b=i-1}^{i+1} \sum_{c=j-1}^{j+1} q_{b, c} - q_{kij} \right) - 0.5 \right] \lambda q_{kij}.$$  \hspace{1cm} (17)$$

If the average output of the surrounding neurons is larger than 0.5, the function $G_1$ increases the neuron output to approach 1 to increase the spatial correlation between neighboring sub-pixels. Otherwise, if the average output of the surrounding neurons is less than 0.5, the function $G_2$ decreases the neuron output to 0.

The multi-class constraint means that for any sub-pixel, the sum of neuron outputs for all $K$ classes should be equal to 1, thus

$$\frac{dM\\{kij}}{dq\\{kij}} = \left( \sum_{n=1}^{K} q_{nij} \right) - 1.$$  \hspace{1cm} (18)$$

Considering the PSF effect, the proportion constraint for class $k$ is expressed as

$$\frac{dP\\{kij}}{dq\\{kij}} = L_k(V_{ij}) - F_k(V_{ij})$$  \hspace{1cm} (19)$$

where $V_{ij}$ is the coarse pixel that sub-pixel at $(i, j)$ falls within, $F_k(V_{ij})$ is the target coarse proportion of class $k$ in the pixel (i.e., spectral unmixing predictions), and $L_k(V_{ij})$ is the coarse proportion estimated as convolution of the current SPM realization. Based on the mechanism in Eq. (6), $L_k(V_{ij})$ is calculated as

$$L_k(V_{ij}) = \frac{1}{2} \left[ 1 + \tanh \left( q_k - 0.5 \right) \lambda \right] * h_{ij}$$

$$= \int_{(x,y)\in V'(i,j)} \frac{1}{2} \left[ 1 + \tanh \left( q_k - 0.5 \right) \lambda \right] h_{ij}(x - i_0, y - j_0) \, dx \, dy \hspace{1cm} (20)$$

in which $V'(i, j)$ is defined in the same way as that in Eq. (9) and $(i_0, j_0)$ is the spatial location of the center of coarse pixel $V_{ij}$. The $tanh$ function is used in Eq. (20) to ensure that if the neuron output $q_k$ is above 0.5, it is counted as an output of 1 (i.e., belongs to class $k$) for the coarse proportion calculation; If $q_k$ is below 0.5, it is
not counted for class $k$ instead. This ensures that the neuron output exceeds 0.5 in order to be counted within the calculations for each class.

In Eq. (19), the target coarse proportion $F_k(V_j)$ is used to examine the current SPM realization. The convolution of the final SPM prediction should be the same as the target coarse proportion. Specifically, if the convolution of the current SPM realization, $L_k(V_j)$, is larger than the target proportion $F_k(V_j)$, a positive gradient is produced and the energy is reduced to decrease the neuron output (see Eq. (12) and Eq. (13)) to cope with this overestimation issue. On the contrary, if $L_k(V_j)$ is smaller than $F_k(V_j)$, a negative gradient is produced and the neuron output is increased correspondingly to cope with underestimation. After a number of iterations (usually over 1000), $L_k(V_j)$ will approach $F_k(V_j)$, suggesting that the convolution of the final prediction is almost identical to the target coarse proportion contaminated by the PSF effect. In this way, the PSF effect in SPM is considered and a more reliable sub-pixel class distribution can be reproduced. It should be noted that the proposed HNN-based SPM method accounting for the PSF is suitable for any PSF. Its implementation is not affected by the specific form of PSF. Once the PSF of the sensor is available or estimated, it can be used readily in the proposed method, as shown in Eq. (20).

