

## International Journal of Remote Sensing

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/tres20>

### A super-resolution mapping method using local indicator variograms

Huiran Jin<sup>a</sup>, Giorgos Mountrakis<sup>a</sup> & Peijun Li<sup>b</sup>

<sup>a</sup> Department of Environmental Resources Engineering, State University of New York College of Environmental Science and Forestry, Syracuse, NY, 13210, USA

<sup>b</sup> Institute of Remote Sensing and GIS, Peking University, Beijing, 100871, PR China

Version of record first published: 16 Jul 2012

To cite this article: Huiran Jin, Giorgos Mountrakis & Peijun Li (2012): A super-resolution mapping method using local indicator variograms, *International Journal of Remote Sensing*, 33:24, 7747-7773

To link to this article: <http://dx.doi.org/10.1080/01431161.2012.702234>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.tandfonline.com/page/terms-and-conditions>

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

## A super-resolution mapping method using local indicator variograms

HUIRAN JIN\*<sup>†</sup>, GIORGOS MOUNTRAKIS<sup>†</sup> and PEIJUN LI<sup>‡</sup>

<sup>†</sup>Department of Environmental Resources Engineering, State University of New York  
College of Environmental Science and Forestry, Syracuse, NY 13210, USA

<sup>‡</sup>Institute of Remote Sensing and GIS, Peking University, Beijing 100871, PR China

(Received 26 July 2011; in final form 18 October 2011)

Super-resolution mapping (SRM) is a recently developed research task in the field of remotely sensed information processing. It provides the ability to obtain land-cover maps at a finer scale using relatively low-resolution images. Existing algorithms based on indicator geostatistics and downscaling cokriging offer an SRM approach using spatial structure models derived from real data. In this article, a novel SRM method is developed based on a sequentially produced with local indicator variogram (SLIV) SRM model. In the SLIV method, indicator variograms extracted from target-resolution classification are produced from a representative local area as opposed to using the entire image. This simplifies the application of the method since limited target-resolution reference data are required. Our investigation on three diverse case studies shows that the local window (approximately 2% of the entire study area) selection process offers comparable accuracy results to those using globally derived spatial structures, indicating our methodology to be a promising practice. Furthermore, comparison of the proposed method with random realizations indicates an improvement of 7–12% in terms of overall accuracy and 15–18% in terms of the kappa coefficient. The evaluation focused on a 270–30 m pixel size reconstruction as a potential popular application, for example moving from Moderate Resolution Imaging Spectroradiometer (MODIS) to Landsat-type resolutions.

### 1. Introduction

Land-cover and land-use (LCLU) information is fundamental to multiple scientific research fields and applications. Remote sensing provides a major source of LCLU data, and usually image classification is the underlying process for generating LCLU thematic maps. Traditionally, there are two general types of image classification for medium- and low-resolution data: pixel-based and subpixel classifiers. Pixel-based classifiers assign each pixel to a single class (Campbell 1996), whereas in reality many pixels in remotely sensed images represent more than one class on the ground (Fisher 1997, Schowengerdt 1997). The common existence of mixed pixels leads to ambiguities in pixel-based classification. Subpixel classifiers (e.g. linear mixture models) predict the proportions of each individual class within a pixel through spectral unmixing (Richards and Jia 1999, Tso and Mather 2001). A variety of spectral unmixing approaches have been developed to place more confidence on the accuracy of the

---

\*Corresponding author. Email: [hjin02@syr.edu](mailto:hjin02@syr.edu)

estimated class proportions (e.g. Foody 1996, 2001, Bateson *et al.* 2000, Haertel *et al.* 2004, Small 2004, Foody and Doan 2007). All provide substantial additional information when compared to pixel-based classifiers; however, they still cannot overcome an inherent constraint, namely that the internal spatial distribution of each land-cover class within a given pixel is still unknown. For example, a pixel may contain 15% of water, but subpixel algorithms offer no location information on how that 15% of water is distributed spatially within that pixel.

Super-resolution mapping (SRM) focuses on prediction of the location of land-cover classes within a pixel from coarse fraction images produced by subpixel classification (Atkinson 1997). It can be considered as a method combining the benefits of both types of image classifiers, and transforming a subpixel classification into a pixel-based classification with a finer resolution (Mertens *et al.* 2006). During this process, every coarse pixel is divided into a number of subpixels according to the scale ratio between two resolution levels. Each subpixel is then assigned to a specific class, subject to the constraint that the class proportions when degrading the resulting map to the coarse scale are consistent with the corresponding class proportions generated on the original coarse-resolution image through spectral unmixing (Atkinson 1997).

As a recently developed research issue, SRM provides the possibility to obtain land-cover information at a finer scale using relatively low-resolution images. Since high spatial resolution images are expensive, difficult to acquire and process consistently/efficiently, especially over large study areas, this reconstruction process can be alternatively used to predict detailed land-cover distribution from images with a lower resolution. Another significant advantage is that typically coarser resolution sensors have a higher revisiting frequency than finer ones (e.g. Moderate Resolution Imaging Spectroradiometer (MODIS) vs Landsat).

Ever since SRM was formally introduced by Atkinson (1997), researchers have proposed several algorithms for the reconstruction process and investigated associated errors. Examples of these methods include subpixel per-field classification (Aplin and Atkinson 2001), linear optimization techniques (Verhoeve and De Wulf 2002), inverse-distance weighting interpolation (Zhan *et al.* 2002), Hopfield neural network optimization (Tatem *et al.* 2001, 2002, 2003), genetic algorithms (Mertens *et al.* 2003), artificial neural network and wavelet transformation (Mertens *et al.* 2004), two-point histogram optimization (Atkinson 2004), pixel-swapping optimization (Atkinson 2001a, 2005, Makido and Shortridge 2005, 2007, Thornton *et al.* 2006), Markov random field model-based approaches (Tso and Mather 2001, Kasetkasem *et al.* 2005) and subpixel/pixel attraction models that consider the spatial interaction across different resolution levels (Mertens *et al.* 2006). The aforementioned approaches are iterative and are based on the assumption of spatial dependence maximization, which refers to the higher possibility of proximate subpixels being the same class as those located far apart (Curran and Atkinson 1998). However, a common concern of these algorithms is their heavy computational cost. This is usually caused by repetitive evaluations of a misfit function assessing the convergence of the simulated and the expected spatial structures (Boucher and Kyriakidis 2006).

As one geostatistical interpolation technique, downscaling cokriging has been applied to image fusion (Memarsadeghi *et al.* 2006), image sharpening (Pardo-Igúzquiza *et al.* 2006) and also SRM (Atkinson *et al.* 2008). In contrast to the aforementioned iterative methods, which are built on the spatial dependence assumption, downscaling cokriging only requires a set of variograms and cross-variograms

at the target fine resolution. The procedure is non-iterative with little computational burden and, more importantly, it explicitly takes into account pixel sizes, auto-correlations and cross-correlations (Pardo-Igúzquiza *et al.* 2006, Atkinson *et al.* 2008). It can also readily accommodate ancillary information and sparse experimental data, e.g. ground-based measurements, to constrain the uncertainty associated with the downscaling process.

Boucher and Kyriakidis (2006) adopted downscaling indicator cokriging (ICK) for SRM using variogram models from real data to describe the spatial structure of land-cover classes at a fine resolution. In their following publications, different prior structural models were tested for the reconstruction, such as analytically formulated indicator variograms and indicator variogram maps (Boucher and Kyriakidis 2007). Training images that conceptually depict the spatial patterns expected to be observed at the target resolution were also investigated by Boucher *et al.* (2008) and Boucher (2009).

In spite of the advantages, a major challenge for these approaches is that the super-resolution process requires fine-scale spatial structures that are not experimentally accessible (Atkinson *et al.* 2008). Unfortunately, none of the existing work refers to a practical means to infer the structural model. More specifically, Boucher and Kyriakidis (2006) extracted the variograms from the reference land-cover map at the target fine resolution, which is not available in real applications. Also, Boucher and Kyriakidis (2007) attempted to use a binary image of a nearby area to create an indicator semivariogram map for the impervious class as representative of the fine-scale pattern expected to occur within the region of interest. However, in reality, the possibility that land covers in two different regions exhibit exactly the same spatial structure is quite low. Synthetic reference maps repetitively showing several simple shapes, such as curvilinear and dendritic, were applied as training images by Boucher *et al.* (2008), but their capability of characterizing complex real land-cover patterns was not demonstrated.

Therefore, this article aims to develop an effective and practical SRM method accounting for localized spatial structure extraction at the target resolution. A sequentially produced with local indicator variogram (SLIV) model is proposed. Indicator variograms are first used to extract the spatial structures of different land-cover classes from high-resolution classification results over a representative local area (as opposed to the entire image). The obtained structural models are then applied to the subsequent downscaling ICK to generate conditional probability maps. A sequential assignment procedure is then undertaken to assign each fine pixel to a specific class according to the ICK-derived probabilities.

To demonstrate the practical advantages and limitations of the proposed method, a further evaluation is carried out on three diverse study areas expressing a range of typical landscapes. Detailed classification accuracy assessment is provided using the produced fine-scale land-cover maps. Of further importance is contrasting local indicator variogram modelling with models using global variograms, as it directly speaks to the applicability of the proposed method (i.e. investigating whether a small sample of high-resolution imagery is sufficient to extract the necessary spatial structures to downscale the entire low-resolution image). The proposed method was assessed using National Land Cover Data (NLCD) 30 m pixel size data as the fine-resolution target and a reproduced 270 m data set as the coarse-resolution data set to simulate downscaling that would take place by integrating, for example, Landsat and MODIS imagery.

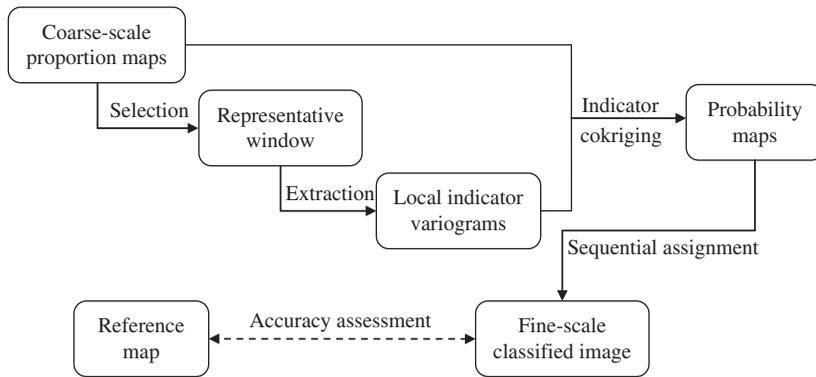


Figure 1. Flow chart of the SLIV model for super-resolution mapping.

