The effect of the point spread function on downscaling continua

Qunming Wang [a], Yijie Tang [a], Peter M. Atkinson [b,c*]

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

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

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

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

Abstract: The point spread function (PSF) is ubiquitous in remote sensing. This paper investigated the effect of the PSF on the downscaling of continua. Geostatistical approaches were adopted to incorporate explicitly, and reduce the influence of, the PSF effect in downscaling. Two general cases were considered: univariate and multivariate. In the univariate case, the input coarse spatial resolution image is the only image available for downscaling. Area-to-point kriging was demonstrated to be a suitable solution in this case. For the multivariate case, a finer spatial resolution image (or images) observed under different conditions (e.g., at a different wavelength) is available as auxiliary data for downscaling. Area-to-point regression kriging was shown to be a suitable solution for this case. Moreover, a new solution was developed for estimating the PSF in image scale transformation. The experiments show that the PSF effect influences downscaling greatly and that downscaling can be enhanced obviously by considering the PSF effect through the geostatistical approaches and the PSF estimation solution proposed.

Keywords: Remote sensing, downscaling, image fusion, point spread function (PSF), geostatistics.

1. Introduction

Downscaling is an important technique in remote sensing (Atkinson, 2013). It refers to the process of predicting pixel values with a reduced (i.e., a downwards or disaggregation process) spatial coverage (i.e., finer spatial scale) based on observed images. By downscaling, each original pixel is broken into multiple sub-pixels and the spatial content is characterized by more data as a result. Downscaling aims to increase the spatial resolution of images and reproduce more information that cannot be presented by the original data. Downscaling can, thus, help users to make more reliable decisions and has gained increasing attention in various domains (not only in remote sensing, but also in ecology (Li et al., 2011; Zhang et al., 2017), hydrology (Jia et al., 2011; Xu et al., 2015), environmental science (Shen et al., 2016), health science (Trang et al., 2016), sociology (Wardrop et al., 2018), etc.) in recent years.

For downscaling in remote sensing, two classes of goal can be distinguished according to the final predictions, that is, downscaling continua and sub-pixel mapping (SPM) (Atkinson, 1997). The former goal means that continua (e.g., in units of reflectance, digital number, radiance, brightness, etc.) are predicted, while the latter goal means that categories (i.e., land cover class labels) are predicted, also termed super-resolution mapping in the remote sensing literature (Atkinson, 2009; Li et al., 2017). For SPM, the land cover class labels of sub-pixels within an original coarse pixel are predicted based on the constraints imposed by the observed coarse data (either in terms of the original coarse spectra (Zhao et al., 2015) or coarse land cover proportions predicted by pre-processing spectral unmixing (Wang et al., 2020a)) and the objective of maximizing spatial dependence (Chen et al., 2018) or matching spatial prior information (Ge et al., 2016). SPM is essentially a hard classification technique performed at a finer spatial resolution, and it is a more specific issue (i.e., land cover mapping) than downscaling continua.

It is widely acknowledged that the point spread function (PSF) effect exists ubiquitously in remote sensing images. This means that the signal attributed to a pixel is not only determined by the contributions from *within* its nominal spatial coverage, but also influenced by its neighboring pixels. As a result, the PSF leads to

ambiguous artifacts in images. The PSF effect is caused mainly by the optical system, motion of the sensor platform, detector and electronics (Huang et al., 2002; Kaiser and Schneider, 2008). It influences greatly the image quality and imposes a fundamental limit on the amount of information in the images (Townshend et al., 2000; Manslow and Nixon, 2002). The main objective of downscaling is to characterize the spatial content (e.g., land cover) in the original pixel with multiple sub-pixels and, thus, to increase the ability to present spatial detail. It is obvious that the PSF effect impacts the ability to perform downscaling. Therefore, it is of great interest to investigate effective solutions to reduce the PSF effect in downscaling and produce more accurate predictions.

In our previous research (Wang and Atkinson, 2017), the PSF effect was considered in SPM, which has been typically ignored in this technique. Specifically, the relation between the coarse land cover proportions and sub-pixel classes was deduced mathematically based on the existence of the PSF effect, and an optimization-based method was developed to modify the coarse proportion constraint in conventional SPM. The accuracy of SPM was increased significantly by considering the PSF effect. For downscaling continua, however, the problem is different to that for SPM and is more general. The prediction of SPM is a hard-classified land cover map and it is used mainly for more reliable thematic mapping. Alternatively, downscaling continua aims to predict data with the *same form* as the original input and the predictions of downscaling continua can be used for more widespread applications, not only land cover mapping, but also various monitoring based on remote sensing retrieval (such as monitoring vegetation dynamics (Gao et al., 2017), land surface temperature (Wang et al., 2020b), soil moisture (Peng et al., 2017), and other variables). The PSF issue in downscaling continua is substantially different to that for SPM. It has been a long standing issue in remote sensing images, but has not gained sufficient attention. Thus, it is important to develop approaches to tackle the PSF effect in downscaling continua. This will also support more reliable use of downscaling predictions in a variety of applications.

For downscaling continua, two cases needs to be identified based on the number of covariates (i.e., input): univariate and multivariate. The former refers to the common case where only a single coarse spatial

resolution image is available for downscaling, while the latter refers to the case where multiple images at different spatial resolutions (normally at a finer spatial resolution) are available as auxiliary data to aid the downscaling process. Specifically, the auxiliary data cover the same scene as the coarse data but are acquired under different conditions, such as by different sensors or channels. A typical example for the multivariate case is multi-resolution image fusion, such as the commonly known pan-sharpening (Vivone et al., 2015) and more general issue of fusing coarse hyperspectral or multispectral with fine spatial resolution multispectral images (Loncan et al., 2015; Wang et al., 2020c).

In this paper, geostatistics-based approaches are developed to account for the PSF effect in downscaling continua. The downscaling process involves a change of pixel size (i.e., change of support in geostatistics), and area-to-point kriging (ATPK)-based solutions have been developed to tackle this problem (Kyriakidis, 2004; Wang et al., 2015). Specifically, ATPK-based solutions can predict a support smaller than that of the original pixel and account for the PSF effect by semivariogram modeling at different scales (Atkinson et al., 2008; Pardo-Iguzquiza et al., 2011). More importantly, they have the appealing advantage of perfect coherence with the original coarse pixel values (Kyriakidis, 2004).

In downscaling, the PSF of interest is not the measurement PSF, but rather the transfer function between images at the original coarse and target fine spatial resolutions. The transfer function is subtly different to the measurement PSF. It is equal to the convolutional difference between the two measurement PSFs. For simplicity, we refer to this transfer function here as the transformation PSF. However, it is important to bear in mind this distinction. To consider the PSF in downscaling, the specific transformation PSF needs to be known or estimated in advance. In the existing literature on downscaling (both SPM and downscaling continua), most methods are performed based on the assumption that the PSF is known exactly or can be determined empirically by referring to additional information such as the parameters provided by the manufacturer (Atkinson, 1997; Aiazzi et al., 2006; Bevilacqua et al., 2014). Due to the aging process, however, the early parameters may not be suitable. Thus, the determination of an effective transformation PSF in downscaling remains an open problem and it is of great value to develop methods for PSF estimation before downscaling.

For the univariate case, it is always challenging to estimate the effective PSF, as no auxiliary data exist at the target fine spatial resolution. For a given fine spatial resolution, the optimal PSF needs to be estimated specifically, if possible. This means that for the univariate case, there may be no universal PSF for various fine spatial resolutions. For the multivariate case, however, there exist auxiliary data covering the same scene at the target fine spatial resolution, although the data are acquired at different wavelengths or by different sensors. The relation between the coarse and fine data can provide valuable information for estimating the effective transformation PSF. Recently, several studies have developed methods for estimation of the PSF. Vivone et al. (2015) proposed a semi-blind deconvolution approach, which assumes that all coarse bands have the same PSF. The assumption is inappropriate, especially when the sensor exhibits different spatial responses across bands (Vivone et al., 2015). To release this strong assumption, in their later research (Vivone et al., 2019), a multiband filter estimation method was proposed to estimate the PSF for each band separately. However, the methods were developed mainly for the pan-sharpening case where only a single fine PAN band is available. It is necessary to develop solutions which can estimate the PSF for each coarse band separately and which are also suitable for the more complex case involving multiple fine bands. This is another goal of this paper.

The contributions of this paper are listed as follows:

1) ATPK-based solutions (including original ATPK and area-to-point regression kriging (ATPRK)) were applied to reduce the PSF effect in downscaling continua (in both univariate and multivariate cases) and produce more accurate downscaling predictions.
2) A new solution was proposed to estimate the effective transformation PSF for the multivariate case. The solution estimates the PSF for each coarse band separately and is also suitable for dealing with multiple fine bands.
3) Case studies on data acquired by the recently launched satellite sensors (including the European Space Agency (ESA) Sentinel-2 and Chinese Gaofen-6 sensors) were performed. The results on these key datasets provide important guidance for their use in applications.

The remainder of this paper is organized into four sections. Section 2 introduces briefly the principles of ATPK-based solutions, including the original ATPK and area-to-point regression kriging (ATPRK), followed by an explicit introduction to our new solution for transformation PSF estimation. The experimental results on several datasets were provided in Section 3. Section 4 further discusses the findings and related issues, as well as openings for future research. Finally, Section 5 concludes the paper.

2. Methods

2.1. Univariate case

ATPK (Kyriakidis, 2004) is adopted for the univariate case of downscaling continua in this paper. ATPK was originally proposed for downscaling data with irregular geographical units, where the sizes and shapes of the observation support vary across geographical locations, such as for population and disease data in different counties (Goovaerts, 2008). In recent years, ATPK was popularized in the remote sensing community (Atkinson, 2013; Wang et al., 2015; Jin et al., 2018), where the data (i.e., images) are composed of regular supports (i.e., pixels with a fixed size and shape). The principle of ATPK in image downscaling is introduced briefly below.

Suppose $Z_C(\mathbf{x}_i)$ is the measurement for coarse pixel C centered at $\mathbf{x}_i$ (i=1,…, M, where M is the number of pixels) in image $\mathbf{Z}$. The downscaling process aims to predict $Z_F(\mathbf{x})$ for all fine pixels F in the image. For a fine pixel centered at $\mathbf{x}_0$, the pixel value predicted by ATPK is a linear combination of the measurements of neighboring coarse pixels

$$\hat{Z}_F(\mathbf{x}_0) = \sum_{i=1}^{N} \lambda_i Z_C(\mathbf{x}_i), \text{ s.t. } \sum_{i=1}^{N} \lambda_i = 1. \quad (1)$$

In Eq. (1), N is the number of coarse neighbors and λ_i is the weight for the ith neighbor centered at $\mathbf{x}_i$. The weights are calculated according to the kriging matrix

$$\begin{bmatrix} \boldsymbol{\gamma}_{CC} & \mathbf{1}^T \\ \mathbf{1} & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda} \\ \theta \end{bmatrix} = \begin{bmatrix} \boldsymbol{\gamma}_{FC} \\ 1 \end{bmatrix} \tag{2}$$

where $\boldsymbol{\gamma}_{CC}=[\gamma_{CC}(\mathbf{s}_{ij}), i=1,...,N,\ j=1,...,N]$ is an $N\times N$ matrix of coarse-to-coarse semivariogram between the coarse pixels centered at $\mathbf{x}_i$ and $\mathbf{x}_j$ ($\mathbf{s}$ denotes the Euclidean distance between the centroids of any two pixels), $\boldsymbol{\gamma}_{FC}=[\gamma_{FC}(\mathbf{s}_{0i}), i=1,...,N]^T$ is an $N\times 1$ vector of fine-to-coarse semivariogram between the fine and coarse pixels centered at $\mathbf{x}_0$ and $\mathbf{x}_i$, $\boldsymbol{\lambda}=[\boldsymbol{\lambda}_i, i=1,...,N]^T$ is an $N\times 1$ vector of weights for the N coarse neighbors, $\mathbf{1}$ is a $1\times N$ vector of ones, and θ is the Lagrange multiplier.