As suggested in Tatem et al. (2001), for the HNN, the network will converge to similar energy minima given any initialization and the network will converge to an accurate prediction in fewer iterations if a proportion-constrained initialization is used. This scheme was employed in the proposed method for faster convergence of the optimization. According to Eq. (6), in the estimated coarse proportions images, if a pixel is presented as a pure pixel (e.g., belongs entirely to class $k$), all sub-pixels in the local window centered at the coarse pixel must belong to class $k$. Thus, for any pure pixel in the coarse proportion image for class $k$, we directly initialize the neuron outputs of all sub-pixels within it as 1 for class $k$ and 0 for other classes. These values are fixed and not changed in the optimization process. For mixed pixels, random initialization (between 0 and 1) was applied.
3. Experiments

Experiments were carried out on three groups of synthetic datasets, including a dataset with six different targets, a land cover map with multiple classes, and a multispectral image. They were used to test the performance of SPM methods on different typical shapes in Section 3.1, multiple classes in Section 3.2 and the case involving realistic unmixing in Section 3.3. As mentioned earlier, the proposed method is suitable for any PSF. The estimation of the PSF of sensors is still an open problem, and uncertainty unavoidably exists in the process. It is beyond the scope of this paper to estimate the PSF. Alternatively, to concentrate solely on the performance of SPM, all coarse data for the three groups of datasets were produced by convolving the available fine spatial resolution data, using a Gaussian PSF shown in Eq. (9), which is a widely used PSF (Campagnolo & Montano, 2014; Huang et al., 2002; Manslow & Nixon, 2002; Townshend et al., 2000; Van der Meer, 2012; Wenny et al., 2015). The width of the PSF was set to half of the coarse pixel size. Based on this strategy, the fine spatial resolution data are known perfectly and can be used as a reference for evaluation. All SPM results were evaluated both visually and quantitatively. For quantitative evaluation, the classification accuracy of each class and the overall accuracy (OA) in terms of the percentage of correctly classified pixels were used. In addition, to emphasize the improvement of the proposed method over the other methods, an index called the reduction in remaining error (RRE) (Wang et al., 2015) was also used, as calculated below

\[
RRE = \frac{RE_1 - RE_0}{RE_1} \times 100\%
\]  

(21)

where \(RE_1\) and \(RE_0\) are the remaining errors of OA (i.e., 100%-OA) for the benchmark method and the proposed method, respectively.

3.1. Experiment on the six targets

In this experiment, six binary images (pixel value 1 for white target while 0 for black background) with different targets were used for validation. They represent typical, basic shapes in nature. All six images have a
spatial size of 56 by 56 pixels and were marked as T1-T6, as shown in Fig. 1(a). Each image was degraded with a Gaussian PSF and a zoom factor of 7, generating coarse images with 8 by 8 pixels, as shown in Fig. 1(b). The coarse images are exactly the simulated coarse proportions of the white targets in each image.

Fig. 3 shows the relation between the PSF-convolved coarse proportions and actual proportions for the six white targets. The actual proportions were generated by averaging the sub-pixel class indicators within each coarse pixel. As observed from the plots, the proportions are very different to the actual values due to the PSF effect. For example, in all six plots, many proportions should be 0 (i.e., pure background pixel and no target pixel in the coarse pixel), but were actually calculated as values larger than 0 (i.e., mixed pixel). The reason is that although no target pixels exist in a coarse pixel, they may be located in neighboring pixels. According to Eq. (6), these neighboring pixels will contaminate the center pixel which will then not be recognized as a pure background pixel. The accuracies of the coarse proportions were quantified in terms of root mean square error (RMSE) and correlation coefficient (CC), as listed in Table 1. The uncertainty in the PSF-convolved coarse proportions motivates the development of SPM methods considering the PSF effect.

<table>
<thead>
<tr>
<th></th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
<th>T5</th>
<th>T6</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>0.0421</td>
<td>0.0602</td>
<td>0.0671</td>
<td>0.0707</td>
<td>0.0924</td>
<td>0.0972</td>
</tr>
<tr>
<td>CC</td>
<td>0.9949</td>
<td>0.9902</td>
<td>0.9820</td>
<td>0.9868</td>
<td>0.9677</td>
<td>0.9668</td>
</tr>
</tbody>
</table>