## 2. Method

The proposed SRM methodology, named the SLIV model, is described in figure 1. The selection and extraction of local indicator variograms is discussed in the next section, followed by two sections on cokriging application and sequential assignment to produce the final fine-resolution map, and another section on accuracy assessment.

### 2.1 Local neighbourhood selection for spatial structure extraction

Spatial structural information at the target resolution has the capability to describe the fine-scale spatial pattern of land-cover distribution, and consequently it is one of the key parameters in the field of SRM. However, such structural information is not directly available in real-world situations; thus, its extraction becomes a major research question. Most existing studies mainly focused on the reconstruction methodology as opposed to the acquisition process of spatial structures.

According to the previous literature analysis, we know that structural information can be obtained either from certain assumptions, such as the assumption of spatial dependence maximization (Atkinson 1997), or from reference data, such as ground observations, existing land-cover/land-use maps, aerial photographs and high-resolution satellite images. Since the spatial dependence assumption method showed limited detail and accuracy in previous SRM attempts (Nguyen *et al.* 2006), in this article the spatial structure of land covers was extracted from reference data, specifically from existing land-cover/land-use maps.

In particular, we employed class distributions from local regions to extract spatial structures, as global structures are usually not available; if the global high-resolution reference map already exists to extract the spatial structures, then there is no reason to go back to the coarse resolution, use the spatial structure and produce a lower quality version of an existent high-resolution map. The selected local regions should be representative and reflect spatial characteristics of the entire area. Therefore, a new method was developed here to identify representative local areas to calculate fine-scale spatial structures.

To elaborate further from a geostatistical point of view, land-cover class is considered to be a regionalized variable, and the spatial variability of land-cover class can be expressed by an indicator variogram (Goovaerts 1997, Boucher *et al.* 2008). The

extraction of indicator variograms is always based on the spatial stationarity assumption (Goovaerts 1997, Boucher *et al.* 2008); that is, the spatial structure of every land-cover class across the region of interest can be measured by a stationary indicator variogram. Stationary in this context means that the variogram model should exhibit very similar fits across different subregions (Atkinson 2001b). Under the stationarity hypothesis, the spatial structure of a subregion is deemed to be consistent with that of the entire study area. Therefore, indicator variograms describing local spatial structures can theoretically approximate the global spatial pattern. The indicator variogram of each land-cover class is first calculated from a high-resolution classified subimage, and then fitted with a proper model to derive a set of parameters that capture the spatial variability (e.g. sill, range and nugget contribution). The subregion-based reconstruction approach assumes that land-cover distributions in the entire area are spatially stationary with no break or outliers. Such a stationarity hypothesis cannot be realistically applied to all regions, especially not to those large areas showing complicated patterns. Nevertheless, it is reasonable to assume a higher similarity of spatial characteristics in land-cover distributions derived from the same study area than derived from regions located far away.

A question then arises as to how to find a representative local neighbourhood from coarse fraction images that is then used to extract spatial structures at the fine (target) resolution. In this study, representative is defined as the approximation of the class percentages within a local window to those of the entire study area. More precisely, for a given local window size, land-cover proportions were first calculated within every moving window on the coarse fraction images derived from spectral unmixing. The squared difference between the proportion of each class within a local window and within the entire image was calculated. The less the total squared difference, the more representative the local area was considered to be. Therefore, the subregion with the minimum total squared difference was picked out as representative for variogram creation. Note that our selection method is only one possible option; it was chosen due to its practicality on real applications, since global and local class proportions can both be easily estimated by performing spectral unmixing on coarse-resolution images over the entire study area. It is beyond the scope of this article to find the most representative window; rather, we focus on testing the performance of locally extracted spatial structures in SRM. The window selection strategy will be further elaborated in future work.

## 2.2 Generation of class probabilities by downscaling indicator cokriging

The downscaling method is based on ICK. It was proposed by Boucher and Kyriakidis (2006) and adopted in our model to produce class probabilities at the fine resolution. Indicator variograms were first calculated from local reference data and modelled with isotropic exponential functions, and then class probabilities at every fine pixel were estimated through ICK.

The indicator variogram of the  $k$ th class pixels is defined as

$$\gamma_k(\mathbf{h}) = \frac{1}{2} \text{E} [I_k(\mathbf{u}) - I_k(\mathbf{u} + \mathbf{h})]^2, \quad (1)$$

where  $I_k(\mathbf{u})$  is a binary class indicator, defined as  $I_k(\mathbf{u}) = 1$  if pixel  $\mathbf{u}$  is identified to be class  $k$ , and zero if not, and denotes the presence or absence of the  $k$ th class at pixel  $\mathbf{u}$ .

The indicator variogram of the  $k$ th class  $\gamma_k(\mathbf{h})$ , calculated by half of the mathematical expectation (E) of the quadratic increments of all pixel pair values at a distance  $\mathbf{h}$ , characterizes the joint probability of any two pixels separated by  $\mathbf{h}$  to be different classes.

Using parametric variogram models obtained from experimental indicator variograms shown in equation (1), the probability for the  $k$ th class occurrence at fine pixel  $\mathbf{u}$  was estimated through ICK.

$$\hat{p}_k(\mathbf{u}) = \eta_k^T \mathbf{a}_k + \lambda_k^T \mathbf{j}_k + \pi_k [1 - \eta_k^T \mathbf{1}_N - \lambda_k^T \mathbf{1}_G]. \quad (2)$$

As shown in equation (2), such a probability was expressed as a weighted linear combination of coarse resolution fraction data  $\mathbf{a}_k$  and, if available, fine resolution sample indicators  $\mathbf{j}_k$  previously provided (observed or simulated) within a predefined neighbourhood.

$$\begin{bmatrix} \Gamma_k^{VV} & \Gamma_k^{Vv} \\ \Gamma_k^{Vv} & \Gamma_k^{vv} \end{bmatrix} \begin{bmatrix} \boldsymbol{\eta}_k \\ \boldsymbol{\lambda}_k \end{bmatrix} = \begin{bmatrix} \gamma_k^{VV}(\mathbf{u}) \\ \gamma_k^{Vv}(\mathbf{u}) \end{bmatrix}. \quad (3)$$

The ICK weights  $\boldsymbol{\eta}_k$  and  $\boldsymbol{\lambda}_k$  for the  $k$ th class were obtained by solving equation (3), which consisted of several indicator variogram matrices that accounted for the correlation between all pairs of observed coarse fractions ( $\Gamma_k^{VV}$ ), the correlation between all pairs of provided fine indicators ( $\Gamma_k^{vv}$ ) and the cross-correlation between all pairs of coarse fractions and fine indicators ( $\Gamma_k^{Vv}$  and  $\Gamma_k^{vV}$ ). For the entire formulation regarding the construction of the ICK system, the reader is referred to Boucher and Kyriakidis (2006, pp. 266–270), Boucher and Kyriakidis (2007, pp. 913–916) and Boucher *et al.* (2008, pp. 273–275).

In this study, indicator variograms extracted from the fine-scale classified image over the selected subregion were modelled using exponential functions. Here, we restricted the functions to exponential for modelling consistency. Variogram parameters (sill, range and nugget contribution) were identified when the fitted exponential model was closest to the calculated semivariogram points and the lowest least square error was attained. Together with coarse class proportions, the variogram fits were then used to generate a set of probability maps through the ICK process, one per land-cover class. Such ICK-derived probabilities show the possibility of the presence of every land cover at every fine-resolution pixel.

### 2.3 Sequential assignment

To create one specific fine-scale classification map that best represented the reference data, a sequential procedure was carried out on these probabilities for class allocation. In every coarse pixel, probabilities of all its fine pixels for all classes were first sorted in a descending order. Each fine pixel was then assigned to a specific class following that sequence, from the one with the highest probability until the total number of subpixels for every class was attained, so as to be consistent with the class fractions of the original coarse pixel.

An example is provided in figure 2 to further explain how the sequential assignment proceeds. Suppose the original coarse pixel represents two land-cover classes on the ground. After spectral unmixing it is found that this coarse pixel contains 25% water and 75% forest. Let us assume that a downscaling task is undertaken to divide that

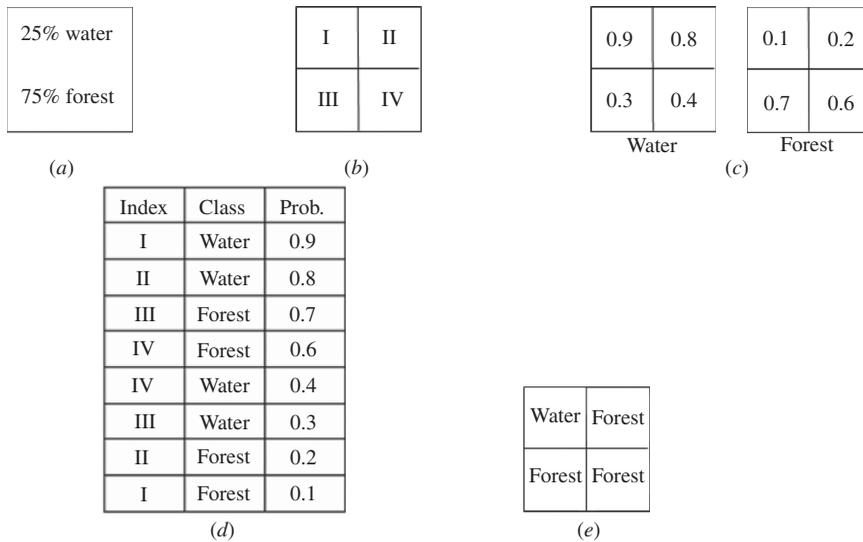


Figure 2. An example of (a) the observed coarse class proportions, (b) indices of the  $2 \times 2$  fine pixels, (c) ICK-derived probabilities, (d) the sequential assignment process and (e) the resultant fine-scale pixels.

pixel into  $2 \times 2 = 4$  fine-resolution pixels. In order to be consistent with the observed class proportions, the final fine-resolution representation should have one fine pixel as water (25% presence) and three fine pixels as forest (75% presence). Figure 2(c) shows the two probability maps (one per class) generated from the ICK process. The sum of class probabilities at every fine pixel is constrained to be 1. Ranking all these probabilities in figure 2(d), we first label fine pixel I as water for its highest probability of 0.9 and remove the last item from the table showing the probability of the forest occurrence at the same location. Due to the occupancy completeness of the water class, the items showing probabilities of the other three fine pixels being water are removed as well. The remaining portion suggests assigning fine pixels II, III and IV to the forest class. Note that our sequential assignment procedure does not consider only the class with the highest probability; in that case, two water fine pixels at I and II would have been selected resulting in 50% coverage by both water and forest. This would not match the observed class proportions (25% for water and 75% for forest), and would make the resultant fine-resolution representation incorrect.