Suppose $h_C(\mathbf{s})$ is the PSF used for transforming the fine spatial resolution data to the coarse spatial resolution. The semivariograms $\gamma_{FC}(\mathbf{s})$ and $\gamma_{CC}(\mathbf{s})$ used in Eq. (2) are calculated by convolving the fine-to-fine semivariogram (denoted as $\gamma_{FF}(\mathbf{s})$) between two fine pixels. All semivariograms $\gamma_{FF}(\mathbf{s})$, $\gamma_{FC}(\mathbf{s})$ and $\gamma_{CC}(\mathbf{s})$ are a function of the distance and direction vector $\mathbf{s}$ (also termed lag in geostatistics), that is, they are not determined by the specific locations of the pixels, but only their lag separation. Moreover, the semivariogram can be viewed as a 2-D image centered at (0, 0), which is composed of values in all directions and at multiple distances. In this respect, $\gamma_{FC}(\mathbf{s})$ and $\gamma_{CC}(\mathbf{s})$ are calculated as

$$\begin{aligned} \gamma_{FC}(\mathbf{s}) &= \gamma_{FF}(\mathbf{s}) * h_C(\mathbf{s}) \\ &= \int_{\mathbf{x}\in V(\mathbf{s})} \gamma_{FF}(\mathbf{x}) \cdot h_C(\mathbf{s}-\mathbf{x}) d\mathbf{x} \\ &= \sum_{x_1=s_1-w_1}^{s_1+w_1} \sum_{x_2=s_2-w_2}^{s_2+w_2} \gamma_{FF}(x_1,x_2) \cdot h_C(s_1-x_1, s_2-x_2) \end{aligned} \tag{3}$$

$$\begin{aligned} \gamma_{CC}(\mathbf{s}) &= \gamma_{FF}(\mathbf{s}) * h_C(\mathbf{s}) * h_C(-\mathbf{s}) \\ &= \gamma_{FC}(\mathbf{s}) * h_C(-\mathbf{s}) \\ &= \int_{\mathbf{x}\in V(\mathbf{s})} \gamma_{FC}(\mathbf{x}) \cdot h_C(\mathbf{x}-\mathbf{s}) d\mathbf{x} \\ &= \sum_{x_1=s_1-w_1}^{s_1+w_1} \sum_{x_2=s_2-w_2}^{s_2+w_2} \gamma_{FC}(x_1,x_2) \cdot h_C(x_1-s_1, x_2-s_2) \end{aligned} \tag{4}$$

in which * is the convolution operator, $V(\mathbf{s})$ is the spatial coverage (with an extent of $2w_1+1$ by $2w_2+1$ fine pixels) of the PSF centered at point $\mathbf{s}=(s_1,s_2)$, and $\mathbf{x}=(x_1,x_2)$ is the center of the fine pixel within the coverage of $V(\mathbf{s})$ in the 2-D semivariogram image $\gamma_{FF}(\mathbf{s})$ or $\gamma_{FC}(\mathbf{s})$.

The fine-to-fine semivariogram $\gamma_{FF}(\mathbf{s})$ is unknown and it can be predicted by deconvolution of the areal semivariogram calculated from the original coarse image (Atkinson et al., 2008; Goovaerts, 2008; Pardo-Iguzquiza et al., 2011; Wang et al., 2015). The optimal solution is determined as the one that, once convolved according to (4), is the same as the known areal semivariogram. More details of the deconvolution strategy used in this paper can be found in Wang et al. (2015).

As pointed out in Kyriakidis (2004), ATPK has the appealing advantage of the perfect coherence property: when the ATPK prediction $\hat{\mathbf{Z}}_F$ is upscaled back to the coarse spatial resolution, it is exactly the same as the original coarse data $\mathbf{Z}_C$:

$$\hat{\mathbf{Z}}_F * h_C = \mathbf{Z}_C . \tag{5}$$

It can be seen from the abovementioned principle of ATPK that the PSF effect in downscaling is accounted for mainly by the semivariogram modeling processes in Eqs. (3) and (4). Different PSFs will lead to different characterizations of the statistical relation between pixels by the semivariograms $\gamma_{FC}(\mathbf{s})$ and $\gamma_{CC}(\mathbf{s})$ in Eq. (2), and further, different characterizations of the contributions of neighbors by the weights in Eq. (2) and finally, different ATPK predictions in Eq. (1). The PSF effect is, thus, accounted for explicitly in ATPK. It should be highlighted that the implementation of ATPK is not influenced by the specific form of PSF, that is, ATPK is suitable for any PSF in theory. This is different from the commonly used interpolation methods such as bicubic or polynomial interpolation, where the PSF is not considered explicitly in the downscaling process.

2.2. Multivariate case

For the multivariate case of downscaling continua, the downscaling process makes use of ancillary fine spatial resolution data, that is, data covering the same area, but acquired in different wavelengths or by different sensors. This issue is known widely as multi-resolution image fusion. The recently developed ATPRK (Wang et al., 2015) is a method for multi-resolution image fusion. ATPRK is an extension to ATPK, which incorporates fine spatial resolution data by a regression process before ATPK. Inheriting the advantages of ATPK, it can account for the PSF effect. Thus, ATPRK is adopted for the multivariate case in this paper. Its principle is introduced briefly in the following.

Without loss of generality, we consider the case that multiple fine spatial resolution images are available. Denote $Z_C^l(\mathbf{x}_i)$ as the measurement of pixel *C* at $\mathbf{x}_i$ in band *l* with a coarse spatial resolution, and $Z_F^k(\mathbf{x}_j)$ as the measurement of pixel *F* at $\mathbf{x}_j$ (j=1,…, MG^2, where *G* is the spatial resolution ratio) in band *k* with a fine spatial resolution (k=1,…, *K*, where *K* is the number of fine bands). The goal of downscaling is to predict $Z_F^l(\mathbf{x})$ for all fine pixels in all coarse bands. In ATPRK, the prediction is accomplished by two steps: regression modeling and residual downscaling (Wang et al., 2015). Suppose the predictions of the two steps are $\hat{Z}^l_{\text{Regression}}(\mathbf{x})$ and $\hat{Z}^l_{\text{Residual}}(\mathbf{x})$, the ATPRK prediction is calculated as

$$\hat{Z}_F^l(\mathbf{x}) = \hat{Z}^l_{\text{Regression}}(\mathbf{x}) + \hat{Z}^l_{\text{Residual}}(\mathbf{x}). \tag{6}$$

The objective of the regression modeling step is to obtain the overall trend of the predictions at the fine spatial resolution. This is achieved by a linear transformation of the existing *K* fine bands acquired in other wavelength

$$\hat{Z}^l_{\text{Regression}}(\mathbf{x}_0) = \sum_{k=1}^{K} a_k Z_F^k(\mathbf{x}_0) + b \tag{7}$$

where a_k is the weighting coefficient for fine band *k* and *b* is a constant. The *K*+1 coefficients are determined according to the regression model constructed between the coarse band and the *K* upscaled bands

$$\mathbf{Z}_C^l = \sum_{k=1}^{K} a_k \mathbf{Z}_C^k + b + \mathbf{R}_C^l. \tag{8}$$

In Eq. (8), $\mathbf{R}_C^l$ is the coarse residual image, and $\mathbf{Z}_C^k$ is the coarse image produced by upscaling the corresponding fine band *k* using a PSF

$$\mathbf{Z}_C^k = \mathbf{Z}_F^k * h_C^l \tag{9}$$

where h_C^l is the transformation PSF between the fine and coarse spatial resolution data. The coefficients a_k and *b* are predicted using the least squares-based fitting method.

In the second step of ATPRK, the coarse residual image $\mathbf{R}_C^l$ produced in the regression model in the first step is downscaled to a fine spatial resolution using the ATPK method

$$\hat{Z}_{\text{Residual}}^l(\mathbf{x}_0) = \sum_{i=1}^{N} \lambda_i \mathbf{R}_C^l(\mathbf{x}_i), \text{ s.t. } \sum_{i=1}^{N} \lambda_i = 1 \tag{10}$$

where $\mathbf{R}_C^l(\mathbf{x}_i)$ is the residual for the *i*th neighbor. Similarly to Eq. (2), the weights are calculated based on the kriging matrix. It should be noted that here the corresponding semivariograms are modeled based on the residuals, rather than the original measurements; that is, the common case in Eqs. (1) and (2).

ATPRK inherits the appealing advantage of ATPK, that is, the perfect coherence property, which was demonstrated mathematically in Wang et al. (2015)

$$\hat{\mathbf{Z}}_F^l * h_C = \mathbf{Z}_C^l . \tag{11}$$

As seen from the principles of ATPRK, the PSF effect is considered in both regression modeling (see Eq. (9)) and ATPK-based residual downscaling. Regarding ATPK-based residual downscaling, the PSF effect was analyzed in Section 2.1. With respect to the regression modeling step, different PSFs will result in different upscaled images $\mathbf{Z}_C^k$ and different regression models (i.e., different estimates of coefficients a_k and *b*, that is, different estimates of $\hat{\mathbf{Z}}_{\text{Regression}}^l$). Moreover, different residual images $\mathbf{R}_C^l$ in the constructed regression model in Eq. (8) will be produced if the PSF is different. This will influence directly the prediction of $\hat{Z}_{\text{Residual}}^l(\mathbf{x})$ in ATPK-based residual downscaling.

2.3. The proposed solution for estimation of the PSF in the multivariate case

From the abovementioned ATPK and ATPRK-based downscaling methods, it is seen that the transformation PSF between the images at different spatial resolutions, rather than the PSF of the original measurements, plays an important role in semivariogram modeling at different scales and cannot be ignored. In practice, however, this PSF may be unknown for downscaling. Thus, it is crucial to determine the exact PSF in advance. As mentioned in the Introduction, PSF estimation for the univariate case is always challenging and remains an open problem. However, it is more realistic to develop solutions to estimate the effective PSF for the multivariate case, as auxiliary data at the target fine spatial resolution are available.

As widely acknowledged, there always exists a statistical relation between different bands covering the same area, especially for bands with proximate in wavelength. It is assumed that the single coarse band $\mathbf{Z}_C^l$ can be characterized as a linear combination of the set of bands $\mathbf{Z}_C^k$ (k=1,…, K) produced by upscaling the existing fine bands $\mathbf{Z}_F^k$ (k=1,…, K) at other wavelengths

$$\mathbf{Z}_C^l = \sum_k \alpha_k \left(\mathbf{Z}_F^k * h_C^l \right) + \boldsymbol{\varepsilon} . \tag{12}$$

Thus, the optimal scale transformation PSF h_C^l for band $\mathbf{Z}_C^l$ can be determined as the one minimizing the fitting error $\boldsymbol{\varepsilon}$

$$\hat{h}_C^l = \arg\min_{h_C^l} f \left(\mathbf{Z}_C^l - \sum_k \alpha_k \left(\mathbf{Z}_F^k * h_C^l \right) \right) \tag{13}$$

where f is an operator for calculating the fitting error and it is determined by the fitting model (e.g., the classical global linear regression model, local weighted regression model, etc.). The optimal PSF h_C^l is determined for each coarse band separately, which may vary for different coarse bands.

In this paper, the PSF in scale transformation is assumed to be a Gaussian filter which has been widely used in remote sensing (Townshend et al., 2000; Huang et al., 2002; Manslow and Nixon, 2002; Kaiser and Schneider, 2008; Van der Meer, 2012). The classical global linear regression model is used as the fitting model

and the fitting error is calculated based on the least squares method, which amounts to maximizing the correlation coefficient (CC) of the fitting model. Once the type of filter is determined, the problem of PSF estimation is essentially the estimation of the scale (width) parameter.

The implementation of the proposed solution is illustrated by an example for fusion of 10 m and 20 m Sentinel-2 images below. The Sentinel-2 satellite sensor acquires four 10 m bands and six 20 m bands (Drusch et al., 2012). The task of downscaling is to fuse the two types of bands to produce the 10-band 10 m data.