The SPM results of the HNN without PSF (that is, with ideal PSF) and with the Gaussian PSF are shown in Fig. 2(c) and Fig. 2(d), respectively. It is seen clearly that by considering the PSF effect, more accurate SPM results can be produced. In the results for target T1-T3, the proposed method produces more compact shapes that are closer to the reference. For T4 and T5, the proposed method produces a straighter line than the HNN without PSF. In T6, the more complex boundaries of the target are reproduced by the proposed method.
Fig. 2 SPM results of six targets. (a) Reference images (56 by 56 pixels). (b) Coarse images (8 by 8 pixels) produced by degrading (a) with a Gaussian PSF and a factor of 7. (c) SPM results produced using the HNN without PSF (or with ideal PSF). (d) SPM results produced using the proposed HNN with the Gaussian PSF. Lines 1-6 are results for T1-T6.
Fig. 3 Relation between the actual proportions (for white target) and proportions (for white target) contaminated due to the PSF effect for the six binary images (z=7). From left to right are results for T1-T6.

Quantitative assessment in terms of OA (accuracy statistics for both target and background in each image) is listed in Table 2. As mentioned earlier, for pure pixels, all sub-pixels within them are simply assigned to the same class to which the pure pixel belongs. As suggested by the existing literature (Mertens et al., 2003), this copy process will only increase the SPM accuracy statistics without providing any useful information on the actual performance of the SPM methods. Hence, for the synthetic coarse images, we did not consider the pure pixels in the accuracy statistics. The quantitative evaluation was conducted for two different zoom factors, z=4 and 7. As observed from the table, the HNN with PSF produces a larger OA than the HNN without PSF. Moreover, the advantage tends to be more obvious when a larger zoom factor is involved. More precisely, for T2 and T5, using the proposed method, the accuracy gains are below 1% for z=4, but increase to be above 2% for z=7. For T6, the proposed method considering the PSF increases the OA by 1.4% for z=4 and 4.7% for z=7. In addition, all RREs are generally large (over 10%) and some even reach over 60%, suggesting that the accuracy increase is obvious.

Table 2 OA (%) of HNN-based SPM with two different schemes for the six binary images

<table>
<thead>
<tr>
<th></th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
<th>T5</th>
<th>T6</th>
</tr>
</thead>
<tbody>
<tr>
<td>z=4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no PSF</td>
<td>99.34</td>
<td>98.72</td>
<td>98.00</td>
<td>98.44</td>
<td>98.30</td>
<td>96.66</td>
</tr>
<tr>
<td>PSF</td>
<td>99.34</td>
<td>99.52</td>
<td>99.26</td>
<td>99.19</td>
<td>98.67</td>
<td>98.08</td>
</tr>
<tr>
<td>RRE</td>
<td>0%</td>
<td>62.50%</td>
<td>63.00%</td>
<td>48.08%</td>
<td>21.76%</td>
<td>42.51%</td>
</tr>
<tr>
<td>z=7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no PSF</td>
<td>99.41</td>
<td>97.31</td>
<td>97.89</td>
<td>97.34</td>
<td>94.48</td>
<td>90.53</td>
</tr>
<tr>
<td>PSF</td>
<td>99.46</td>
<td>99.26</td>
<td>98.16</td>
<td>97.91</td>
<td>97.51</td>
<td>95.18</td>
</tr>
<tr>
<td>RRE</td>
<td>8.47%</td>
<td>72.49%</td>
<td>12.80%</td>
<td>21.43%</td>
<td>54.89%</td>
<td>49.10%</td>
</tr>
</tbody>
</table>
Fig. 4 The accuracy of the two SPM methods in relation to the width of the PSF (in units of coarse pixel, which was produced by degrading Fig. 2(a) with a factor of 7). From left to right are results for T1, T2 and T5, respectively.

The influence of the width of the Gaussian PSF was also analyzed, as shown in Fig. 4. The fine spatial resolution T1, T2 and T5 images were degraded with a factor of 7. Three PSF sizes, 0.5, 0.75 and 1 coarse pixels, were considered for each target. Three observations can be made from the results in Fig. 4. First, for all PSF sizes, the consideration of the PSF effect can lead to consistently larger OA than that produced without the PSF. Second, the accuracy of SPM methods decreases as the width of the PSF increases, no matter whether the PSF is accounted for or not. This is because more neighboring sub-pixels are involved in the convolution process and the SPM problem becomes more complex as the width increases. Third, as the width increases, the advantage of considering the PSF is more obvious.