#### 2.4 Accuracy assessment

Accuracy assessment is an indispensable step in image classification, as accuracy is always a key criterion to quantitatively evaluate the quality of the resultant classified map. Usually, an error matrix is produced by comparing the classification result with the reference data, and the overall accuracy, the kappa coefficient, as well as the producer's and user's accuracies for all classes are computed from the matrix.

Traditional assessment is undertaken over the entire image to provide a practical overall assessment. However, pure pixels (100% presence of a single class at the coarse resolution) do not influence the results based on the methodology used in super-resolution reconstruction. For pure pixels, any approach can result in 100% accuracy;

hence, statistical differences among the results using different algorithms cannot be fully assessed. To compensate for that, the same statistics are calculated exclusively on non-pure pixels, i.e. excluding fine pixels inside every pure coarse pixel that is completely occupied by a single class with no other class presence. Such statistics are considered to be a better reflection of the performance of different super-resolution algorithms.

At the final step, the produced fine-resolution classification map was compared with the reference data. To further evaluate the difference between the two classification results, a Z-test was performed (Congalton and Green 1999). This statistic tests whether one classification is significantly different from the other and it is expressed as

$$Z = \frac{|\hat{K}_1 - \hat{K}_2|}{\sqrt{\hat{\text{var}}(\hat{K}_1) + \hat{\text{var}}(\hat{K}_2)}}, \quad (4)$$

where  $\hat{K}_1$  and  $\hat{K}_2$  denote the estimates of the kappa coefficient for two independent error matrices, and  $\hat{\text{var}}(\hat{K}_1)$  and  $\hat{\text{var}}(\hat{K}_2)$  denote the corresponding estimates of the sample variances, respectively (Congalton and Green 1999). The critical value at the 99% confidence level is 2.58; therefore, values larger than 2.58 indicate a statistically significant difference.

A detailed step-by-step example follows in §4.1, with additional tests on two other sites in §4.2. The degree as to whether selective local structures can replace globally acquired ones is investigated in §4.3.

### 3. Materials

#### 3.1 National Land Cover Data in 2001

As a product of the Multi-Resolution Land Characteristics Consortium (MRLC), the National Land Cover Database 2001 (NLCD 2001) is a raster-based land-cover classification with a 30 m × 30 m spatial resolution over all 50 US states and Puerto Rico. It was created from a set of data layers, including multi-season Landsat 5 and Landsat 7 images mostly acquired in 2001, digital elevation model (DEM)-based derivatives (e.g. slope and aspect), as well as other ancillary data sets (Homer *et al.* 2004, 2007). Here, a subset of the NLCD 2001 over the study area was chosen as the reference map that for the purposes of this article we assume presents the fine-resolution land-cover distribution. We understand that NLCD is not a perfectly accurate product; however, it was deemed appropriate for our proof-of-concept testing.

Particularly, the NLCD 2001 shows the spatial distribution of 16 land-cover classes across the 48 conterminous states with approximately 27 billion pixels. Since we were more interested in several generalized land covers, the original NLCD data set was converted to a simpler classification scheme, shown in table 1. Note that woody wetlands were grouped into the forest class according to some previous studies on NLCD accuracy assessments, which indicated the feasibility of a woody wetland pixel being labelled as either the wetland or forest class (Yang *et al.* 2001, Stehman *et al.* 2003). From these seven grouped classes, ice/snow and wetland were not present in our study areas. The soil class had a minimal footprint (0.02%, 0.08% and 1.19% for the three study sites) and based on visual supervision all associated pixels were reassigned to the forest class. Therefore, at the end only four classes were present – water, urban,

Table 1. The NLCD class grouping strategy.

Index	2001 land-cover class	Grouped class
11	Open water	Water
12	Perennial ice/snow	Ice/snow
21	Developed, open space	Urban
22	Developed, low intensity	
23	Developed, medium intensity	
24	Developed, high intensity	
31	Barren land	Soil
41	Deciduous forest	Forest
42	Evergreen forest	
43	Mixed forest	
52	Shrub/scrub	
90	Woody wetlands	
71	Grassland/herbaceous	Agriculture
81	Pasture hay	
82	Cultivated crops	
95	Emergent herbaceous wetlands	Wetland

forest and agriculture – a reasonable simplification for this initial practical assessment. It is expected that future assessments should diversify further to other classes and regions.

The coarse-resolution representation from where super-resolution starts is assigned to  $270 \text{ m} \times 270 \text{ m}$  for all study cases in this article. To simulate a potential MODIS–Landsat application, each coarse pixel consists of  $9 \times 9 = 81$  fine pixels. Here, coarse-resolution class fractions are directly calculated by degrading the selected fine-resolution NLCD data sets. By doing so, errors generated by spectral unmixing are excluded from accuracy assessment.

### 3.2 Study areas

Three study areas were selected to evaluate different spatial distribution patterns. All three study areas were based on the NLCD 2001 class grouping scheme previously discussed resulting in four classes: water, urban, forest and agriculture. Each land-cover class map has a ground extent of  $20.25 \text{ km} \times 20.25 \text{ km}$ , including  $675 \times 675$  fine-resolution pixels with a pixel size of  $30 \text{ m} \times 30 \text{ m}$ . A total of  $75 \times 75$  pixels are present on every coarse fraction image corresponding to the  $9 \times 9$  upsampling scheme and a  $270 \text{ m} \times 270 \text{ m}$  pixel size. The first study area is located in South Carolina centred at  $34^\circ 43' 5'' \text{ N}$ ,  $83^\circ 0' 31'' \text{ W}$  (figure 3(a)). It is a relatively flat area with typical suburban satellite development while forest and agriculture classes are spatially mixed. A significant waterbody is also present.

To further evaluate the proposed method on different land-cover distribution patterns, two additional study areas in Ohio centred at  $40^\circ 13' 52'' \text{ N}$ ,  $82^\circ 58' 25'' \text{ W}$  (figure 3(b)) and California centred at  $39^\circ 23' 42'' \text{ N}$ ,  $120^\circ 12' 6'' \text{ W}$  were tested (figure 3(c)). More specifically, the selected region in Ohio exhibits a higher developed pattern of urban areas and a wide and continuous distribution of agricultural areas. In the California site, forest appears to be the major class with an overwhelming coverage. Also, there are some unique spatial characteristics displayed in this California site: multiple lakes, scattered agricultural areas and a complicated road network.

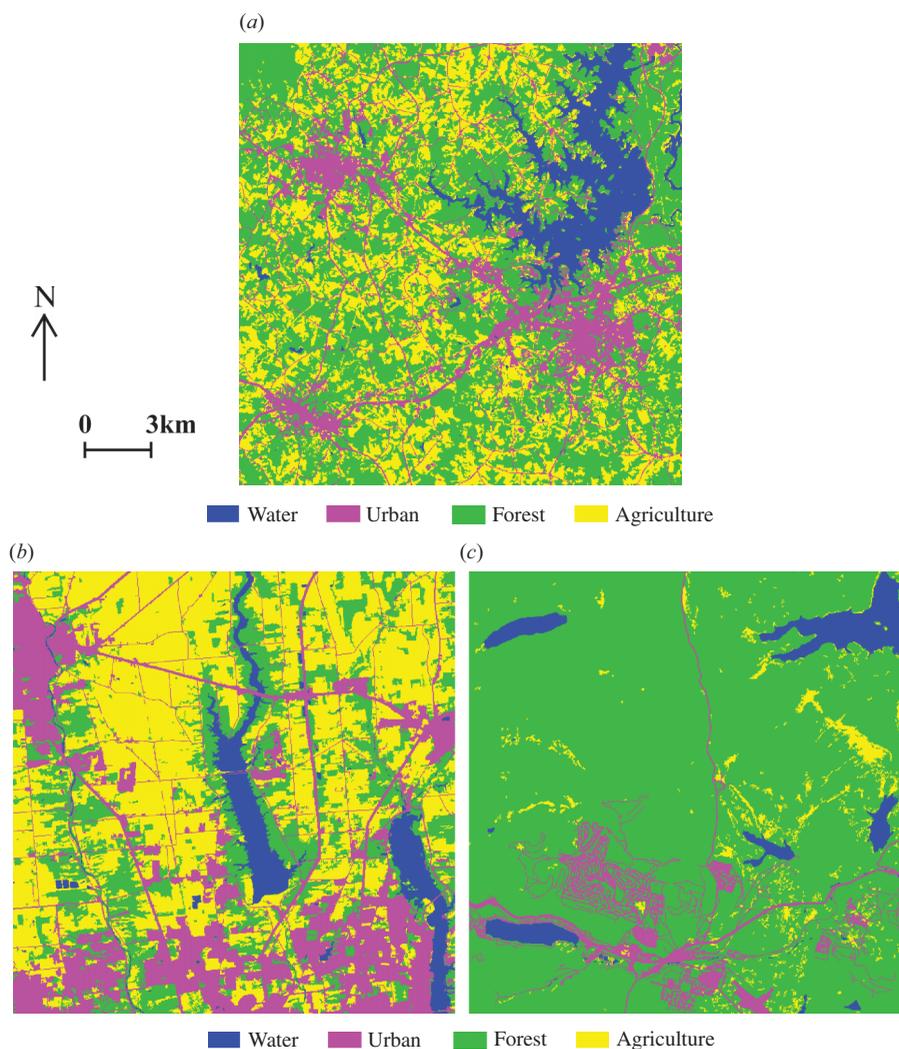


Figure 3. Land-cover class maps in (a) South Carolina, (b) Ohio and (c) California (NLCD 2001).

#### 4. Experimental results

All experiments used the NLCD regrouped maps (figure 3) as the target (fine) resolution with  $30\text{ m} \times 30\text{ m}$  pixel size. The original (coarse) resolution representation was assigned to  $270\text{ m} \times 270\text{ m}$ . Here, each coarse pixel consists of  $9 \times 9 = 81$  fine pixels, which is similar to a potential Landsat–MODIS integration. To avoid errors associated with spectral unmixing, the coarse-resolution fraction images in all experiments were created directly by upsampling fine-resolution representations. As no experimental errors were involved, the class fraction images could be viewed as the outputs of a perfect spectral unmixing process of data with a  $270\text{ m}$  spatial resolution; thus, the following tests focused only on the SRM performance. It should be noted that global target-resolution representations (figure 3) are not necessarily required; instead,

class proportions can be produced from the original coarse-resolution image. In our study, the role of the NLCD land-cover maps is to serve as the reference data set for accuracy assessment. All algorithmic development used MathWorks' MATLAB software (Natick, MA, USA).