1) All fine spatial resolution bands (e.g., four 10 m Sentinel-2 bands) are convolved with a Gaussian PSF (with scale parameter σ_i) and upscaled to the coarse spatial resolution (e.g., 20 m).
2) For each coarse band, a linear regression model is fitted between the multiple upscaled images (e.g., four 20 m Sentinel-2 images) and the observed coarse image (e.g., one of the six observed 20 m Sentinel-2 images). The CC is calculated.
3) Step 2) is conducted for all parameter candidates of σ. For the visited coarse band, the optimal σ is estimated as the one leading to the largest CC in 2).
4) Steps 2) and 3) are conducted for all observed coarse bands, in turn.

If we have the prior information that all coarse bands have the same PSF, the solution mentioned above can also be modified by assuming that all coarse bands share the same PSF parameter σ. In this case, for a given candidate σ_i, the CCs for all coarse bands can be averaged to produce a mean CC, and the optimal σ is determined as the one leading to the largest mean CC.

Generally, the proposed solution for PSF estimation in the multivariate case holds several advantages.

1) It is a general method applicable for the case involving multiple fine bands (e.g., downscaling a 20 m Sentinel-2 band using four 10 m Sentinel-2 bands), not only for the conventional pan-sharpening case with only one fine band.
2) It releases the assumption that all coarse bands have the same PSF, and can be performed for each coarse band separately, that is, predicting a specific PSF for each coarse band.

3. Experiments

3.1. Univariate case

Three WorldView-2 images were used in the experiment for the univariate case, as shown in Fig. 1. All images were acquired in April, 2011. They contain eight multispectral bands with a spatial resolution of 2 m and the spatial extent of each study area is 500 by 500 pixels (e.g., 1 km by 1 km). The first and second images cover two urban areas in Shenzhen, China (hereafter, called Region 1) and the third image covers an urban area in Hong Kong, China (hereafter, called Region 3).

The images were degraded with a PSF filter and a zoom factor of *S*. The Gaussian filter with a width of 0.5 pixel size was used, and two zoom factors (S=2 and 4) were considered. The task of downscaling is to restore the 2 m images from the input 4 m (or 8 m) coarse images with a zoom factor of 2 (or 4). Using this strategy, the reference of the fine spatial resolution images (i.e., 2 m images) are known perfectly and the evaluation can be performed objectively.

Fig. 2 shows the 2 m downscaling results for the 4 m coarse images in the three regions. The ATPK method with ideal square wave filter (i.e., no PSF) and Gaussian PSF (width is 0.5 coarse pixel size) produced the results in the second and third columns. It is clear that by considering the PSF effect, the downscaling predictions are more accurate. For example, the buildings in Fig. 2(c) are clearer than those in Fig. 2(b), and closer to the reference in Fig. 1(a). The downscaling results for the 8 m coarse images are shown in Fig. 3, where ATPK with a PSF also produces more accurate downscaling results than without the PSF.

The CC, universal image quality index (UIQI), relative global-dimensional synthesis error (ERGAS) and spectral angle mapper (SAM) were used for quantitative evaluation of the predictions. For CC and UIQI, they were first calculated for each band and the values for all bands were then averaged. For the SAM, its value was calculated for each pixel first and the values for all pixels were then averaged. To evaluate the ability to honour

the original coarse data, an index called coherence was also used. It was calculated by upscaling the prediction to the original coarse spatial resolution and then comparing to the input coarse image based on CC.

Table 1 shows the quantitative accuracy assessment for the three WorldView-2 images. The widely used bicubic and polynomial (with 23 coefficients) interpolation methods were also tested. Generally, four observations can be made.

First, the accuracies of all downscaling methods decrease when the zoom factor is larger. For example, the CCs of all methods decrease by over 0.06 from S=2 to S=4, and the corresponding SAMs increase by around 0.02. This is because the downscaling issue becomes more complex and the uncertainty increases, as the values of more sub-pixels need to be predicted.

Second, the two ATPK-based methods can produce more accurate results than the bicubic and polynomial interpolation methods in terms of all indices in both cases. For example, for Region 1, the UIQIs of ATPK are at least 0.01 larger than the two benchmark methods for the case of S=2, and the differences increase to be at least 0.025 for the case of S=4. For Region 3, the increase in UIQIs can exceed 0.02 for S=2. The more accurate results reveal the advantages of ATPK in downscaling.

Third, the ATPK with PSF method produces the more accurate results than ATPK without PSF for both zoom factors. More precisely, for Region 1, when considering the PSF, the CCs increase by 0.015 and 0.017 for S=2 and 4, respectively, while the corresponding UIQIs increase by 0.025 and 0.040, respectively. For Region 3, the increase in CCs is 0.018 and 0.022 for S=2 and 4 when the PSF is considered.

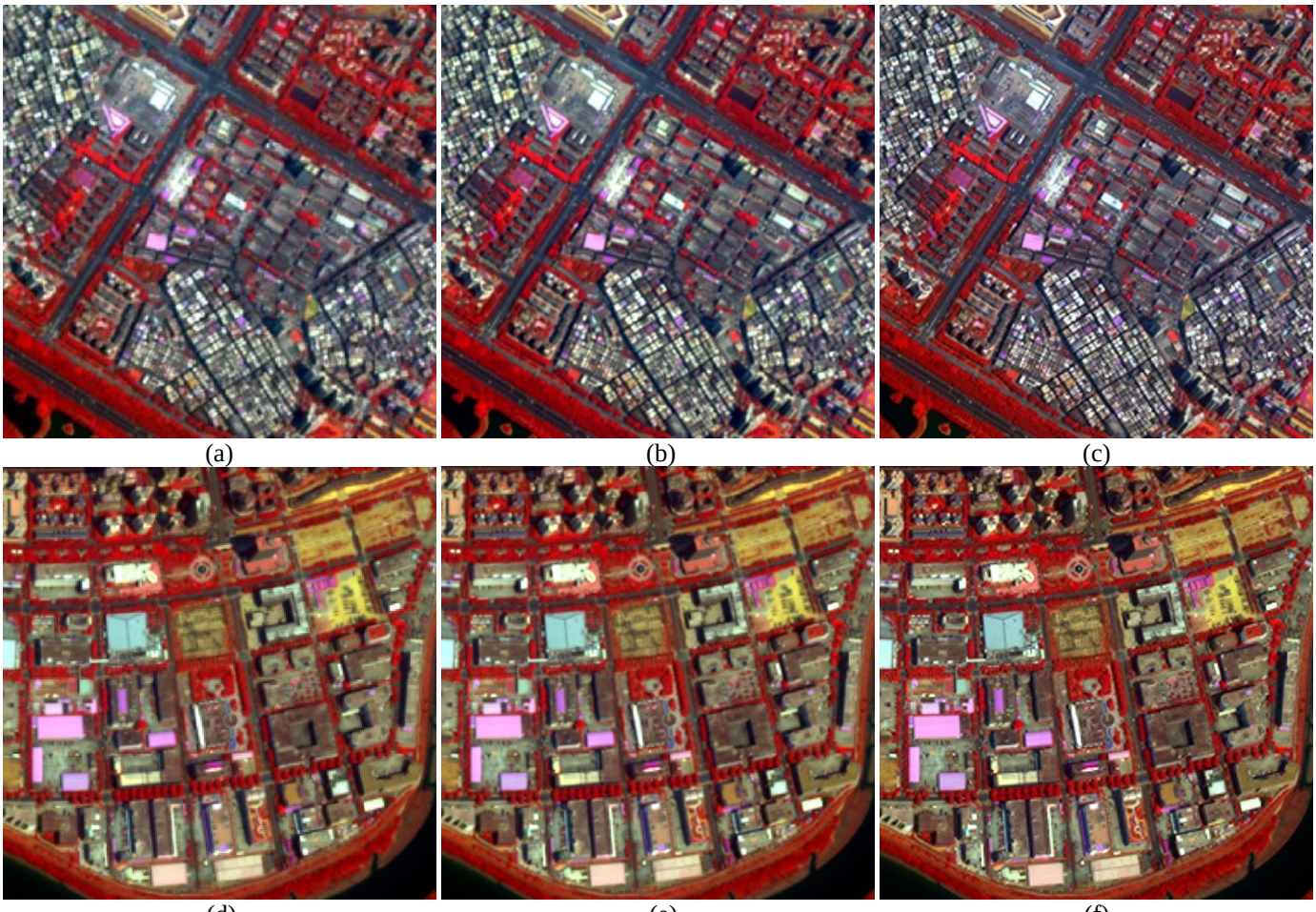
Fourth, the ATPK with PSF method produces almost an ideal coherence value of 1 for both cases, honouring the perfect coherence property of ATPK, as mentioned in Section 2 and demonstrated in Kyriakidis (2004).

The benefit of considering the PSF effect is also illustrated based on the performance of land cover classification. Specifically, k-means based unsupervised classification was performed on the downscaling predictions as well as the 2 m reference images in Fig. 1. The land cover maps of the former were compared with those of the latter to evaluate the classification accuracy (i.e., by calculating the overall accuracy (OA)).

The OAs for the three regions are listed in Table 2. It can be concluded from the results that by considering the PSF effect, the land cover maps of the downscaling predictions are closer to the reference land cover maps.

(a) (b) (c)

Fig. 1. The three WorldView-2 images (bands 8, 4 and 1 as RGB). (a) Region 1. (b) Region 2. (c) Region 3.

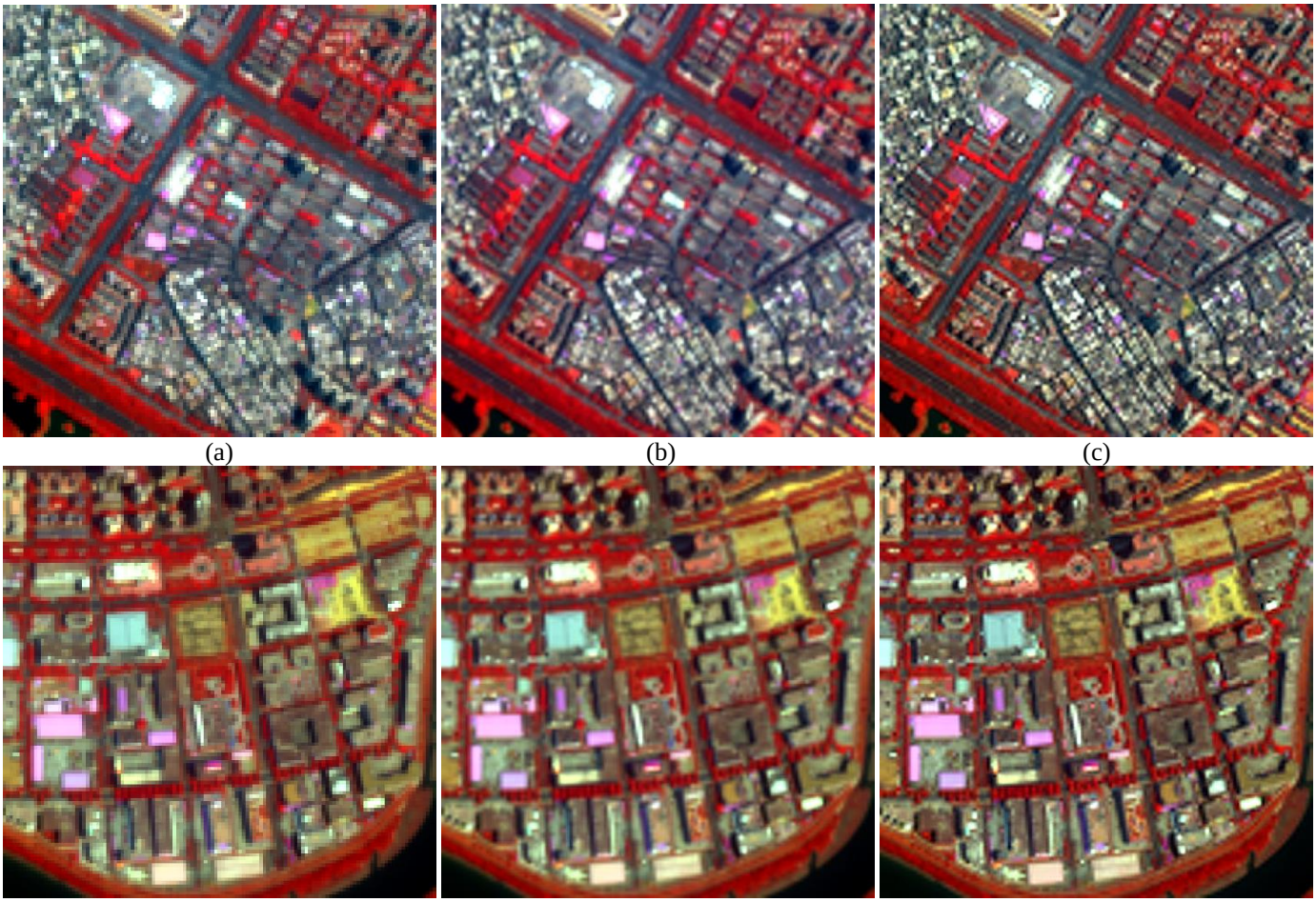


(a) (b) (c)