3.2. Experiment on the land cover map with multiple classes

A land cover map produced from an aerial photograph covering an area in Bath, UK was used in this experiment, as shown in Fig. 5. The image has a spatial resolution of 0.6 m, and a spatial size of 360 by 360 pixels. Four classes were identified in the land cover map, including roads, trees, buildings and grass. The map was degraded with a factor of 8 and a Gaussian PSF, generating four proportion images at a spatial resolution of 4.8 m. The relation between the simulated coarse proportions and actual proportions is shown in the plots in Fig. 6. Table 3 lists the accuracies of the proportions. Similarly to the observation in Fig. 3 and Table 1, the two types of proportions are very different due to the PSF effect.
Fig. 5 The land cover map. (a) Original aerial photograph. (b) Reference map drawn manually from (a).

Fig. 6 Relation between the actual proportions and proportions contaminated due to the PSF effect for the land cover map (z=8). From left to right are results for roads, trees, buildings and grass.

Table 3 Accuracy of the coarse proportion images for the land cover map

<table>
<thead>
<tr>
<th></th>
<th>Roads</th>
<th>Trees</th>
<th>Buildings</th>
<th>Grass</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>0.0440</td>
<td>0.0576</td>
<td>0.0591</td>
<td>0.0924</td>
<td>0.0633</td>
</tr>
<tr>
<td>CC</td>
<td>0.9866</td>
<td>0.9867</td>
<td>0.9844</td>
<td>0.9792</td>
<td>0.9842</td>
</tr>
</tbody>
</table>

The proposed method is not only compared to the original HNN method, but also to several typical SPM methods. As mentioned in the Introduction, there are mainly two families of SPM methods for post-processing of spectral unmixing. One is the method considering the relation between sub-pixels and neighboring pixels, while the other considers the relation between sub-pixels. The SPSAM (Mertens et al., 2006) and RBF (Wang et al., 2014a) methods were selected for the former and PSA (Makido et al., 2007) was selected for the latter.
For these three benchmark methods, the number of sub-pixels for each class needs to be determined first. The numbers were calculated according to Eq. (10), where uncertainty from the PSF effect exists in the coarse proportions. The SPM results of all five methods are presented in Fig. 7.

Fig. 7 SPM results of the land cover map with 360 by 360 pixels (z=8). (a) SPSAM (no PSF). (b) RBF (no PSF). (c) PSA (no PSF). (d) HNN (no PSF). (e) The proposed HNN with PSF. (f) Reference.

Due to the uncertainty in the coarse proportions, the SPM predictions of SPSAM and RBF contain a number of jagged artifacts and the boundaries of classes are rough. Although PSA is able to enhance the performance and produce a more compact result, the class boundaries are still not smooth and some noisy pixels exist. Different from the three methods, the original HNN method produces a cleaner result and the class boundaries
are smoother, although the PSF is not considered. The reason is that the HNN is not absolutely slavish to the coarse proportions, and can remove isolated pixels and produce a compact result by the two spatial clustering functions $G_1$ and $G_2$ (see Eq. (16) and Eq. (17)). However, there are some needle-like artifacts for restoration of roads and trees in the HNN result. Using the proposed method accounting for the PSF effect, the result is more accurate than the original HNN without the PSF method. For example, the prediction of the roads by the proposed method is smoother, and needle-like artifacts are removed. The result of the proposed method is the closest to the reference in Fig. 7(f) amongst all five methods.

Table 4 SPM accuracy (%) of the different SPM schemes for the experiment on the land cover map