#### 4.1 Implementation example using the South Carolina case study

This section provides a detailed insight into the proposed methodology using the South Carolina study area. Initially, the global proportion of each class was calculated from the NLCD reference map of figure 3(a): water 8.53%, urban 14.38%, forest 49.35% and agriculture 27.74%. The fraction images, shown in figure 4, were generated by calculating the class proportions within every coarse pixel. In figure 4, black indicates the complete absence of a certain class, and white indicates a complete coverage. Fraction values between 0 and 1 are displayed with different grey levels. Fraction values between 0 and 1 are displayed with different grey levels.

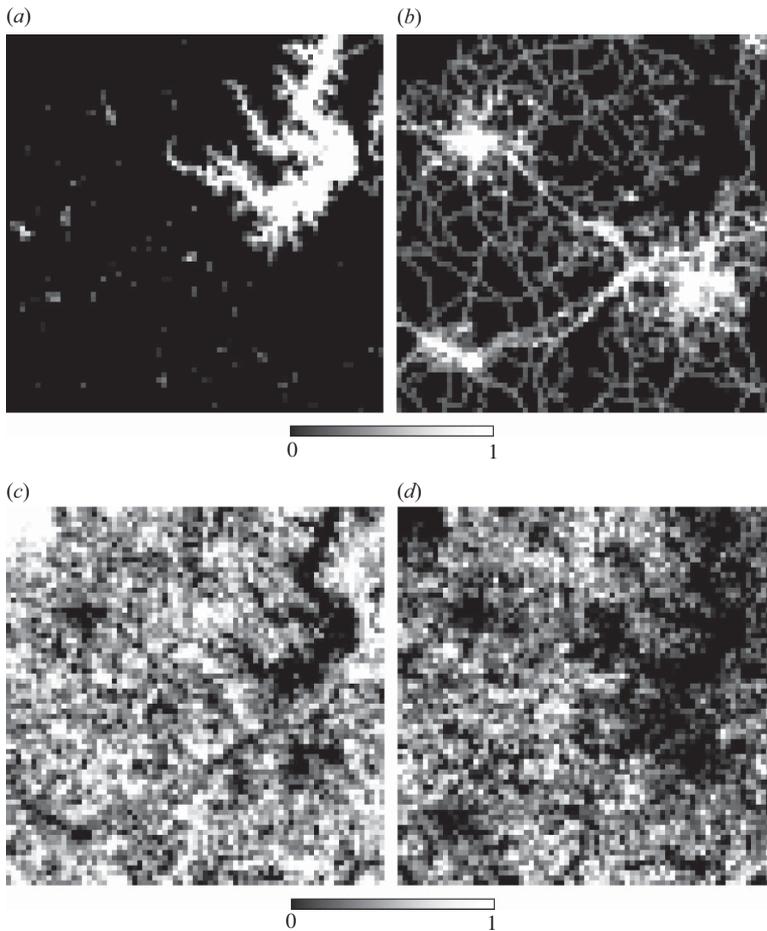


Figure 4. Fraction images via the  $9 \times 9$  upsampling scheme with pixel size  $270 \text{ m} \times 270 \text{ m}$  (coarse resolution). (a) Water, (b) urban, (c) forest and (d) agriculture.

For our implementation, a moving window of  $11 \times 11$  coarse pixels (consisting of  $99 \times 99$  fine pixels) was first examined throughout the study area. The class proportions of every window were calculated and compared with the proportions of the entire image using the summation of squared differences as a criterion of representativeness. A lower total squared difference would result in a more representative local area to the global distribution, and therefore the subregion with the minimum total squared difference was picked out and considered to be most representative. Note that our neighbourhood searching strategy is one possible solution, and in a real implementation a major advantage of the proposed method is that only coarse-resolution fraction images are required to identify the location of the local window. High-resolution data are necessary only within the identified local window to extract the spatial structure. Under the condition that very minor changes in class proportions would be brought in when the selected window is extended by one pixel in both directions, we used a local window size of  $100 \times 100$  instead of  $99 \times 99$  in the following experiments to facilitate an investigation on spatial structure extraction with a negligible effect. The selection of 100 over 99 as the window size allowed us to progressively test several window sizes of relative proportions. Sizes of  $25 \times 25$ ,  $50 \times 50$ ,  $75 \times 75$ ,  $100 \times 100$  and  $125 \times 125$  were tested to investigate the impact of the local window size on the quality of the resultant fine-resolution land-cover map. It was found that the  $100 \times 100$  and  $125 \times 125$  windows offered comparable accuracy results with those using globally derived spatial structures, which will be further described in later sections. The smaller sizes of  $25 \times 25$ ,  $50 \times 50$  and  $75 \times 75$  led to lower accuracies, probably due to the insufficient information on spatial patterns those window sizes could provide. Since our goal was to identify the smallest functional window size to increase method applicability and decrease computational load, we chose  $100 \times 100$  as the local window size for the following analysis. Figure 5 shows the selected local window for this study site, 9.22% of which is occupied by water, 14.24% by urban, 48.59% by forest and 27.95% by agriculture.

The selected local window of figure 5 was then transformed to a set of binary images, shown in figure 6, from which fine-resolution local indicator variograms were extracted for every land-cover class using equation (1). The maximum lag distance

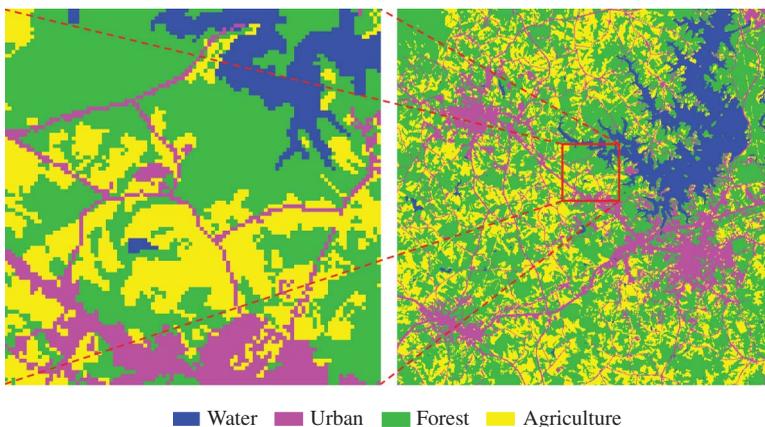


Figure 5. The selected  $100 \times 100$  local window (fine resolution) on the left, identified as a red square on the right.

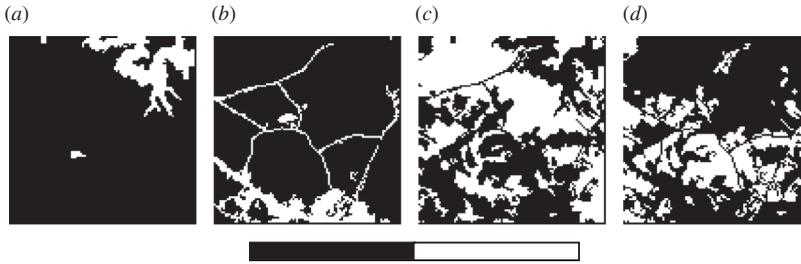


Figure 6. Binary images of the selected local window. At each fine pixel, black indicates the absence of a particular class, and white indicates the presence (fine resolution). (a) Water, (b) urban, (c) forest and (d) agriculture.

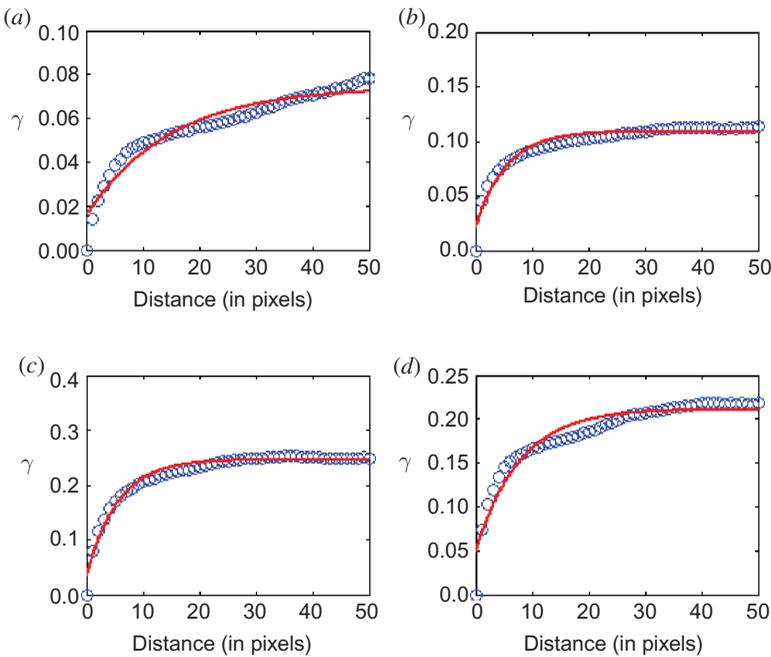


Figure 7. Indicator variograms extracted from the best local window. (a) Water, (b) urban, (c) forest and (d) agriculture.

was specified to be half of the window size, i.e. 50 fine pixels, to ensure enough pairs at large distances. An isotropic exponential function was applied to model every set of semivariogram values by changing the parameters to automatically find the best fit in terms of least square error. Indicator variogram values and the corresponding exponential fits are shown in figure 7. For each land-cover class, blue circles denote the local indicator semivariogram points at progressively increasing pixel distances, and the red continuous line represents the corresponding best exponential fit.

The variogram model parameters, nugget, sill and range, are given in table 2. The nugget effect is relatively small when compared with the associated sill value, suggesting that every class, especially water, exhibits only a small component of purely random spatial variability within the selected local window. The highest sill value

Table 2. Parameters of the local indicator variogram models.