(d) (e) (f)

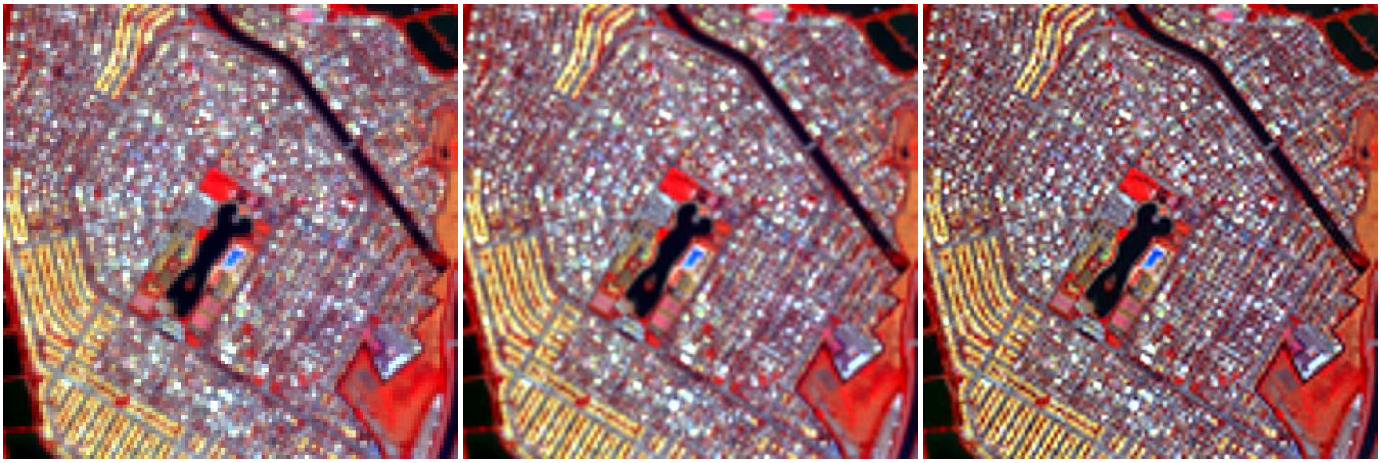
(g) (h) (i)

Fig. 2. Downscaling results (2 m) for the 4 m coarse WorldView-2 images (S=2, bands 8, 4 and 1 as RGB) simulated with Gaussian PSF. (a), (d) and (g) 4 m coarse image. (b), (e) and (h) ATPK without PSF. (c), (f) and (i) ATPK with PSF. (a)-(c) Results for Region 1. (d)-(f) Results for Region 2. (g)-(i) Results for Region 3.



(a) (b) (c)

(d) (e) (f)



(g) (h) (i)

Fig. 3. Downscaling results (2 m) for the 8 m coarse WorldView-2 images (*S*=4, bands 8, 4 and 1 as RGB) simulated with Gaussian PSF. (a), (d) and (g) 8 m coarse image. (b), (e) and (h) ATPK without PSF. (c), (f) and (i) ATPK with PSF. (a)-(c) Results for Region 1. (d)-(f) Results for Region 2. (g)-(i) Results for Region 3.

Table 1 Quantitative assessment for the downscaling results for the three WorldView-2 images (the Gaussian PSF case)

			CC	ERGAS	UIQI	SAM	Coherence
		Ideal	1	0	1	0	1
S=2	Region 1	Bicubic	0.9329	4.0250	0.9179	0.0334	0.9886
		Polynomial	0.9358	3.9362	0.9218	0.0328	0.9882
		ATPK (noPSF)	0.9436	3.6724	0.9334	0.0306	0.9934
		ATPK (PSF)	**0.9585**	**3.0665**	**0.9579**	**0.0264**	**0.9993**
	Region 2	Bicubic	0.9587	3.1108	0.9534	0.0259	0.9938
		Polynomial	0.9598	3.0661	0.9549	0.0257	0.9934
		ATPK (noPSF)	0.9647	2.8720	0.9610	0.0240	0.9964
		ATPK (PSF)	**0.9721**	**2.5158**	**0.9716**	**0.0213**	**0.9995**
	Region 3	Bicubic	0.9091	4.2231	0.8832	0.0355	0.9846
		Polynomial	0.9130	4.1357	0.8890	0.0349	0.9841
		ATPK (noPSF)	0.9225	3.8769	0.9046	0.0328	0.9911
		ATPK (PSF)	**0.9401**	**3.3020**	**0.9387**	**0.0288**	**0.9988**
S=4	Region 1	Bicubic	0.8265	3.1017	0.7808	0.0519	0.9831
		Polynomial	0.8252	3.0988	0.7831	0.0517	0.9805
		ATPK (noPSF)	0.8435	2.9440	0.8097	0.0488	0.9910
		ATPK (PSF)	**0.8606**	**2.7502**	**0.8501**	**0.0453**	**0.9986**
	Region 2	Bicubic	0.8977	2.3855	0.8796	0.0405	0.9905
		Polynomial	0.8950	2.4047	0.8783	0.0408	0.9877
		ATPK (noPSF)	0.9073	2.2675	0.8944	0.0382	0.9949
		ATPK (PSF)	**0.9161**	**2.1413**	**0.9120**	**0.0360**	**0.9989**
	Region 3	Bicubic	0.7658	3.2142	0.6944	0.0542	0.9759
		Polynomial	0.7653	3.2061	0.6991	0.0543	0.9734
		ATPK (noPSF)	0.7883	3.0581	0.7324	0.0514	0.9867
		ATPK (PSF)	**0.8104**	**2.8622**	**0.7892**	**0.0485**	**0.9977**

Table 2 Classification accuracies (in terms of OA) of the downscaling results for the three WorldView-2 images

	Region 1	Region 2	Region 3

	ATPK (noPSF)	ATPK (PSF)	ATPK (noPSF)	ATPK (PSF)	ATPK (noPSF)	ATPK (PSF)
S=2	79.62%	**82.70%**	83.68%	**86.20%**	76.73%	**80.50%**
S=4	66.40%	**68.23%**	73.25%	**75.12%**	42.16%	**51.53%**

Fig. 4 shows the semivariograms of all eight bands for the ATPK-based predictions for Region 1. It is clear that for each band, when the PSF is considered, the semivariograms at all lags are larger, indicating the prediction reproduces more spatial variance.

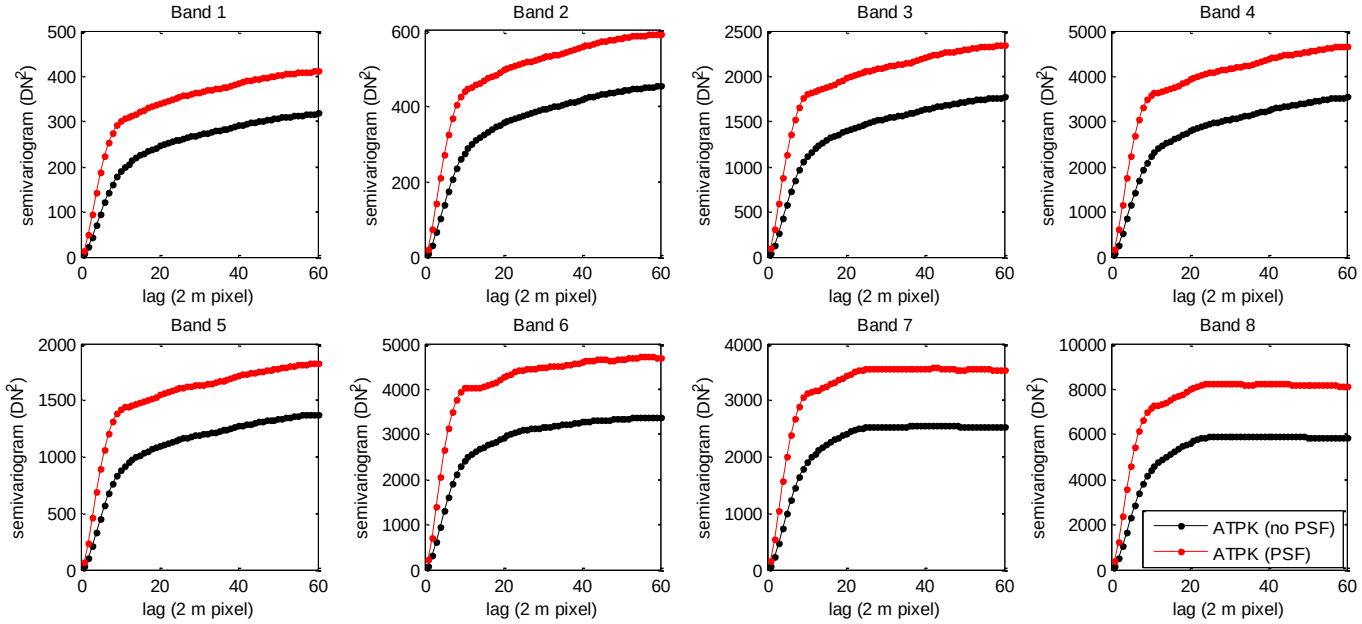


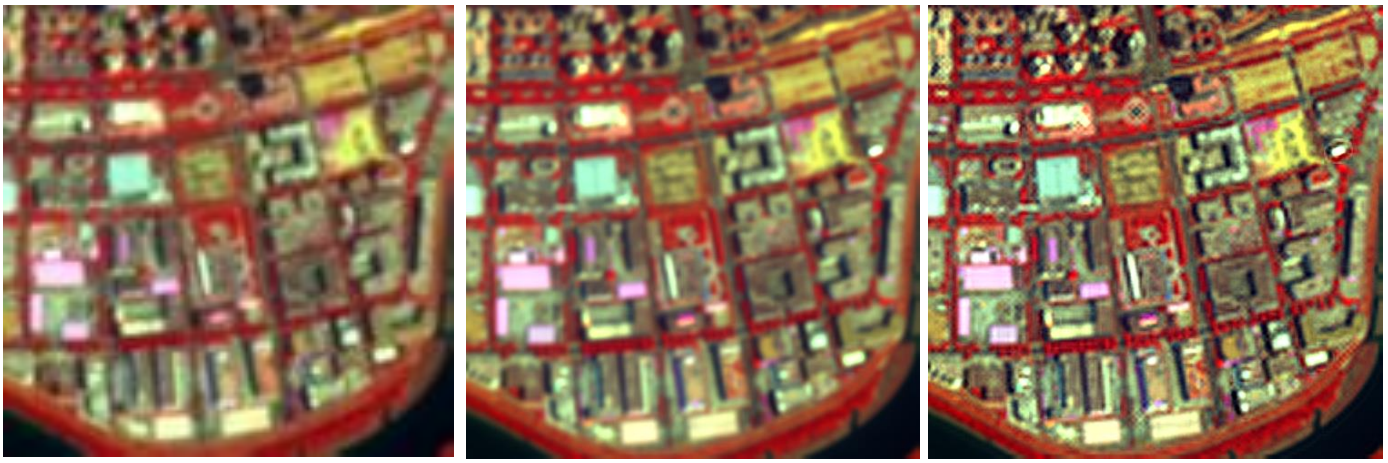
Fig. 4. Semivariograms of the ATPK-based downscaling predictions for the WorldView-2 image simulated with Gaussian PSF in Region 1 (*S*=4).