<table>
<thead>
<tr>
<th></th>
<th>SPSAM</th>
<th>RBF</th>
<th>PSA</th>
<th>HNN</th>
<th>HNN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(no PSF)</td>
<td>(no PSF)</td>
<td>(no PSF)</td>
<td>(no PSF)</td>
<td>(PSF)</td>
</tr>
<tr>
<td>Roads</td>
<td>94.59</td>
<td>94.64</td>
<td>93.81</td>
<td>98.92</td>
<td>99.78</td>
</tr>
<tr>
<td>Trees</td>
<td>92.16</td>
<td>92.53</td>
<td>91.45</td>
<td>94.38</td>
<td>95.41</td>
</tr>
<tr>
<td>Buildings</td>
<td>91.20</td>
<td>91.69</td>
<td>90.43</td>
<td>93.50</td>
<td>95.28</td>
</tr>
<tr>
<td>Grass</td>
<td>93.82</td>
<td>94.10</td>
<td>93.22</td>
<td>97.86</td>
<td>98.57</td>
</tr>
<tr>
<td>OA</td>
<td>93.17</td>
<td>93.47</td>
<td>92.50</td>
<td>96.62</td>
<td>97.58</td>
</tr>
<tr>
<td>RRE (in OA)</td>
<td>64.57%</td>
<td>62.94%</td>
<td>67.73%</td>
<td>28.40%</td>
<td></td>
</tr>
<tr>
<td>Roads</td>
<td>88.24</td>
<td>88.86</td>
<td>87.26</td>
<td>95.13</td>
<td>98.03</td>
</tr>
<tr>
<td>Trees</td>
<td>86.20</td>
<td>87.48</td>
<td>86.13</td>
<td>85.20</td>
<td>90.08</td>
</tr>
<tr>
<td>Buildings</td>
<td>82.51</td>
<td>83.81</td>
<td>81.83</td>
<td>82.92</td>
<td>88.64</td>
</tr>
<tr>
<td>Grass</td>
<td>91.77</td>
<td>90.98</td>
<td>91.74</td>
<td>95.97</td>
<td>95.85</td>
</tr>
<tr>
<td>OA</td>
<td>89.42</td>
<td>89.15</td>
<td>89.23</td>
<td>92.15</td>
<td>94.00</td>
</tr>
<tr>
<td>RRE (in OA)</td>
<td>43.29%</td>
<td>44.70%</td>
<td>44.29%</td>
<td>23.57%</td>
<td></td>
</tr>
</tbody>
</table>

Table 4 is the accuracy for the five SPM methods, where two zoom factors, 4 and 8, were tested. Again, the pure pixels were not considered in the accuracy statistics. For both zoom factors, the proposed HNN with PSF method produces the greatest accuracy for all classes, and thus, the largest OA amongst all methods. More specifically, SPSAM, RBF and PSA have very similar accuracies. The three methods produce OA of around 93% for $z=4$ and 89% for $z=8$. Compared to the three benchmark methods, the HNN without PSF method is more accurate. For $z=4$ and 8, the OA is increased by around 3.5% and 3%, respectively. With respect to the proposed HNN with PSF method, the accuracies of roads, trees, buildings and grass are increased by at least
5%, 3%, 3.5% and 4% in comparison with SPSAM, RBF and PSA for $z=4$, and correspondingly, the OA is increased by over 4%. In addition, with the PSF, the OA is 1% larger than the original HNN method. For $z=8$, the accuracies of roads, trees and buildings for the proposed method are 3%, 5% and 6% larger than for the HNN without PSF method. Moreover, the OA of the proposed method is 4.6%, 4.9%, 4.8% and 1.9% larger than the SPSAM, RBF, PSA and HNN without PSF methods, and the corresponding RREs are 43.29%, 44.70%, 44.29% and 23.57%.

3.3. Experiment on the multispectral image

In this experiment, to control the experimental analysis and ensure the perfect reliability of the reference, a synthesized multispectral image was used. The original multispectral image was acquired by the Landsat-7 Enhanced Thematic Mapper sensor in August 2001 and covers a farmland in the Liaoning Province, China. The spatial resolution is 30 m. The studied area has a spatial size of 240 by 240 pixels and covers mainly four land cover classes (marked as C1-C4). Fig. 8(a) and Fig. 8(b) show the original multispectral image and the corresponding manually digitized reference map. From the reference map, the mean and variance of each land cover class in the original 30 m Landsat image (bands 1-5 and 7 were considered) were calculated. Referring to Fig. 8(b), the six-band 30 m resolution multispectral image was synthesized based on the random normal distribution and the mean and variance of the classes. Fig. 8(c) shows the synthesized multispectral image. A 240 m coarse image, a spatial resolution comparable to that of medium-spatial-resolution systems such as Moderate Resolution Imaging Spectroradiometer (MODIS), was created by degrading the synthesized 30 m image with a factor of 8, using a Gaussian PSF. Fig. 8(d) shows the 240 m image. With this strategy, the reference land cover map at 30 m is known perfectly for accuracy assessment.