	Nugget	Sill	Range (in pixels)
Water	0.0157	0.0586	44.7155
Urban	0.0204	0.0885	15.9573
Forest	0.0320	0.2141	16.3397
Agriculture	0.0487	0.1625	22.9940

appears at the forest and agriculture classes, suggesting that spatial variability is more evident in these classes. Consistent with the visual interpretation of figure 5, urban has the lowest value of range, indicating that the scale of spatial correlation of urban is smaller than any other types. Also, all ranges are less than 50 pixels, indicating that it was reasonable to use half of the window size as the maximum lag distance to calculate indicator variograms in this experiment.

The first step of SRM, ICK, was then carried out to estimate the probability that each fine pixel belongs to a particular class. The resulting probability maps are shown in figure 8. These ICK-derived probability maps were produced using (i) the semivariograms extracted from the  $100 \times 100$  local window acting as a representative of the entire fine-resolution map and (ii) the coarse fraction images via the  $9 \times 9$  upsampling scheme. For each class, the brighter the pixel, the higher the probability that the pixel belongs to that class. The sum of all class probabilities at every fine pixel is constrained to be 1.

The next step was to generate one specific fine-resolution land-cover class map that could best represent the reference data over the study area. Sorting all class probabilities within each coarse pixel ( $9 \text{ pixels} \times 9 \text{ pixels} \times 4 \text{ classes} = 324$  probability values in total) in a descending order, each fine pixel was assigned to a particular class, following the sequence from the one with the highest probability until the total number of subpixels for every class was attained. The same procedure was repeated for every coarse pixel; thus, all fine pixels across the region of interest were recognized as one of the four mutually exclusive classes. In this way, spatial patterns were encapsulated in the indicator variogram models for probability estimation, while class fractions of the original coarse pixels can be strictly maintained.

Since the accuracy of the resulting fine-resolution classified map is always associated with several task-related factors, e.g. spatial structure, upsampling scheme and the extent of the given study site, there is no established process to assess the algorithmic performance of SRM methods. In this article, we opted for a randomly generated reconstruction simply conditional to the  $9 \times 9$  upsampled coarse class fractions without any consideration of spatial structural patterns. The comparison between the SLIV- and random-classified maps directly shows us differences with and without local structural information being integrated into the downscaling process, and therefore the performance of the proposed method can be evaluated. A total of 1000 random simulations were created, from which the one with the highest overall accuracy was selected for accuracy assessment as the benchmark. Comparing the most accurate random simulation and the SLIV-based reconstruction with the fine-resolution NLCD reference land-cover map of figure 3(a), two error matrices were produced, and accuracy statistics, e.g. overall accuracy, kappa coefficient and per-class producer's and user's accuracies, were calculated and are presented in table 3. Statistical increases from the random to the SLIV classifications indicate the improvement in

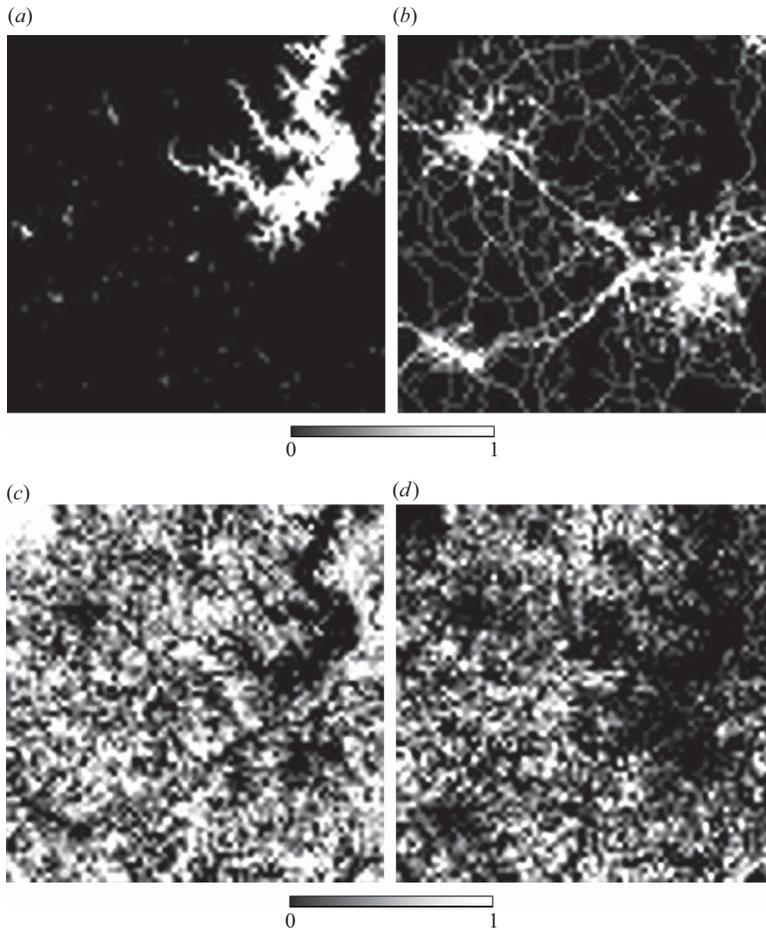


Figure 8. Conditional probabilities of class occurrence derived by ICK (fine resolution). (a) Water, (b) urban, (c) forest and (d) agriculture.

algorithmic performance achieved by using the SLIV model. Such improvement was further assessed through the  $Z$ -statistic.

The statistics in table 3 show a considerable increase in all accuracy values for the SLIV-classified image when indicator variograms extracted from the representative window were incorporated into the reconstruction process, in comparison with the most accurate random simulation simply using class fractions. The overall accuracy increases from 62.7% for the random classification to 71.4% for the SLIV-based classification, and the kappa coefficient increases from 0.427 for the random classification to 0.562 for the SLIV-based classification, showing an approximate 9% and 13.5% improvement, respectively. Also, the producer's and user's accuracies of the SLIV-reconstructed image increase around 10% for each class.

The evaluation in table 3 summarizes performance over the entire study area. However, pure pixels (100% presence in the coarse resolution) were included in that assessment; these pixels can be accurately recreated by any method. Table 4 summarizes the statistics calculated exclusively on non-pure pixels, i.e. excluding fine pixels

Table 3. Accuracy assessment of classification results over the entire image (South Carolina).

	Water	Urban	Forest	Agriculture	Total	User's accuracy (%)
Statistics for the most accurate random simulation						
Water	30 396	896	5 805	1 774	38 871	78.2
Urban	872	33 195	18 699	12 754	65 520	50.7
Forest	5 851	18 647	155 278	45 076	224 852	69.1
Agriculture	1 752	12 782	45 070	66 778	126 382	52.8
Total	38 871	65 520	224 852	126 382	455 625	
Producer's accuracy (%)	78.2	50.7	69.1	52.8		
	Overall accuracy (%)		<b>62.7</b>	Kappa coefficient		<b>0.427</b>
Statistics for the SLIV-based reconstruction						
Water	33 801	462	3 474	1 134	38 871	87.0
Urban	570	39 334	14 353	11 263	65 520	60.0
Forest	3 313	13 882	173 011	34 646	224 852	76.9
Agriculture	1 187	11 842	34 014	79 339	126 382	62.8
Total	38 871	65 520	224 852	126 382	455 625	
Producer's accuracy (%)	87.0	60.0	76.9	62.8		
	Overall accuracy (%)		<b>71.4</b>	Kappa coefficient		<b>0.562</b>
Z-statistic						87.89 ( $p < 0.0001$ )

Table 4. Accuracy assessment of classification results excluding pure pixels (South Carolina).

	Water	Urban	Forest	Agriculture	Total	User's accuracy (%)
Statistics for the most accurate random simulation						
Water	15 816	896	5 805	1 774	24 291	65.1
Urban	872	29 226	18 699	12 754	61 551	47.5
Forest	5 851	18 647	129 601	45 076	199 175	65.1
Agriculture	1 752	12 782	45 070	65 482	125 086	52.4
Total	24 291	61 551	199 175	125 086	410 103	
Producer's accuracy (%)	65.1	47.5	65.1	52.4		
	Overall accuracy (%)		<b>58.6</b>	Kappa coefficient		<b>0.357</b>
Statistics for the SLIV-based reconstruction						
Water	19 221	462	3 474	1 134	24 291	79.1
Urban	570	35 365	14 353	11 263	61 551	57.5
Forest	3 313	13 882	147 334	34 646	199 175	74.0
Agriculture	1 187	11 842	34 014	78 043	125 086	62.4
Total	24 291	61 551	199 175	125 086	410 103	
Producer's accuracy (%)	79.1	57.5	74.0	62.4		
	Overall accuracy (%)		<b>68.3</b>	Kappa coefficient		<b>0.508</b>
Z-statistic						90.93 ( $p < 0.0001$ )

inside every  $9 \times 9$  upsampled pure coarse pixel. Compared with the corresponding error matrix in table 3, only the diagonal values decrease due to the exclusion of pure pixels. Such exclusion results in a lower overall accuracy and kappa coefficient. However, there is still an approximate 10% increase in overall accuracy from 58.6% for the most accurate random classification to 68.3% for the SLIV-based classification, and the kappa coefficient is more than 15% higher when local indicator variograms were used in the reconstruction process. As all pure pixels have been excluded from the statistical evaluation, the improvements are directly achieved by the incorporation of spatial characteristics extracted from the selected window.

Tables 3 and 4 also present the Z-statistic values used for determining whether the SLIV-based reconstruction is significantly better than the most accurate random simulation. The two pairwise tests result in a Z-statistic value of 87.89 for the entire image assessment and a value of 90.93 for the assessment exclusively using non-pure pixels. Since both are considerably higher than 2.58, the SLIV-classified land-cover map is deemed significantly better than the random simulation at the 99% confidence level.

Figure 9 shows the pair of super-resolution realizations with and without prior spatial structural models. In the random case, pixels at the target resolution are randomly distributed within the coarse pixels ( $9 \times 9$  upsampling; see figure 9(a)), since they only satisfy the requirement that the proportion of each class within every coarse pixel in the classified image is consistent with the original class fraction. Since no prior spatial correlation was taken into consideration, detailed spatial patterns cannot be well identified. On the contrary, the SLIV-generated realization contains more detailed information and less salt and pepper noise, especially on the boundary of two immediately adjacent land-cover types. The reconstructed classification map, as shown in figure 9(b), improves significantly the land-cover distribution at the finer spatial resolution.