A non-Gaussian PSF was also considered to examine the performance of ATPK-based downscaling. Assuming the PSF to be fairly flat across the center, with the curvature occurring at the edges, the PSF (hereafter, called the center flat PSF) is expressed as

$$h_C(\mathbf{x}) = \begin{cases} \dfrac{1}{2\pi\sigma^2}, & \text{if } |x_1|,|x_2| < 0.5S \text{ and } \mathbf{x} \in V(\mathbf{x}_0) \\ \dfrac{1}{2\pi\sigma^2}\exp\left[-\left(\dfrac{(x_1-0.5S)^2+(x_2-0.5S)^2}{2\sigma^2}\right)\right], & \text{if } |x_1|,|x_2| > 0.5S \text{ and } \mathbf{x} \in V(\mathbf{x}_0) \\ 0, & \mathbf{x} \notin V(\mathbf{x}_0) \end{cases} \quad (14)$$

where S is the zoom factor and $V(\mathbf{x}_0)$ is the spatial coverage of the PSF centered at point $\mathbf{x}_0$.

Similarly, the three WorldView-2 images in Fig. 1 were degraded with the center flat PSF filter (σ=0.5) and the zoom factor of *S*=4 was considered. The results for Region 2 are shown in Fig. 5. Noticeably, the prediction produced by considering the PSF is visually clearer than that produced without considering the PSF, and the former is much closer to the reference in Fig. 1(b). Table 3 lists the accuracy assessment results for all three regions. The advantage of ATPK with PSF is obvious when compared with the benchmark methods. Therefore, the applicability of ATPK is not affected by the specific form of PSF.



(a) (b) (c)

Fig. 5. Downscaling results (2 m) for the 8 m coarse WorldView-2 image in Region 2 (*S*=4, bands 8, 4 and 1 as RGB) simulated with the center flat PSF. (a) 8 m coarse image. (b) ATPK without PSF. (c) ATPK with PSF.

Table 3 Quantitative assessment for the downscaling results for the three WorldView-2 images (*S*=4, the center flat PSF case)

		CC	ERGAS	UIQI	SAM	Coherence
	Ideal	1	0	1	0	1
Region 1	Bicubic	0.7763	3.4598	0.7131	0.0586	0.9940
	Polynomial	0.7763	3.4550	0.7148	0.0585	0.9914
	ATPK (noPSF)	0.7914	3.3559	0.7350	0.0566	0.9889
	ATPK (PSF)	**0.8391**	**2.9393**	**0.8255**	**0.0490**	**0.9992**
Region 2	Bicubic	0.8680	2.6836	0.8401	0.0462	0.9963
	Polynomial	0.8669	2.6894	0.8399	0.0463	0.9938
	ATPK (noPSF)	0.8783	2.5843	0.8544	0.0442	0.9932
	ATPK (PSF)	**0.9065**	**2.2547**	**0.9018**	**0.0385**	**0.9995**
Region 3	Bicubic	0.6982	3.5677	0.6088	0.0607	0.9927
	Polynomial	0.7003	3.5560	0.6127	0.0607	0.9903
	ATPK (noPSF)	0.7174	3.4684	0.6346	0.0588	0.9866
	ATPK (PSF)	**0.7802**	**3.0467**	**0.7532**	**0.0527**	**0.9989**

3.2. Multivariate case

1) Validation of the solution to PSF estimation. The solution to PSF parameter estimation was evaluated using a hyperspectral image (HSI). The HSI image contains 100 bands and covers a 6 km by 6 km area in Moffett Field, as shown in Fig. 6. The spatial resolution is 20 m and the spatial size of the image is 300 by 300 pixels. A four-band 20 m multispectral image was simulated by averaging 12 consecutive bands of the first 48 bands. For objective evaluation where the reference PSF needs to be known perfectly, simulated coarse images contaminated by different PSFs were used. Specifically, a set of coarse 100-band HSI images was simulated by convolving the 20 m HSI with a Gaussian PSF (with different parameters) and different degradation factors. More precisely, four PSF parameters (σ =0.2, 0.4, 0.6 and 0.8 coarse pixel) and four degradation factors (S=2, 3, 4 and 5) were considered.

Fig. 6. The 20 m Moffett Field HSI image (bands 33, 15, and 4 as RGB).

The proposed method used the simulated four-band 20 m multispectral image (i.e., fine bands) for PSF estimation of the simulated coarse bands. It was performed for all 16 cases and the results are shown in Fig. 7. The optimal parameter in each case was selected in the pool composed of 10 candidates (0.1, 0.2, …, and 1 coarse pixel). The optimal one was determined as the one maximizing the CC in the y-axis; see the values marked by a red asterisk in Fig. 7. From the curves in Fig. 7, it is seen clearly that the proposed method can find the correct parameter consistently in all cases.

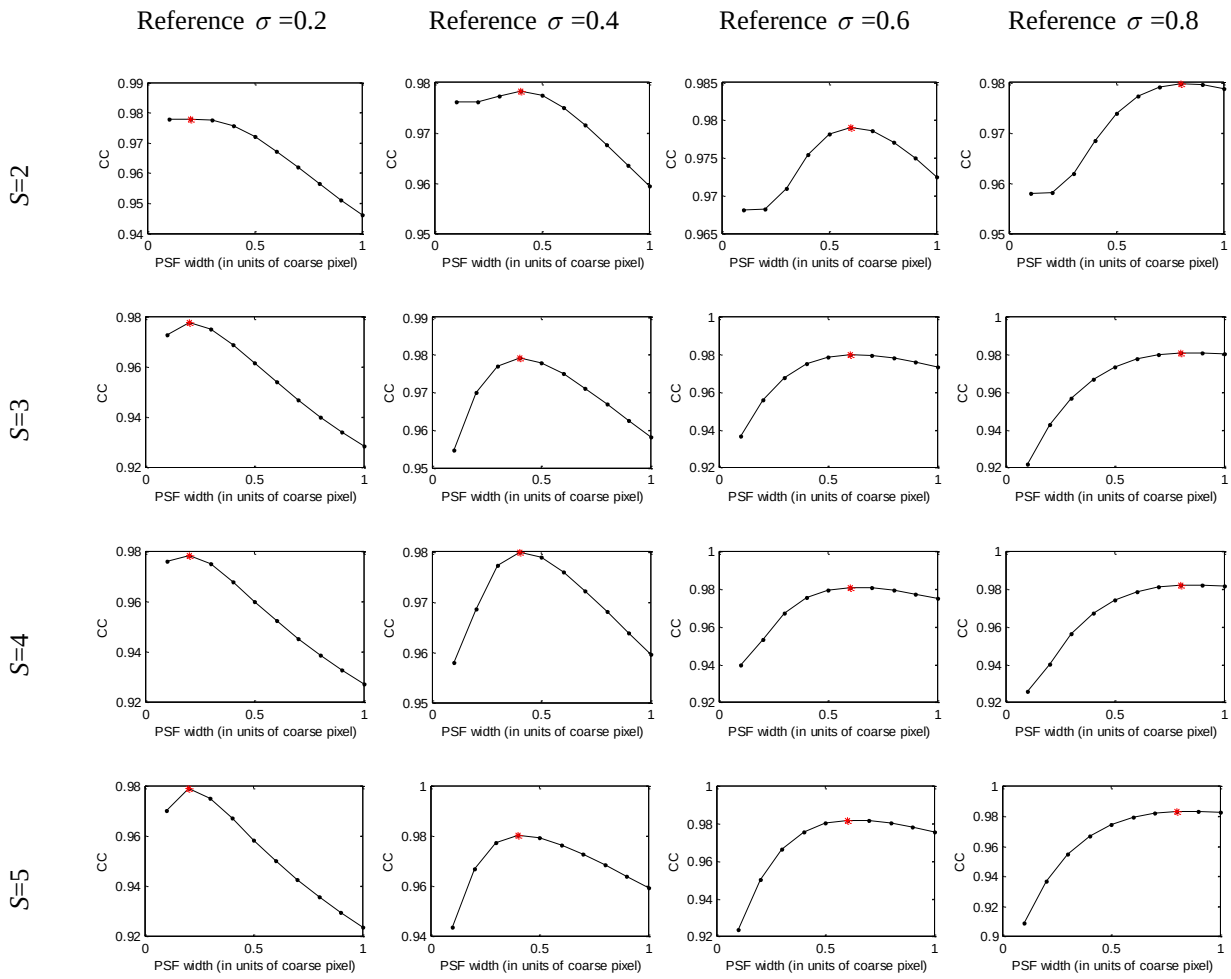


Fig. 7. Plots of CC (between predicted and observed coarse data) against PSF width used for PSF estimation for the HSI image (the red asterisk represents the estimated optimal PSF parameter in each case).

2) Application to Sentinel-2 image fusion. The proposed method was applied for fusion of two Sentinel-2 images. The first image was acquired on 18 August 2015 and covers a 4 km by 4 km urban area in Verona, Italy. The second image was acquired on 24 May 2019 and covers a 4 km by 4 km urban area in Shanghai, China. The task of this experiment is to downscale the six 20 m bands (bands 5, 6, 7, 8a, 11 and 12) to 10 m bands by fusing with the four observed 10 m bands (bands 2, 3, 4 and 8).

The Gaussian PSF was considered for the two Sentinel-2 images. From Fig. 8, for the Verona Sentinel-2 image, the optimal PSF parameters were determined as σ=0.4, 0.4, 0.4, 0.4, 0.6 and 0.6 coarse pixels for bands 5, 6, 7, 8a, 11 and 12, respectively. According to Fig. 9, the corresponding PSF parameters were determined as σ=0.4, 0.4, 0.4, 0.5, 0.6, 0.6 coarse pixels. The results suggest that the effective PSF for

different coarse bands can vary and it is important to estimate the PSF separately for each coarse band. Moreover, the optimal CC values in most cases are above 0.97, suggesting that a very large correlation exists between 10 m and 20 m bands and further, supporting the rational of the proposed PSF estimation method. The 10 m downscaling results for PSF parameters of 0.2 and 0.8 (the parameter is the same for all six bands) for the two Sentinel-2 images are shown in Fig. 10 for visual validation of the selected parameter set. Compared with the case of 0.2, the optimal parameter set can lead to downscaling results containing more spatial details, where the urban fabric (e.g., roads and buildings) and vegetation features can be more clearly observed. For the parameter of 0.8, it produces results with over-highlighted boundaries between classes, and are clearly too noisy. The results reveal that the appropriate PSF parameter is critical for the performance of multi-resolution image fusion (i.e., the multivariate case for downscaling continua).

The ATPRK without PSF method was also performed and the results are exhibited in Fig. 11(b) and Fig. 11(e) for visual comparison with the result of ATPRK with PSF (based on the estimated optimal PSF parameter set) in Fig. 11(c) and Fig. 11(f). Obviously, the image is visually more satisfactory when the PSF is considered, where the spatial detail can be observed more clearly (see, for example, the marked sub-areas).

Since the 10 m reference data for the six Sentinel-2 bands are not available in reality, quantitative evaluation cannot be performed using indices in Table 1. As shown in Fig. 4, however, the semivariogram characterizes the spatial variance of the image at various lags and thus, can reflect the gains of the downscaling prediction. More importantly, it is calculated without the need for any reference data. Thus, the semivariogram was used to quantitatively evaluate the downscaling predictions here. The results of both ATPRK-based methods are shown in Figs. 12 and 13. It is observed clearly that for all bands, the semivariograms are larger by considering the PSF effect, suggesting the benefit of considering the PSF. More precisely, by considering the PSF effect, the global (also known as the *a priori*) variances (i.e., the value of the semivariogram for the infinitely great lag) for bands 5, 6, 7, 8a, 11 and 12 of the Verona dataset are increased by 21.13%, 15.26%, 13.32%, 11.46%, 44.22% and 49.01%, respectively.