Spectral unmixing was first performed on the 240 m coarse image using the classical linear spectral mixture model. Fig. 9 is the relation between the estimated proportions and actual proportions for the four classes. The errors are visually larger than those in Fig. 6. This is because the errors in proportions in this experiment originate not only from the PSF effect, but also from the spectral unmixing process. This can be supported by
the results in Table 5, where the accuracies of the proportions were evaluated using two different references. The convolved proportions were produced by convolving the 30 m land cover map in Fig. 8(b) with the Gaussian PSF. Using the convolved proportions as reference, errors can be observed from only spectral unmixing. As shown in the table, the spectral unmixing errors are particularly large for C1 and C2 (with RMSEs of 0.0418 and 0.0303, respectively). Furthermore, as the PSF introduces additional uncertainty, the accuracies evaluated using the actual proportions as reference are smaller than those for the convolved proportions.

![Image](image.png)

Fig. 8 The multispectral image used for experiment 3. (a) Original 30 m multispectral image (bands 432 as RGB). (b) Reference map drawn manually from (a). (c) Synthesized 30 m multispectral image (bands 432 as RGB). (d) 240 m coarse image produced by degrading (c) with a factor of 8 using a Gaussian PSF.

![Image](image.png)

Fig. 9 Relation between the actual proportions and proportions estimated from spectral unmixing for the multispectral image (z=8). From left to right are results for C1-C4.

Based on the estimated proportions, SPM was implemented and the results of the five methods are shown in Fig. 10. Failing to account for the PSF, the SPSAM, RBF and PSA predictions are dominated by noisy pixels.
and spurs on boundaries between classes. As the original HNN method is not completely slavish to the coarse proportions, it can remove the noisy pixels and spurs and produce a smoother result than the SPSAM, RBF and PSA. In Fig. 10(d), however, there are jagged boundaries, such as for C3 and C4. The prediction is further enhanced by considering the PSF effect in the proposed method. The quantitative assessment in Table 6 suggests that the proposed method produces a larger accuracy for all four classes as well as a larger OA than the other four SPM methods. Note that in this experiment, all pixels were considered in accuracy statistics, including pure pixels. This is because whether a pixel is pure or not is determined by spectral unmixing and we are concerned about the performance of spectral unmixing. The OA of the proposed method is 91.86%, with a gain of more than 5% over SPSAM, RBF and PSA. Compared with the HNN without PSF, the proposed method increases the accuracies of C1 and C4 by 4.5% and 5.5% and the OA by 1.3%. All RREs (in OA) are large, which means that the proposed method reduces the errors obviously. Note that due to the uncertainty in spectral unmixing, the OA of the proposed method is smaller than that for the land cover map in Table 2 (z=8).

Table 5 Accuracy of the coarse proportion images for the multispectral image

<table>
<thead>
<tr>
<th>Actual proportion</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C4</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>0.0873</td>
<td>0.0950</td>
<td>0.0517</td>
<td>0.0529</td>
<td>0.0717</td>
</tr>
<tr>
<td>CC</td>
<td>0.9703</td>
<td>0.9766</td>
<td>0.9844</td>
<td>0.9692</td>
<td>0.9751</td>
</tr>
</tbody>
</table>

Convolved proportion

| RMSE              | 0.0418| 0.0303| 0.0085| 0.0206| 0.0253|
| CC                | 0.9885| 0.9957| 0.9992| 0.9913| 0.9937|
Fig. 10 SPM results of the multispectral image with 240 by 240 pixels (z=8). (a) SPSAM (no PSF). (b) RBF (no PSF). (c) PSA (no PSF). (d) HNN (no PSF). (e) The proposed HNN with PSF. (f) Reference.