Figure 10 shows a zoom-in comparison of much smaller spatially coincident portions of the random, SLIV and reference data, each covering  $100 \times 100$  fine pixels with all classes included (black rectangle in figure 9). It is obvious that the random

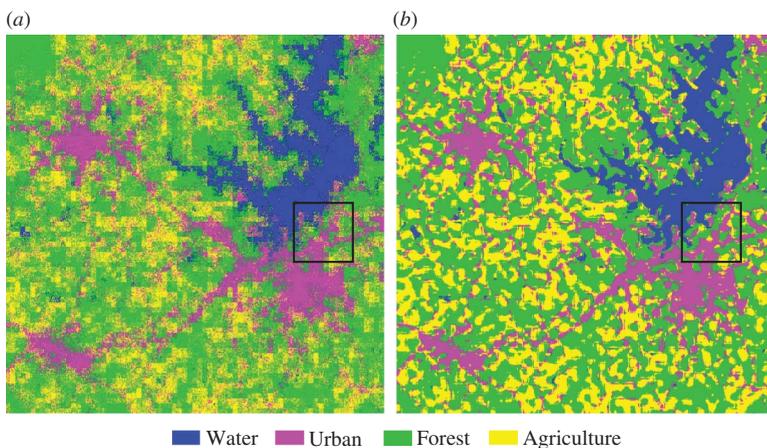


Figure 9. Visual comparison of random and SLIV-based reconstruction for South Carolina (fine resolution). Black squares indicate the locations of the focused area in figure 10. (a) The random simulation and (b) the SLIV-based reconstruction.

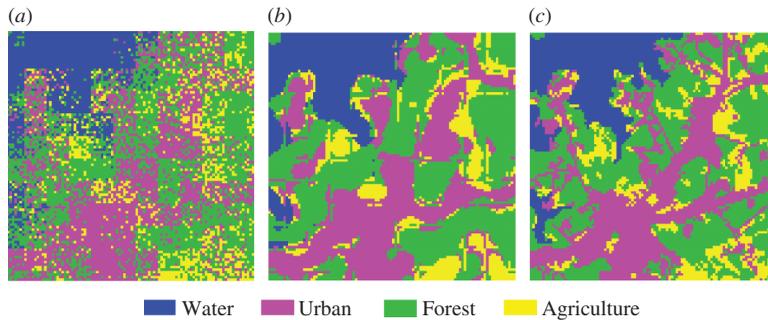


Figure 10. Visual comparison of (a) random and (b) SLIV-based reconstruction with (c) reference data for a focused area in South Carolina (fine resolution) identified as black squares in figure 9.

portion of figure 10(a) exhibits substantial noise, and two immediately adjacent land-cover types can hardly be separated from each other, even though the selected random simulation had the highest overall accuracy from all random simulations. In contrast, figure 10(b) created by the SLIV model is visually less noisy. We can see clearly the land-cover boundaries and, more importantly, the class distributions that are closer to the reference shown in figure 10(c). For example, the spatial pattern of water at the top-left corner is much better reproduced than the random case due to the encapsulation of spatial correlation of water in the downscaling process.

#### 4.2 Additional evaluation using the Ohio and California study sites

To further evaluate the proposed methodology, two other study areas with different spatial patterns from the previous case were tested in Ohio and California State, respectively, following the same process described above. Again, 1000 random simulations were produced on the basis of coarse class proportions, and the one with the highest overall accuracy was selected as the benchmark in each case.

In the Ohio case, the statistics over the entire image show a considerable increase in accuracy using the new technique (see table 5). The overall accuracy of the resulting land-cover map increases by 9% from 72.5% for the random classification to 81.6% for the SLIV-based classification using local indicator variograms as the global spatial structure measurement. The kappa coefficient increases from 0.591 for the random-classified map to 0.726 for the SLIV-classified map, with an approximate 13.5% improvement. Also, the per-class producer's and user's accuracies increase around 10% for all land-cover types.

To focus on the evaluation of the performance of the proposed method, accuracy assessment was also carried out over non-pure pixels under the  $9 \times 9$  upsampling scheme. Table 6 presents the statistics with pure pixels excluded from all classes. Although the overall accuracy and kappa coefficient are lower than those over the entire image, there is still an approximate 12% increase in overall accuracy and an approximate 18% increase in kappa coefficient. The per-class accuracies are significantly improved as well. In particular, the improvement of the producer's and user's accuracy of water is nearly 17%.

Table 5. Accuracy assessment of classification results over the entire image (Ohio).

	Water	Urban	Forest	Agriculture	Total	User's accuracy (%)
Statistics for the most accurate random simulation						
Water	21 660	1 390	3 718	1 336	28 104	77.1
Urban	1 411	79 830	12 343	18 396	111 980	71.3
Forest	3 712	12 398	66 849	25 451	108 410	61.7
Agriculture	1 321	18 362	25 500	161 948	207 131	78.2
Total	28 104	111 980	108 410	207 131	455 625	
Producer's accuracy (%)	77.1	71.3	61.7	78.2		
	Overall accuracy (%)		<b>72.5</b>	Kappa coefficient		<b>0.591</b>
Statistics for the SLIV-based reconstruction						
Water	24 313	872	2 194	725	28 104	86.5
Urban	882	90 461	8 285	12 352	111 980	80.8
Forest	2 261	8 259	80 337	17 553	108 410	74.1
Agriculture	648	12 388	17 594	176 501	207 131	85.2
Total	28 104	111 980	108 410	207 131	455 625	
Producer's accuracy (%)	86.5	80.8	74.1	85.2		
	Overall accuracy (%)		<b>81.6</b>	Kappa coefficient		<b>0.726</b>
Z-statistic						103.38 ( $p < 0.0001$ )

Table 6. Accuracy assessment of classification results excluding pure pixels (Ohio).

	Water	Urban	Forest	Agriculture	Total	User's accuracy (%)
Statistics for the most accurate random simulation						
Water	9 429	1 390	3 718	1 336	15 873	59.4
Urban	1 411	47 268	12 343	18 396	79 418	59.5
Forest	3 712	12 398	56 967	25 451	98 528	57.8
Agriculture	1 321	18 362	25 500	103 790	148 973	69.7
Total	15 873	79 418	98 528	148 973	342 792	
Producer's accuracy (%)	59.4	59.5	57.8	69.7		
	Overall accuracy (%)		<b>63.4</b>	Kappa coefficient		<b>0.456</b>
Statistics for the SLIV-based reconstruction						
Water	12 082	872	2 194	725	15 873	76.1
Urban	882	57 899	8 285	12 352	79 418	72.9
Forest	2 261	8 259	70 455	17 553	98 528	71.5
Agriculture	648	12 388	17 594	118 343	148 973	79.4
Total	15 873	79 418	98 528	148 973	342 792	
Producer's accuracy (%)	76.1	72.9	71.5	79.4		
	Overall accuracy (%)		<b>75.5</b>	Kappa coefficient		<b>0.636</b>
Z-statistic						109.26 ( $p < 0.0001$ )

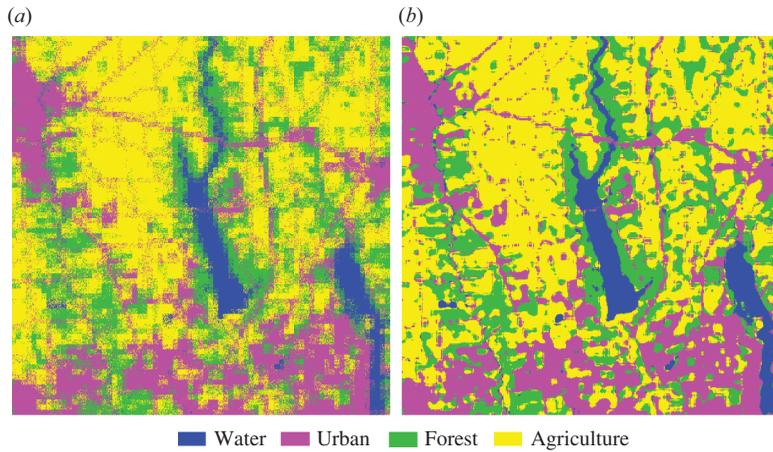


Figure 11. Visual comparison of random and SLIV-based reconstruction for Ohio (fine resolution). (a) The random simulation and (b) the SLIV-based reconstruction.

The  $Z$ -statistic values in tables 5 and 6 show that a significant difference exists between the two land-cover maps generated with and without incorporating local spatial structures into the super-resolution reconstruction, which further demonstrates the effectiveness of the proposed method from the statistical perspective. The higher  $Z$  value in table 6 suggests a higher increase in accuracy when only non-pure pixels were used as the assessment data set.

Figure 11 visually compares the classification maps generated with and without using prior structural models. Similarly, the random simulation appears quite noisy, and the land-cover patterns are poorly reproduced since no spatial structures were taken into account in the downscaling process. In contrast, the SLIV-generated realization contains more spatial details, and is much closer to the NLCD reference land-cover map of figure 3(b).

Similarly, accuracy assessments of the random and SLIV-based classification maps were carried out in the California site both over the entire image and over non-pure pixels. The statistics are summarized in tables 7 and 8. Approximately there is a 3% increase in overall accuracy and a 10.5% increase in kappa coefficient over the entire image (see table 7). In table 8, the overall accuracy and kappa coefficient calculated excluding pure pixels are more than 7% and 14% higher than the statistics in the random case, due to the incorporation of the representative local semivariograms. The  $Z$ -test results also support that such increase in accuracy is statistically significant.

From figure 12, we can see the difference between the random and the SLIV-based classification results. The reconstructed realization on the right side contains more spatial details. In particular, the edges of the water class are well delineated and very close to the reference data using the proposed method, compared with the random-generated image in which the edges of water are blocked and noisy, resulting in higher accuracy statistics shown in tables 7 and 8. The reference data are presented in figure 3(c). Also, some linear features of the urban class, i.e. roads, are better recreated with stronger connectivity; however, there are also limitations of our method as it does not fully capture the grid-type road network.

Table 7. Accuracy assessment of classification results over the entire image (California).

	Water	Urban	Forest	Agriculture	Total	User's accuracy (%)
Statistics for the most accurate random simulation						
Water	19 897	275	2 844	330	23 346	85.2
Urban	277	15 874	14 092	499	30 742	51.6
Forest	2 837	14 092	357 587	9 967	384 483	93.0
Agriculture	335	501	9 960	6 258	17 054	36.7
Total	23 346	30 742	384 483	17 054	455 625	
Producer's accuracy (%)	85.2	51.6	93.0	36.7		
	Overall accuracy (%)		<b>87.7</b>	Kappa coefficient		<b>0.560</b>
Statistics for the SLIV-based reconstruction						
Water	22 210	89	924	123	23 346	95.1
Urban	126	18 426	11 774	416	30 742	59.9
Forest	773	11 766	363 834	8110	384 483	94.6
Agriculture	237	461	7 951	8 405	17 054	49.3
Total	23 346	30 742	384 483	17 054	455 625	
Producer's accuracy (%)	95.1	59.9	94.6	49.3		
	Overall accuracy (%)		<b>90.6</b>	Kappa coefficient		<b>0.664</b>
Z-statistic						46.99 ( $p < 0.0001$ )

Table 8. Accuracy assessment of classification results excluding pure pixels (California).