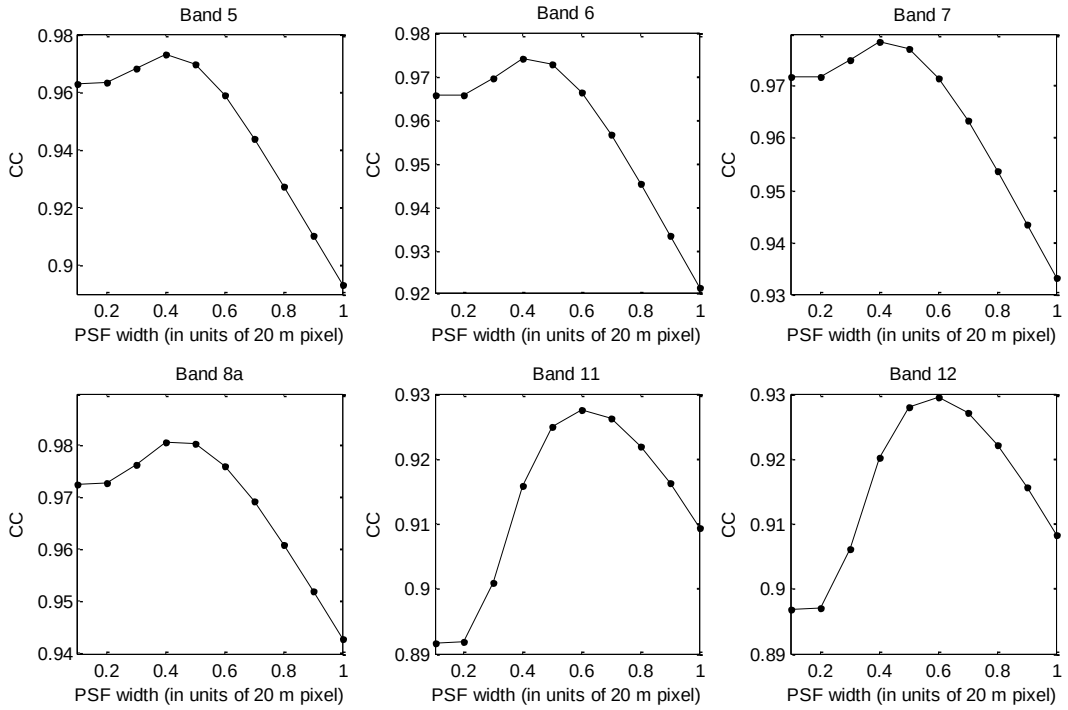


Fig. 8. Plots of CC (between predicted and observed coarse data) against PSF width used for PSF estimation for the Sentinel-2 image in Verona, Italy.

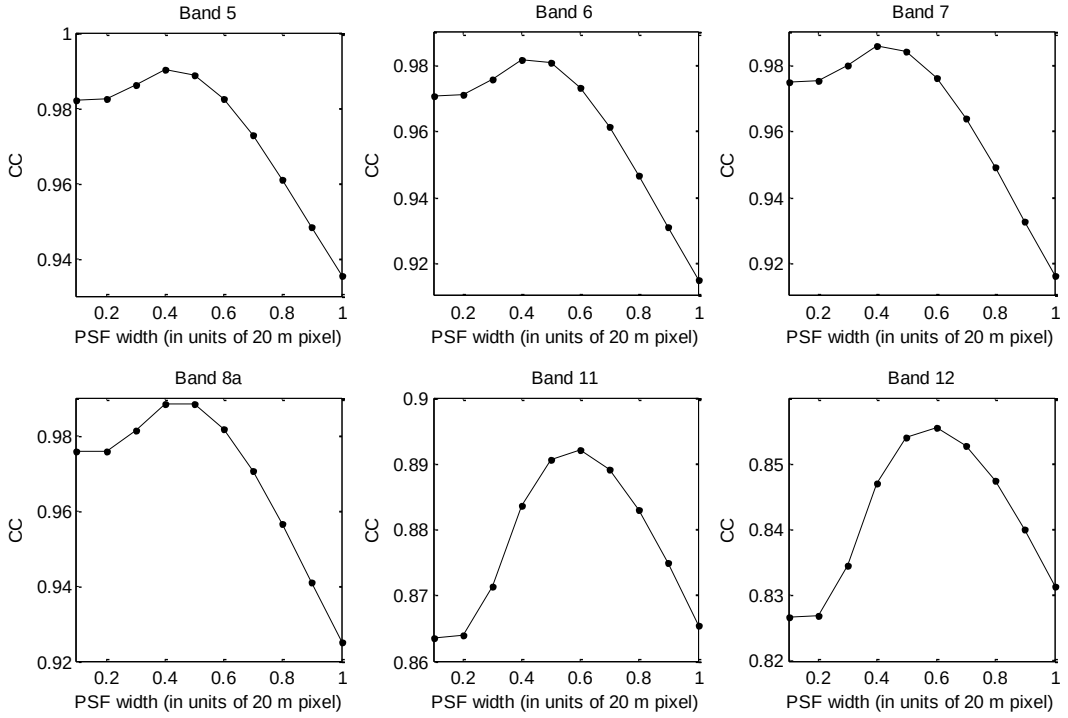
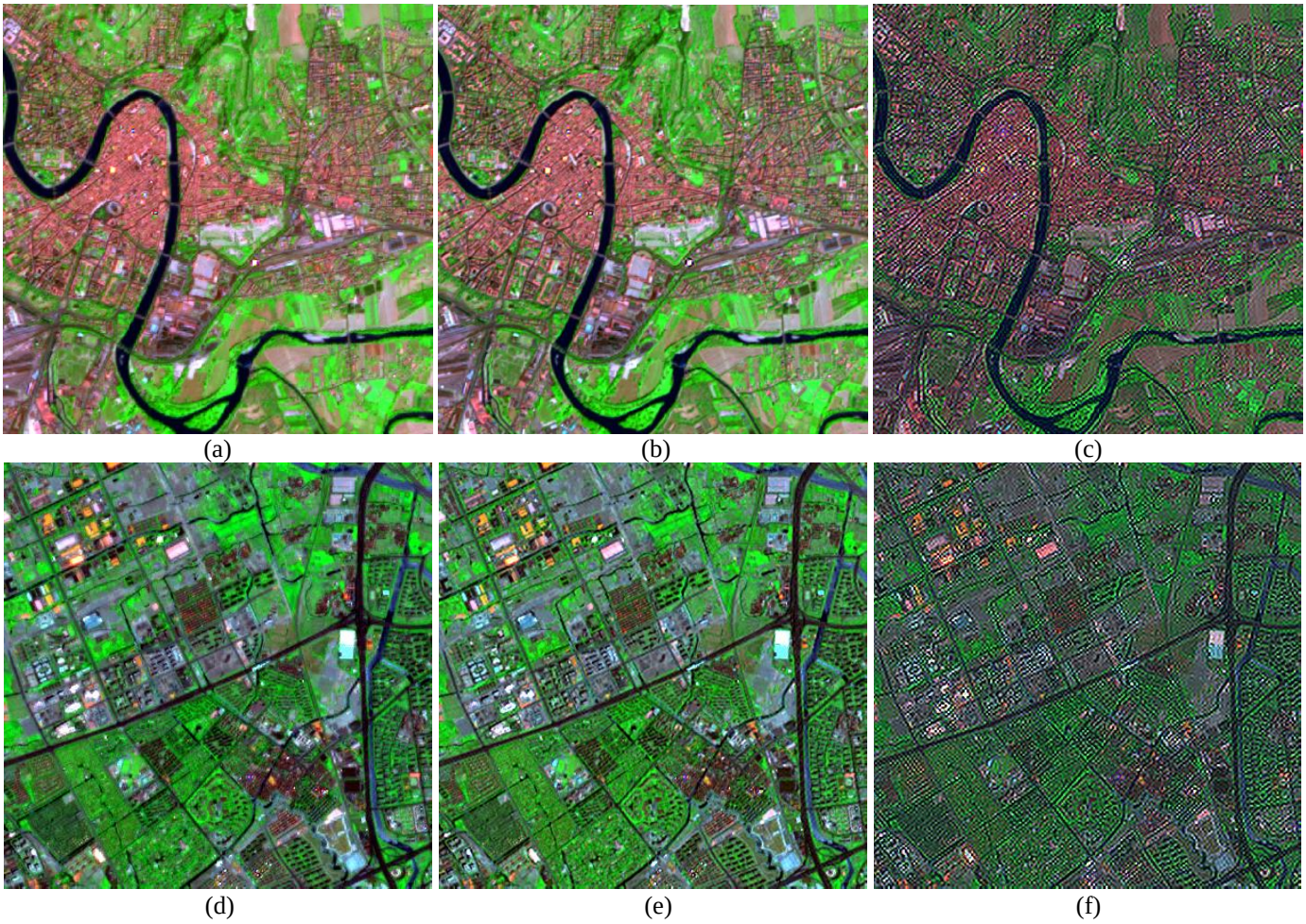


Fig. 9. Plots of CC (between predicted and observed coarse data) against PSF width used for PSF estimation for the Sentinel-2 image in Shanghai, China.



(a) (b) (c)

(d) (e) (f)

Fig. 10. Downscaling result (10 m) for different PSF parameters (bands 12, 8a and 5 as RGB). (a) and (d) σ =0.2 for all six bands. (b) and (e) σ = estimated optimal parameters. (c) and (f) σ =0.8 for all six bands. (a)-(c) Results for the Sentinel-2 image in Verona, Italy. (d)-(f) Results for the Sentinel-2 image in Shanghai, China.

(a) (b) (c)

(d) (e) (f)

Fig. 11. Downscaling result (10 m) for the two Sentinel-2 images (bands 12, 8a and 5 as RGB). (a) and (d) 20 m coarse image. (b) and (e) ATPRK without PSF. (c) and (f) ATPRK with PSF. (a)-(c) Results for the Sentinel-2 image in Verona, Italy. (d)-(f) Results for the Sentinel-2 image in Shanghai, China.

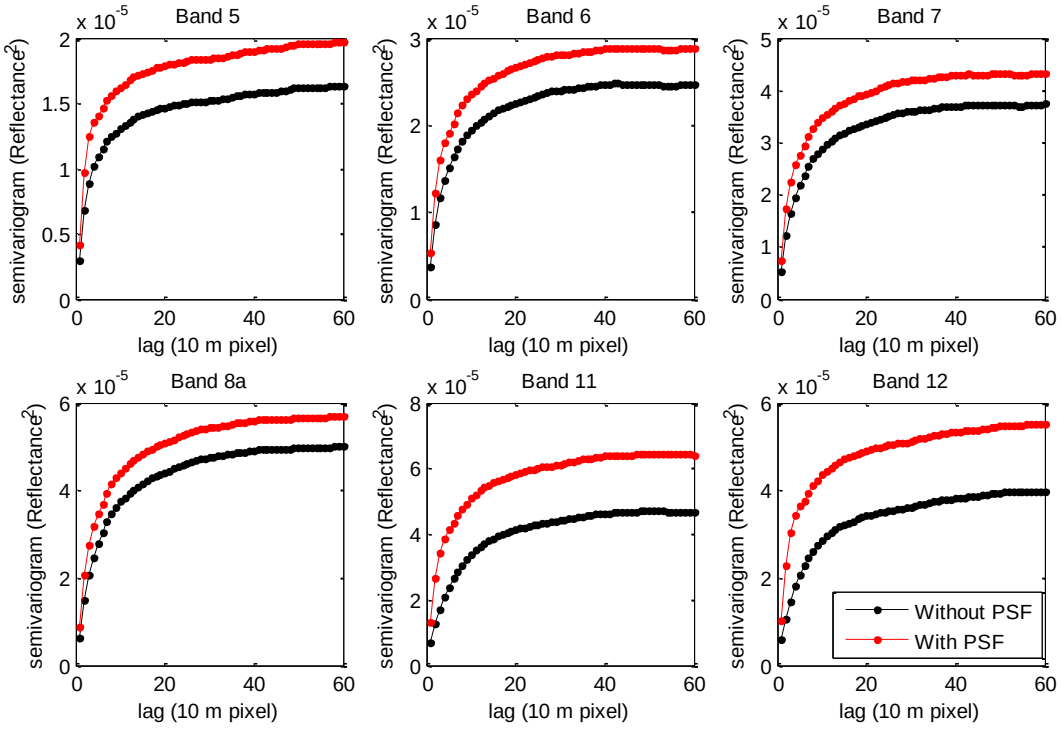


Fig. 12. Semivariograms of the ATPRK-based downscaling predictions for the Sentinel-2 image in Verona, Italy.