Table 6 SPM accuracy (%) of the different SPM schemes for the experiment on the multispectral image

<table>
<thead>
<tr>
<th></th>
<th>SPSAM (no PSF)</th>
<th>RBF (no PSF)</th>
<th>PSA (no PSF)</th>
<th>HNN (no PSF)</th>
<th>HNN (PSF)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>75.88</td>
<td>77.65</td>
<td>75.98</td>
<td>76.89</td>
<td>81.40</td>
</tr>
<tr>
<td>C2</td>
<td>90.92</td>
<td>91.71</td>
<td>90.92</td>
<td>96.57</td>
<td>96.26</td>
</tr>
<tr>
<td>C3</td>
<td>82.50</td>
<td>84.30</td>
<td>81.68</td>
<td>87.08</td>
<td>89.81</td>
</tr>
<tr>
<td>C4</td>
<td>71.18</td>
<td>73.38</td>
<td>72.95</td>
<td>77.75</td>
<td>83.15</td>
</tr>
<tr>
<td>OA</td>
<td>85.91</td>
<td>87.07</td>
<td>85.95</td>
<td>90.57</td>
<td>91.86</td>
</tr>
<tr>
<td>RRE (in OA)</td>
<td>42.23%</td>
<td>37.05%</td>
<td>42.06%</td>
<td>13.68%</td>
<td></td>
</tr>
</tbody>
</table>

4. Discussion

The PSF effect exists ubiquitously in remote sensing images and the pixel signal is contaminated by its neighbors as a result. Existing SPM methods are generally performed based on the assumption of the ideal square wave filter-based PSF and treat the signal of a coarse pixel as the mean of sub-pixel signals within it. This assumption ignores the effect from neighbors and limits unnecessarily the accuracy of SPM predictions. As shown in the experimental results, the SPSAM, RBF and PSA predictions contain noticeable errors where the PSF is ignored. Note that the performances of the SPSAM, RBF and PSA methods are not as satisfactory as
demonstrated in the literature where they were proposed (Makido et al., 2007; Mertens et al., 2006; Wang et al., 2014a). The reason is that in the experimental design in the literature, the coarse images were simulated using the ideal square PSF and, thus, the sub-pixel maps can be satisfactorily reproduced by ignoring the PSF effect.

The proposed SPM approach that considers the PSF effect is developed based on the iterative HNN method. The computing time is the product of the average time consumed in each iteration and the number of required iterations. When considering the PSF effect, more sub-pixels are involved in the convolution process (see Eq. (20)) and more time is needed in each iteration. Moreover, the convergence rate is slower and a larger number of iterations is required. Thus, the proposed method has larger computational cost than the original HNN method without the PSF. In the experiments, the original HNN method is able to converge after 1000 iterations, but the proposed method needs 3000 iterations instead. For SPM of the multispectral image, the running time of the original HNN and the proposed method is 1 and 5 hours, respectively. Note that the computational cost of a SPM method is positively related to the spatial size of the input coarse image, number of classes and zoom factor.

This paper considers using HNN to cope with the PSF effect. The solution is identified as modifying the proportion constraint term in the original HNN using a convolution process where contributions from neighboring coarse pixels are accounted for. It would be worthwhile to develop other solutions based on existing SPM methods, such as PSA and the STHSPM methods. For example, in the original PSA method, pixel swapping is allowed only within a coarse pixel. To account for the PSF, the SPM method can be extended by allowing sub-pixel swaps not only within a coarse pixel, but also between coarse pixels. The number of allowable sub-pixel swap between coarse pixels would be a key issue that needs to be addressed. Moreover, it would also be interesting to consider developing artificial intelligence-based methods (e.g., genetic algorithm and PSO, etc.). In such methods, the solutions in the set (called population in genetic algorithm and swarm in PSO) can be updated iteratively according to the fitness values (calculated according to the pre-defined objective) by using related operators (such as crossover and mutation in genetic algorithm and position update in PSO). An optimal solution is expected to be produced after a number of generations. A critical issue for this
type of approach would be the definition of the objective, where a constraint term is needed for taking the PSF effect into consideration. Both the prediction accuracy and computational cost will be important indices to identifying an effective method.