	Water	Urban	Forest	Agriculture	Total	User's accuracy (%)
Statistics for the most accurate random simulation						
Water	7 504	275	2 844	330	10 953	68.5
Urban	277	14 659	14 092	499	29 527	49.7
Forest	2 837	14 092	103 409	9 967	130 305	79.4
Agriculture	335	501	9 960	6 096	16 892	36.1
Total	10 953	29 527	130 305	16 892	187 677	
Producer's accuracy (%)	68.5	49.7	79.4	36.1		
	Overall accuracy (%)		<b>70.2</b>	Kappa coefficient		<b>0.380</b>
Statistics for the SLIV-based reconstruction						
Water	9 817	89	924	123	10 953	89.6
Urban	126	17 211	11 774	416	29 527	58.3
Forest	773	11 766	109 656	8 110	130 305	84.2
Agriculture	237	461	7 951	8 243	16 892	48.8
Total	10 953	29 527	130 305	16 892	187 677	
Producer's accuracy (%)	89.6	58.3	84.2	48.8		
	Overall accuracy (%)		<b>77.2</b>	Kappa coefficient		<b>0.527</b>
Z-statistic						50.72 ( $p < 0.0001$ )

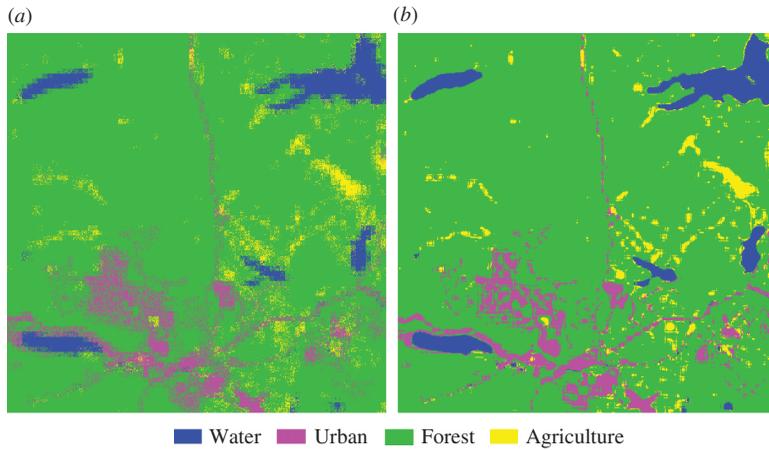


Figure 12. Visual comparison of random and SLIV-based reconstruction for California (fine resolution). (a) The random simulation and (b) the SLIV-based reconstruction.

### 4.3 Contrasting local versus global windows for spatial structure extraction

An important evaluation for realistic implementation of the proposed SLIV-based reconstruction is the trade-off from moving from a global spatial structure extraction to a local one, where only a small portion of the study area is used. This is particularly important as high-resolution reference data are needed only in the local window, which significantly increases the efficiency and applicability of the proposed method. Figure 13 displays this assessment using overall accuracies and kappa coefficients of the SLIV-generated reconstructions from local and global indicator variogram models. The local indicator variogram was extracted using our aforementioned selection process, while the global indicator variogram used the entire fine-resolution image to extract the spatial structure. The accuracy statistics derived by importing variograms from the selected  $100 \times 100$  window and the global reference map are almost

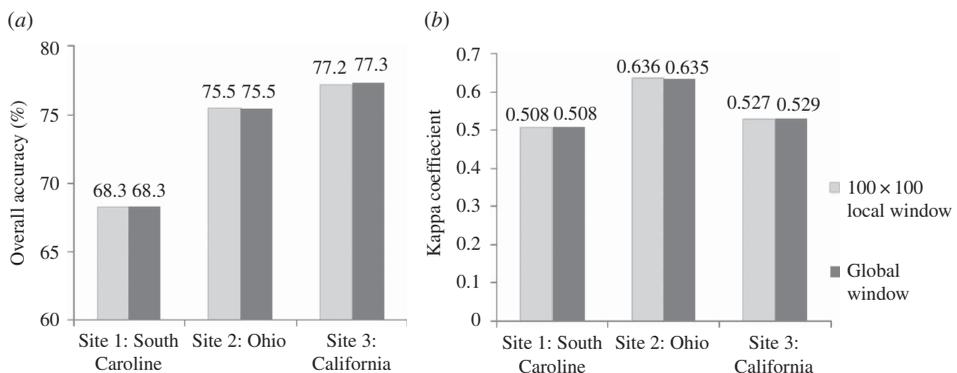


Figure 13. (a) Overall accuracies and (b) kappa coefficients of the SLIV-based reconstructions using indicator variogram models extracted from the selected  $100 \times 100$  local window and the global fine-resolution reference map.

identical. This validates the selection process for the local window and enforces the assumption that spatial stationarity is a reasonable hypothesis for the given problem.

## 5. Discussion and conclusions

SRM is an inherently under-determined inverse problem, where uncertainty (or ambiguity) is commonly associated with the coarse-to-fine downscaling process. To limit the inherent uncertainty, it is necessary to invoke some prior information of spatial structure pertaining to the target (fine) resolution. Such prior information can be implicit assumptions or explicit reference data. In this article, a set of local indicator variogram models extracted from the fine-scale classification map of a representative local subregion were used to extract structural information. Under the hypothesis of stationarity, an extension of the geostatistical intrinsic hypothesis to remote-sensing applications, it is theoretically reasonable that such models may capture the global spatial variability of land-cover classes at the target scale.

The results indicated an improvement of 7–12% in terms of overall accuracy and 15–18% in terms of kappa coefficient when pure pixels were excluded from the assessment data set. The magnitude of improvements suggests that the approach is statistically superior to random assignments of pixel classes. Therefore, spatial structure does contribute significantly in SRM. Looking beyond statistics, two important discussion topics are of interest: the practical implementation feasibility and limitations of the existing approach. From the practical perspective, SRM is an ambiguous process with coarse class proportions as the only constraint. In our method, the spatial structure is extracted from local variograms, a significant step forward towards wide adoption of the methodology. It is not realistic to expect that large portions of the study area will be offered in both coarse and fine resolution. As one possible option, our neighbourhood selection method picked out one specific window with a size of  $100 \times 100$  fine-resolution pixels, approximately 2% of the entire study area ( $675 \times 675$ ). The identical results in accuracy showed that only a small portion of fine-resolution data is necessary. Furthermore, the simulated coarse-to-fine resolution example moving from 270 to 30 m is a plausible case with existing sensors, for example reproducing subpixel classifications from MODIS- to Landsat-type resolutions. Such methods are essential for mapping at continental to global scales, where limitations in availability and consistency of high-resolution data mandate alternative sensors that exhibit larger spatial footprints with higher acquisition frequency.

The aforementioned MODIS–Landsat integration is a potential application of the proposed SRM algorithm to land-cover classification. Although most pixel-based and subpixel classifiers can be employed on the MODIS imagery, an inherent constraint commonly exists that spatial distribution at a finer scale, e.g. Landsat-type resolutions, is still unpredictable. The method developed in this study provides an alternative to increase the spatial resolution of the resultant land-cover classification maps. In contrast to the existing studies using practically unavailable global structures, we first identify a small neighbourhood where high-resolution data should be acquired. Spatial structures delineated from the selected window are then incorporated into the downscaling super-resolution process representing the global patterns of different land-cover classes, and therefore a more accurate result can be generated at the fine resolution compared with random simulations simply subject to the constraint on class proportions. The improvement in accuracy and the additional spatial information gained by the proposed SRM method is considerable in real

applications of land-cover classification, especially when data (e.g. cloud coverage) and cost limitations do not allow detailed analysis over a large study area.

It is necessary to point out that there is no established process for SRM algorithmic evaluation. Some existing studies selected one algorithm previously published by other researchers as the benchmark, while others simply used results generated from pixel-based classifiers at the target resolution. For example, Boucher and Kyriakidis (2006, 2007) and Boucher *et al.* (2008) generated a set of realizations. All were considered as being equally reconstructed with no accuracy assessment performed. Other papers (e.g. Verhoeve and De Wulf 2002, Kasetkasem *et al.* 2005, Makido and Shortridge 2007), which used real remote-sensing data sets rather than artificial or simulated images with very simple objects, created 70–80% accurate results depending on the complexity of land-cover distribution. These accuracies are comparable with those of our South Carolina case. Nevertheless, we acknowledge that it is not easy to identify one of the SRM approaches for comparison due to constraints related to data and model availability, model assumptions and task-related factors such as spatial structure of a given site, difference in the upsampling scheme, pixel sizes and the extent of the study region. Furthermore, when performance was assessed in existing work, it was done mostly on a single (occasionally two) study site. In this article, we opted for a random simulation assessment on three sites with different characteristics to widen the applicability of SRM from the practical perspective. As further SRM studies emerge, our simulated data set and task may act as an initial common assessment data set and metric – we plan on providing the reference and predicted data freely to other researchers to facilitate comparisons and advance the state of the art.

The proposed method also exhibits some limitations. First, although our model can be used for SRM in any study area, spatial variability of land-cover classes does impact the resulting accuracy. Generally, classes presenting irregular and discrete distribution are more difficult to be correctly predicted. For example, reproductions of urban and agriculture in the California case result in low accuracies, which is probably caused by the high complexity of patterns of the road network and agricultural areas that cannot be characterized by local isotropic variograms. Second, the sequential assignment procedure in this study was undertaken using the ICK-derived probability maps without any consideration of additional object connectivity. This might be partially responsible for the break-ups of roads in all three cases. Finally, the SLIV-based reconstructions exhibit some horizontal and vertical artefacts on the boundaries of two immediately adjacent coarse pixels, as shown in figure 10(b), possibly due to the algorithmic inability to look outside the coarse pixels when estimating class probabilities of an inner fine pixel by the ICK method.