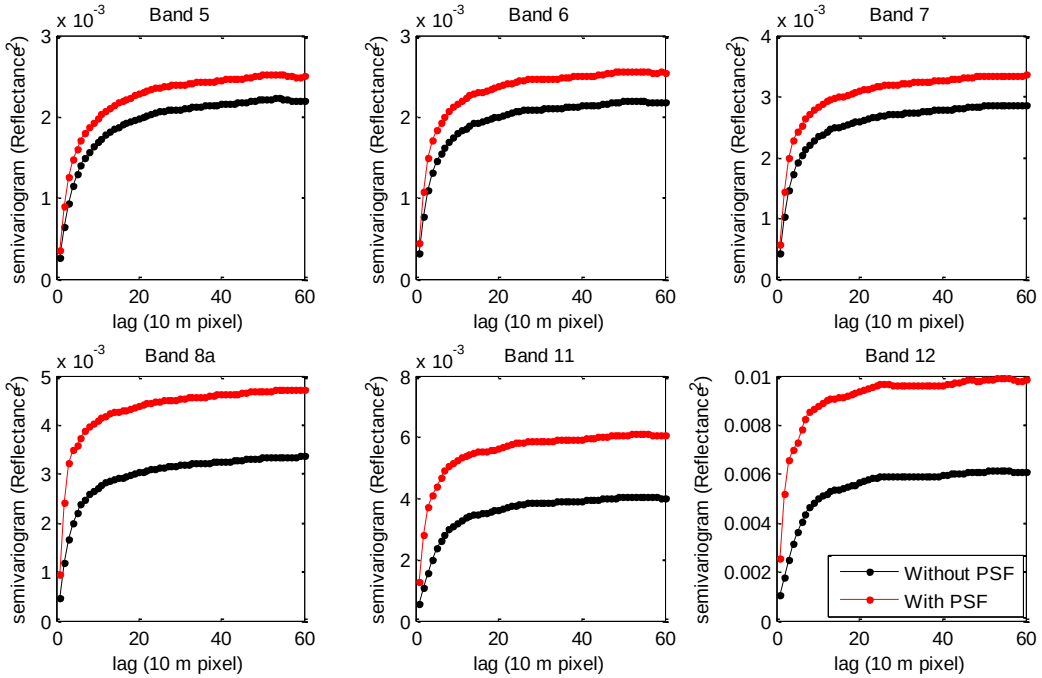
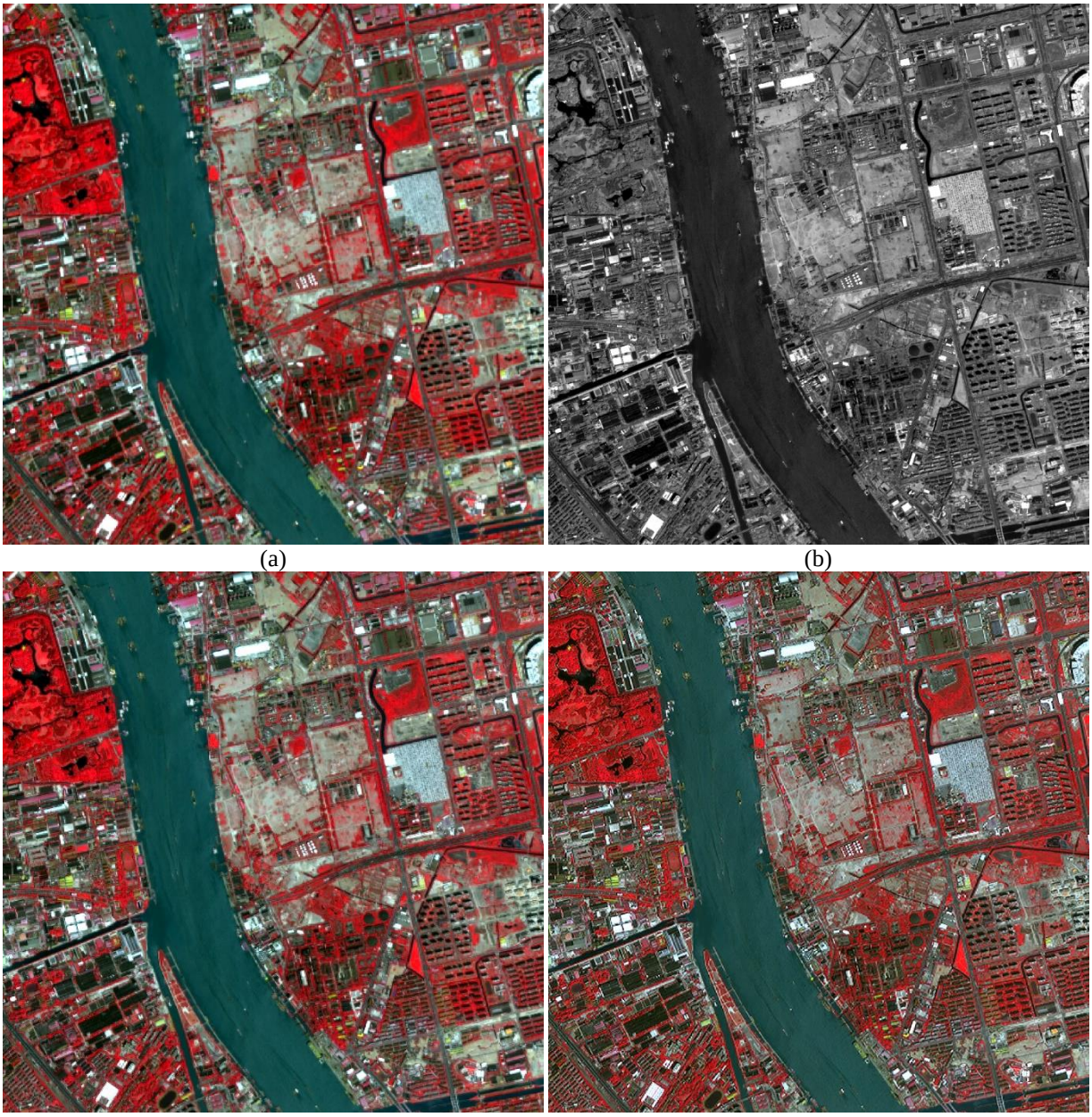


Fig. 13. Semivariograms of the ATPRK-based downscaling predictions for the Sentinel-2 image in Shanghai, China.

3) Application to Gaofen-6 image fusion. The application of the ATPRK method to the image acquired by the recently launched Chinese Gaofen-6 satellite was also investigated. The Gaofen-6 satellite carries two cameras: one acquires fine resolution multispectral images composed of four 8 m bands (blue, red, green and NIR) and one 2 m PAN band, while the other acquires medium resolution multispectral images composed of four 16 m bands (blue, red, green and NIR). In this experiment, the study area is a 4 km by 4 km city area of Shanghai, China. The Gaofen-6 test data were acquired on 18 April 2019 and we downscaled the 8 m bands (see Fig. 14(a)) to 2 m using the 2 m PAN band (see Fig. 14(b)). The results are shown in Fig. 14(c) and (d). It is seen that the two 2 m predictions are clearer than the original 8 m multispectral image. Moreover, the prediction produced by considering the PSF effect presents more spatial details (e.g., the boundaries of buildings can be observed more clearly).



(a) (b)

(c) (d)

Fig. 14. Downscaling result (2 m) for the Gaofen-6 image (bands 4, 3 and 2 as RGB). (a) 8 m multispectral image. (b) 2 m PAN image. (c) ATPRK without PSF. (d) ATPRK with PSF.

4. Discussion

4.1. Remarks on experimental design

In the experiments, for validating the proposed solution to PSF estimation in the multivariate case, coarse images contaminated by a *known* PSF were used, as the estimated PSF could then be evaluated objectively by comparing to the reference. The hyperspectral image with the same spatial resolution across all bands was used for testing. The reason for not using a multi-resolution image (e.g., Sentinel-2 image) is that the observed fine bands are inherently contaminated by the PSF effect. For convenience of illustration, the Sentinel-2 image is used here as an example. When simulating coarse bands contaminated by the PSF effect, the observed 20 m bands (bands 5, 6, 7, 8a, 11 and 12) need to be upscaled to 40 m by the known PSF. Meanwhile, the corresponding fine bands used for PSF estimation have to be simulated from the observed 10 m bands (bands 2, 3, 4 and 8) as well. That is, the four 10 m bands have to be upscaled to 20 m correspondingly, which will be used as covariates. However, the observed six 20 m bands are already contaminated by the PSF effect and this specific PSF is unknown in reality. When upscaling the observed four 10 m bands to 20 m, another transformation PSF needs to be identified accurately to maximize their relation with original 20 m bands 5, 6, 7, 8a, 11 and 12, producing reliable simulated covariates of four 20 m bands 2, 3, 4 and 8. This is challenging for validation of the proposed solution (which estimates exactly the transformation PSF) as extra uncertainty will be introduced. This is not the case for using hyperspectral image in the experiments, as no scale change was performed for the observed fine bands and we could focus solely on the performance of PSF estimation.

4.2. Computational cost

Due to the PSF effect, the signal of a pixel is contaminated by its neighbors. When considering the PSF effect in downscaling, more pixels from neighbors are involved. For ATPK-based solutions (original ATPK and ATPRK), this is achieved by the semivariogram modeling process. As shown in Eqs. (3) and (4), more fine pixels from neighbors within the spatial coverage the PSF are considered in convolution, rather than only the center coarse pixel in conventional ATPK (i.e., without PSF). The consideration of more pixels will undoubtedly increase the computational cost. Table 4 shows the computing time for the two Sentienl-2 images

in the multivariate case. By considering the PSF effect, the computing time increases from about 15 seconds to 6 minutes. This reveals that computing efficiency is indeed sacrificed when considering the PSF effect such as to produce more accurate downscaling predictions, but the final computational cost (i.e., several minutes) is generally acceptable.

Table 4 Computing time of ATPRK for the two Sentinel-2 images

Verona	ATPRK(PSF)	353.4s
	ATPRK(no PSF)	14.4s
Shanghai	ATPRK(PSF)	334.5s
	ATPRK(no PSF)	14.9s

4.3. Fitting model in PSF estimation

In this paper, for PSF estimation in the multivariate case, the fitting model for determining the relation between the coarse and fine bands in Eq. (13) was selected as the classical global linear regression model and the fitting error was calculated using the least squares method accordingly. In theory, other choices exist for quantifying the fitting error such as local regression or even geographically weighted regression (Brunsdon et al., 1996; Foody, 2003). These models may be more advantageous when the study area is large and the conditions of data acquisition (e.g., viewing angle) vary locally or regionally, and the effective PSF is a function of location as a result. In this case, the PSF can be estimated based on a pixel or object basis. This spatially adaptive scheme, however, may be computationally very expensive, especially when the spatial size of the study area is large. The gain in accuracy of PSF estimation and cost of computational burden need to be appropriately balanced according to the requirements of users.