Each sensor has its own PSF size. Based on the assumption of a Gaussian filter, some studies were conducted to estimate the PSF size of satellite sensor images. Radoux et al. (2016) estimated the PSF size of Landsat 8 and Sentinel-2 images. It was found that the full-width at half-maximum (FWHM) for Landsat 8 red band is 1.70 pixels (amounting to a width of 0.72 pixel) and ranges from 1.67 to 2.21 pixels (i.e., a width from 0.71 to 0.94 pixel) for Sentinel-2 bands. Campagnolo & Montano (2014) estimated the PSF size of MODIS images at different view zenith angles. The size of the PSF is a critical factor affecting the performance of the proposed method, as displayed in Fig. 4. As the size increases, more neighboring sub-pixels contaminate the center coarse pixel, and the accuracy of the proposed method will decrease correspondingly. However, the advantage of the proposed method over the traditional method (i.e., accuracy gain) is obvious across different PSF sizes, especially for large ones. The encouraging performance will promote the application of the proposed method for sensors with various PSF sizes.

The main objective of this paper was to find a generic solution to address the PSF effect in SPM and increase the accuracy of SPM predictions. A Gaussian PSF was assumed for convenience in the experimental validation. In reality, the PSF may not be the Gaussian filter, especially for sensors with a rotating/scanning mirror which will ensure that the shape has a directional component. For example, the MODIS sensor PSF was claimed to be triangular in the along-scan direction but rectangular in the along-track direction in Tan et al. (2006). As seen from Eq. (20), however, the proposed method is suitable for any PSF. Thus, in real cases, if the PSF is known or estimated reliably, it can be used readily in the proposed method. This paper provides a first guidance to increase SPM accuracy by considering the PSF effect. It could also help to initiate research on PSF estimation that would ultimately provide the information needed for the use of the proposed method. The PSF estimation mainly includes determination of the specific form of the function and related parameters. This is part of our ongoing research.
The spatial clustering functions $G_1$ and $G_2$ in the HNN make the proposed method more appropriate for SPM in the H-resolution case, where the objects of interest are larger than the pixel size of the input coarse image (Atkinson, 2009). In the L-resolution case where the objects of interest are smaller than the pixel size, alternative prior spatial structure information-based models, such as the semivariogram-based model in Tatem et al. (2002), are more suitable, and should be applied. It would also be worthy of developing models that can adaptively cope with different spatial patterns. Ge et al. (2016) performed a pioneering study for this issue.

Although the proposed method can increase the SPM accuracy, uncertainty still exists. Based on Eq. (6), the ideal SPM solution is identified as the one that when convolved with the PSF, predicts the coarse proportion. There can be multiple solutions satisfying this condition, leading to the perfect coherence constraint. The uncertainty related to this issue can be further reduced by using additional information, such as sub-pixel shifted remote sensing images (Ling et al., 2010; Zhong et al., 2014), panchromatic images (Ardila et al., 2011; Li et al., 2014; Nguyen et al., 2011), high resolution color images (Mahmood et al., 2013), digital elevation models (Huang et al., 2014), time-series images (Wang et al., 2016), segmentation data (Aplin & Atkinson, 2001; Robin et al., 2008) or shape information (Ling et al., 2012). It would be an interesting challenge to design the appropriate model to incorporate such additional information into the proposed method considering the PSF effect for possible enhancement in future research.

5. Conclusion

This paper presents a HNN-based method to account for the PSF effect in SPM and increase the SPM predictions. Based on the recognition of the PSF as a real effect, the coarse proportions are viewed as the convolution of sub-pixels within the local window centered at the coarse pixel, rather than of only the sub-pixels within the coarse pixel. In the proposed HNN-based method, the interim SPM realization is convolved with the PSF and compared with coarse proportions for further updating. The final solution is
identified as the one that when convolved with the PSF, is the same as the input coarse proportion. The proposed method is a generic method suitable for any PSF. The effectiveness of the proposed method was validated using three groups of datasets.

Acknowledgment

This work was supported in part by the Research Grants Council of Hong Kong under Grant PolyU 15223015. The authors would also like to thank the handling editor and three anonymous reviewers for their valuable comments which greatly improved the work.

References