Future work will focus on using anisotropic variograms to better characterize the spatial pattern of complicated land covers in reality. Also, we plan to import image segmentation algorithms, e.g. region growing and morphological operations, to the generation of a classification map from a set of ICK-derived probabilities. To further reduce the horizontal and vertical artefacts, the coverage of surrounding subpixels could be expanded to a circular area with class variogram range as radius and used to estimate the probability of every class occurrence at the currently visited subpixel. An appropriately designed filter might also be useful to smooth out some blocky artefacts. Additional tests with many local variograms extracted from different subsets should be undertaken to account for the more common spatial non-stationary case and enhance the applicability of this work.

## Acknowledgements

This research was supported by the National Science Foundation of China (Grant number 40372130) and by the National Aeronautics and Space Administration through the New Investigator and Biodiversity Programs (awards NNX08AR11G and NNX09AK16G).

## References

- APLIN, P. and ATKINSON, P.M., 2001, Sub-pixel land cover mapping for per-field classification. *International Journal of Remote Sensing*, **22**, pp. 2853–2858.
- ATKINSON, P.M., 1997, Mapping sub-pixel boundaries from remotely sensed images. In *Innovations in GIS IV*, Z. Kemp (Ed.), pp. 166–180 (London: Taylor and Francis).
- ATKINSON, P.M., 2001a, Super-resolution target mapping from soft-classified remotely sensed imagery. In *Proceedings of the 6th International Conference on GeoComputation*, 24–26 September 2001, Brisbane, QLD (Greenwich: GeoComputation CD-ROM).
- ATKINSON, P.M., 2001b, Geographical information science: GeoComputation and nonstationarity. *Progress in Physical Geography*, **25**, pp. 111–122.
- ATKINSON, P.M., 2004, Super-resolution land cover classification using the two-point histogram. In *GeoENV IV: Geostatistics for Environmental Applications*, X. Sanchez-Vila, J. Carrera and J.J. Gomez-Hernandez (Eds.), pp. 15–28 (The Netherlands: Kluwer).
- ATKINSON, P.M., 2005, Sub-pixel target mapping from soft-classified, remotely sensed imagery. *Photogrammetric Engineering and Remote Sensing*, **71**, pp. 839–846.
- ATKINSON, P.M., PARDO-IGÚZQUIZA, E. and CHICA-OLMO, M., 2008, Downscaling cokriging for super-resolution mapping of continua in remotely sensed images. *IEEE Transactions on Geoscience and Remote Sensing*, **46**, pp. 573–580.
- BATESON, C.A., ASNER, G.P. and WESSMAN, C.A., 2000, Endmember bundles: a new approach to incorporating endmember variability into spectral mixture analysis. *IEEE Transactions on Geoscience and Remote Sensing*, **38**, pp. 1083–1094.
- BOUCHER, A., 2009, Sub-pixel mapping of coarse satellite remote sensing images with stochastic simulations from training images. *Mathematical Geosciences*, **41**, pp. 265–290.
- BOUCHER, A. and KYRIAKIDIS, P.C., 2006, Super-resolution land cover mapping with indicator geostatistics. *Remote Sensing of Environment*, **104**, pp. 264–282.
- BOUCHER, A. and KYRIAKIDIS, P.C., 2007, Integrating fine scale information in super resolution land cover mapping. *Photogrammetric Engineering and Remote Sensing*, **73**, pp. 913–921.
- BOUCHER, A. and KYRIAKIDIS, P.C., 2008, Geostatistical solutions for super-resolution land cover mapping. *IEEE Transactions on Geoscience and Remote Sensing*, **46**, pp. 272–283.
- CAMPELL, J.B., 1996, *Introduction to Remote Sensing* (London: Taylor and Francis).
- CONGALTON, R.G. and GREEN, K., 1999, *Assessing the Accuracy of Remotely Sensed Data: Principles and Practice*, pp. 49–53 (Boca Raton, FL: Lewis Publishers, CRC Press).
- CURRAN, P.J. and ATKINSON, P.M., 1998, Geostatistics and remote sensing. *Progress in Physical Geography*, **22**, pp. 61–78.
- FISHER, P.F., 1997, The pixel: a snare and a delusion. *International Journal of Remote Sensing*, **18**, pp. 679–685.
- FOODY, G.M., 1996, Approaches for the production and evaluation of fuzzy land-cover classifications from remotely sensed data. *International Journal of Remote Sensing*, **17**, pp. 1317–1340.
- FOODY, G.M., 2001, Monitoring the magnitude of land-cover change around the southern limits of the Sahara. *Photogrammetric Engineering and Remote Sensing*, **67**, pp. 841–847.
- FOODY, G.M. and DOAN, H.T.X., 2007, Variability in soft classification prediction and its implications for sub-pixel scale change detection and super resolution mapping. *Photogrammetric Engineering and Remote Sensing*, **73**, pp. 923–933.

- GOOVAERTS, P., 1997, *Geostatistics for Natural Resources Evaluation* (New York: Oxford University Press).
- HAERTEL, V., SHIMABUKURO, Y.E. and ALMEIDA-FILHO, R., 2004, Fraction images in multi-temporal change detection. *International Journal of Remote Sensing*, **25**, pp. 5473–5489.
- HOMER, C., DEWITZ, J., FRY, J., COAN, M., HOSSAIN, N., LARSON, C., HEROLD, N., MCKERROW, A., VANDRIEL, J.N. and WICKHAM, J., 2007, Completion of the 2001 National Land Cover Database for the coterminous United States. *Photogrammetric Engineering and Remote Sensing*, **73**, pp. 337–341.
- HOMER, C., HUANG, C., YANG, L., WYLIE, B. and COAN, M., 2004, Development of a 2001 National Land Cover Database for the United States. *Photogrammetric Engineering and Remote Sensing*, **70**, pp. 829–840.
- KASETKASEM, T., ARORA, M. and VARSHNEY, P., 2005, Super-resolution land cover mapping using a Markov random field approach. *Remote Sensing of Environment*, **96**, pp. 302–314.
- MAKIDO, Y. and SHORTRIDGE, A., 2005, Land cover mapping at subpixel scales: unraveling the mixed pixel. In *Proceedings of the 8th International Conference on GeoComputation*, 1–3 August 2005, Ann Arbor, MI (Greenwich: GeoComputation CD-ROM).
- MAKIDO, Y. and SHORTRIDGE, A., 2007, Assessing alternatives for modeling the spatial distribution of multiple land cover classes at sub-pixel scales. *Photogrammetric Engineering and Remote Sensing*, **73**, pp. 935–943.
- MEMARSADEGHI, N., LE MOIGNE, J., MOUNT, D.M. and MORISSETTE, J., 2006, A new approach to image fusion based on cokriging. In *Proceedings of the 8th International Conference on Information Fusion*, 25–29 July 2005, Philadelphia, PA (Sunnyvale, CA: International Society of Information Fusion).
- MERTENS, K.C., DE BAETS, B., VERBEKE, L.P.C. and DE WULF, R.R., 2006, A sub-pixel mapping algorithm based on sub-pixel/pixel spatial attraction models. *International Journal of Remote Sensing*, **27**, pp. 3293–3310.
- MERTENS, K.C., VERBEKE, L.P.C., DUCHEYNE, E.I. and DE WULF, R.R., 2003, Using genetic algorithms in sub-pixel mapping. *International Journal of Remote Sensing*, **24**, pp. 4241–4247.
- MERTENS, K.C., VERBEKE, L.P.C., WESTRA, T. and DE WULF, R.R., 2004, Sub-pixel mapping and sub-pixel sharpening using neural network predicted wavelet coefficients. *Remote Sensing of Environment*, **91**, pp. 225–236.
- NGUYEN, M.Q., ATKINSON, P.M. and LEWIS, H.G., 2006, Superresolution mapping using a Hopfield neural network with fused images. *IEEE Transactions on Geoscience and Remote Sensing*, **44**, pp. 736–749.
- PARDO-IGÚZQUIZA, E., CHICA-OLMO, M. and ATKINSON, P.M., 2006, Downscaling cokriging for image sharpening. *Remote Sensing of Environment*, **102**, pp. 86–98.
- RICHARDS, J.A. and JIA, X., 1999, *Remote Sensing Digital Image Analysis*, 3rd ed. (Berlin: Springer-Verlag).
- SCHOWENGERDT, R.A., 1997, *Remote Sensing: Models and Methods for Image Processing* (San Diego, CA: Academic).
- SMALL, C., 2004, The Landsat ETM+ spectral mixing space. *Remote Sensing of Environment*, **93**, pp. 1–17.
- STEHMAN, S.V., WICKHAM, J.D., SMITH, J.H. and YANG, L., 2003, Thematic accuracy of the 1992 National Land-Cover Data (NLCD) for the eastern United States: statistical methodology and regional results. *Remote Sensing of Environment*, **86**, pp. 500–516.
- TATEM, A.J., LEWIS, H.G., ATKINSON, P.M. and NIXON, M.S., 2001, Super resolution target identification from remotely sensed images using a Hopfield neural network. *IEEE Transactions on Geoscience and Remote Sensing*, **39**, pp. 781–796.
- TATEM, A.J., LEWIS, H.G., ATKINSON, P.M. and NIXON, M.S., 2002, Super resolution land cover pattern prediction using a Hopfield neural network. *Remote Sensing of Environment*, **79**, pp. 1–14.

- TATEM, A.J., LEWIS, H.G., ATKINSON, P.M. and NIXON, M.S., 2003, Increasing the spatial resolution of agricultural land cover maps using a Hopfield neural network. *International Journal of Geographical Information Science*, **17**, pp. 647–672.
- THORNTON, M.W., ATKINSON, P.M. and HOLLAND, D.A., 2006, Sub-pixel mapping of rural land cover objects from fine spatial resolution satellite sensor imagery using super-resolution pixel-swapping. *International Journal of Remote Sensing*, **27**, pp. 473–491.
- TSO, B. and MATHER, P., 2001, *Classification Methods for Remotely Sensed Data* (London: Taylor and Francis).
- VERHOEYE, J. and DE WULF, R., 2002, Land cover mapping at sub-pixel scales using linear optimization techniques. *Remote Sensing of Environment*, **79**, pp. 96–104.
- YANG, L., STEHMAN, S.V., SMITH, J.H. and WICKHAM, J.D., 2001, Short communication: thematic accuracy of MRLC land-cover for the eastern United States. *Remote Sensing of Environment*, **76**, pp. 418–422.
- ZHAN, Q., MOLENAAR, M. and LUCIEER, A., 2002, Pixel unmixing at the sub-pixel scale based on land cover class probabilities: application to urban areas. In *Uncertainty in Remote Sensing and GIS*, G.M. Foody and P.M. Atkinson (Eds.), pp. 59–76 (Chichester: John Wiley & Sons Ltd.).