4.4. PSF characterization

The geostatistical solutions (including ATPK and ATPRK) presented in this paper are applicable to the case involving the PSF effect. The PSF always differs between bands and sensors. It is necessary to characterize the

PSF accurately to guarantee the performance of the geostatistical solutions. In this paper, the Gaussian filter was assumed for PSF estimation in the multivariate case, as it is considered widely to approximate the real PSF satisfactorily, as claimed in most of the existing literature (Townshend et al., 2000; Huang et al., 2002; Manslow and Nixon, 2002; Kaiser and Schneider, 2008; Van der Meer, 2012; Vivone, 2015, 2019). In reality, however, the real PSF may not be the Gaussian filter, but a more complex form instead. Since the implementation of ATPK and ATPRK is not influenced by the specific form of PSF, these techniques can be applied potentially to different cases with various PSFs. For example, in the experiments for the univariate case, we examined the performance of ATPK for the assumed center flat PSF in Eq. (14). It was demonstrated that ATPK is also suitable for this type of PSF. Thus, once the real PSF is exactly known or estimated, the geostatistical solutions can be applied readily to reduce the PSF effect on downscaling continua.

(a) (b) (c)

Fig. 15. Downscaling result (10 m) for the Sentinel-2 images (bands 12, 8a and 5 as RGB) in Shanghai, China. (a) ATPRK without PSF. (b) ATPRK with Gaussian PSF. (c) ATPRK with the center flat PSF.

Additionally, in the multivariate case, we investigated the center flat PSF in Eq. (14) based on ATPRK. The result for the Shanghai Sentinel-2 image is shown in Fig. 15(c). From visual comparison, it was observed that the center flat PSF can lead to greater accuracy than without considering the PSF, but the result is ambiguous in comparison with the Gaussian PSF (see the marked sub-areas). The semivariograms for both PSFs are shown in Fig. 16. It is seen that the both semivariograms are very close for bands 5, 6 and 7 (the global

variances of the center flat PSF are only 3.83%, 4.52% and 4.87% larger than those of the Gaussian PSF for bands 5, 6 and 7, respectively). However, for bands 8a and 11, the Gaussian PSF leads to obvious larger semivariances (the global variances are 16.26% and 16.32% larger than those of the center flat PSF). The results suggest that the Gaussian PSF is not necessarily inferior to the center flat PSF, at least from the perspective of spatial variation reproduction.

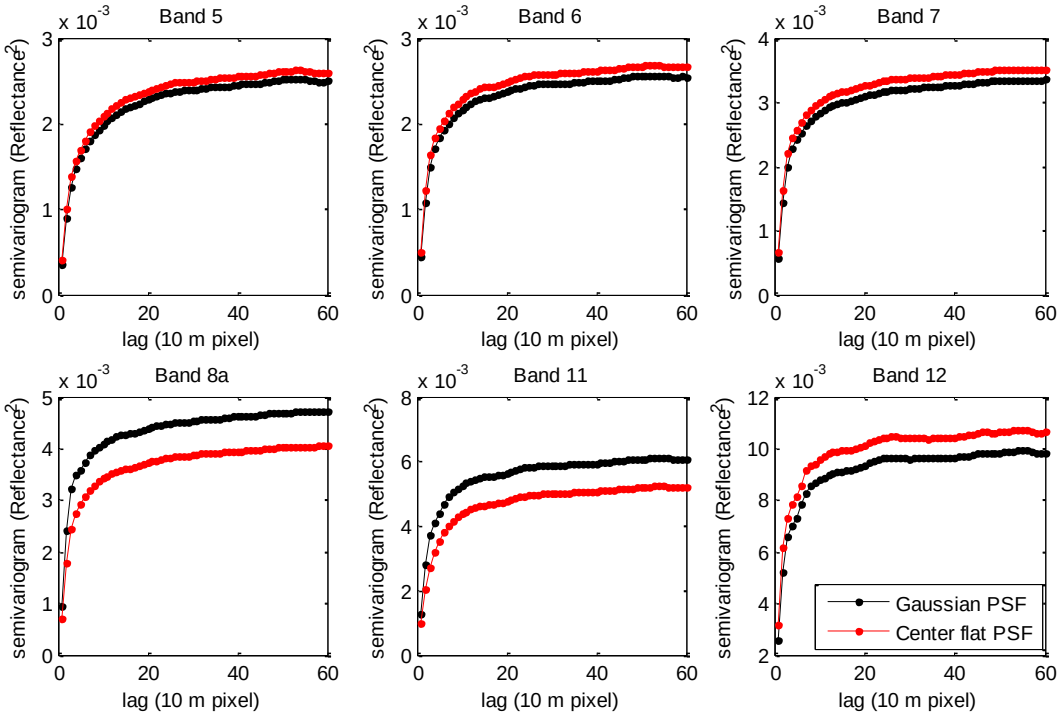


Fig. 16. Semivariograms of the ATPRK-based downscaling predictions using different PSFs for the Sentinel-2 image in Shanghai, China.

It is also interesting to consider different PSFs for the across-track and along-track directions (Inamdar et al., 2020). Moreover, nested models could also be investigated for characterizing the PSF. For example, the PSF can be a mixture of a Gaussian filter with different widths or even a mixture of different filters such as the Gaussian and triangular filter (Tan et al., 2006), or others (e.g., the ideal square wave filter). On the other hand, it should be highlighted that in downscaling, we are interested in the transformation PSF between the original coarse and target fine spatial resolution, not the PSF of the original measurements (i.e., from point on ground to pixel in image). This means that the PSF characterization and estimation depends, in addition to the sensor,

on the quality of both the observed coarse and fine bands, such that estimation needs to be done on a case-by-case basis. It is not clear how greatly the potential alternatives such as nested models can enhance the reliability of PSF characterization in practice and also how great the generalization ability is for dealing with data acquired by different sensors. This is part of our ongoing research.

4.5. Target fine spatial resolution in the multivariate case

For the multivariate case, this paper investigates the general case that the target spatial resolution of downscaling is exactly the same as the observed fine bands and, thus, the transformation PSF between the two spatial resolutions can be estimated reasonably. In some applications, the required target spatial resolution of downscaling is beyond that of any observed bands (Atkinson et al., 2008), and the proposed solution for PSF estimation cannot be used directly, as the data at the target fine spatial resolution are unavailable. In this case, the estimation of the required PSF can refer to the transformation PSF estimated between the available coarse and intermediate spatial resolutions, as the PSF is actually characterized by a continuous function and can be adopted to any discretization based on the pixel size. For example, the latter can be transferred to the former case straightforwardly or by slight empirical modification. The effectiveness of the potential solutions needs to be validated rigorously in future research. Undoubtedly, it is still encouraged to seek as much auxiliary data or information as possible at target fine spatial resolution to maximize the reliability of PSF estimation in this case.

4.6. Spectral distance between coarse and fine bands

The experiments show that the proposed PSF estimation method is applicable to the challenging case where the fine bands do not have the same wavelength as the coarse bands (e.g., in the hyperspectral and Sentinel-2 image cases). This is different from the special pan-sharpening case where the PAN waveband usually covers

almost all wavelengths of the coarse multispectral bands. However, it should be noted that normally the spectral distance between the coarse and fine bands cannot be too large. For example, the spatial texture information present in visual and near-infrared bands is substantially different from that in thermal bands and cannot be used for enhancing downscaling thermal images.

4.7. Uncertainties in downscaling

The univariate case exists commonly in reality, especially for sensors acquiring images with a single spatial resolution across all bands. In this paper, the PSF was assumed to be known exactly for the univariate case, as in most of the existing literature (Atkinson, 1997; Aiazzi et al., 2006; Bevilacqua et al., 2014). For more practical application of ATPK-based downscaling, the specific PSF needs to be known reliably beforehand. If no additional information (e.g., parameters provided by the manufacturer) is available, PSF estimation can be achieved by checking typical features in observed coarse images such as linear features. For example, a set of simulated fine spatial resolution features can be upscaled by convolving with various PSF candidates and then compared with the observed features to determine the optimal PSF. The question is whether the estimated PSF is applicable to other features in the coarse observations.

Downscaling is essentially an ill-posed, inverse problem. Compared to the multivariate case where fine bands are available to aid the downscaling process, the uncertainty in the univariate case is much greater. An increasing number of learning-based methods have been developed for downscaling in recent years (Kwon et al., 2015; Kwon and Tai, 2015; Yuan et al., 2020; Zhang et al., 2020a,b). This class of methods takes advantage of spatial structure information in available real fine spatial resolution images (i.e., training data) to train a model and can effectively reduce the uncertainty in downscaling. A potential advantage of learning-based methods is that the relation between the coarse (PSF-contaminated) and fine training data is captured directly by learning models, and thus, the specific PSF may not need to be known. The key issue is the quality of the training data. They need to have similar characteristics (e.g., spatial structure and spectral

wavelength) to the study area. If such training data are available, they can be used readily for enhancing downscaling.

As a typical branch of machine learning, deep learning has been developed widely for downscaling (Metari and Deschenes, 2007; He et al., 2011; Shajkofci and Liebling, 2018; Voynov et al., 2019). The main advantages of deep learning-based methods are the same as common learning-based methods, as elaborated above. The geostatistics methods presented in this paper, however, are substantially different from deep learning-based methods. The differences can be explained from three aspects. First, deep learning-based methods are data-based, while geostatistics methods are model-based. The former is performed mainly using a pre-defined black box, but the latter is a principled approach with analytical solutions. Second, in most cases, deep learning requires a large number of training data for network fitting. The geostatistics methods, however, are performed using solely the observed data and, thus, can be implemented readily in most applications. Third, since deep learning involves training, it is generally computationally expensive, whereas geostatistics methods generally require less computational cost, such as only several minutes as listed in Table 4. Considering the powerful non-linear fitting ability of deep learning for complex problems, however, it should still be worthwhile to investigate this class of method in future research.

4.8. Downscaling continua vs SPM

As mentioned in the Introduction, there are two classes of goal for downscaling: downscaling continua and SPM. The two goals can be related: the goal of SPM (i.e., hard classified land cover maps at finer spatial resolution) can be realized by downscaling continua and then hard classification. For reducing the influence of the PSF effect, it would be interesting to compare this potential downscaling-then-classification scheme with the recently developed SPM approach accounting for the PSF (Wang and Atkinson, 2017). The main difference is that for the former scheme, the PSF effect is considered in original multispectral images (e.g., in units of reflectance, digital number, etc.), while for the latter scheme, the PSF effect is considered in land cover

classification. The predictions of the two schemes also rely heavily on other factors, such as the hard classification method in the former scheme and the spectral unmixing and SPM methods in the latter scheme. Therefore, rigorous case studies need to be designed for fair comparison of the two schemes, and to check whether the downscaling-then-classification scheme is a more accurate alternative for reducing the PSF effect in SPM or when it is more advantageous.

5. Conclusion

It is acknowledged widely that the PSF effect exists ubiquitously in remote sensing images, but the development of effective solutions to reduce this effect in downscaling has not gained sufficient attention. This paper investigates ATPK-based geostatistical solutions to account for the PSF effect in downscaling continua and, thereby, enhance downscaling. Furthermore, a new solution is proposed for PSF estimation in the multivariate case. The experimental results suggest that the PSF effect should not be ignored and that it is beneficial to consider PSF in downscaling. The key findings can be summarized as follows.

1) Downscaling can be enhanced by considering the PSF effect. The predictions produced with accounting for the PSF are both visually and quantitatively more accurate than those without the PSF. More spatial details can be reproduced by considering the PSF in downscaling, which can provide more abundant information for decision-making.
2) The encouraging results suggest that the ATPK-based methods including the original ATPK and ATPRK are satisfactory solutions to the univariate and multivariate cases, respectively. For example, ATPK was demonstrated to be more accurate than the classical bicubic and polynomial interpolation methods and can also preserve the original coarse data perfectly.
3) It is critical to consider the correct PSF, and the proposed PSF estimation solution is a reliable choice for the multivariate case. The results in Fig. 10 show that the gain in restoration of spatial detail is limited if

the width of the Gaussian PSF is too small, while the texture is over-highlighted (presenting obvious noises) if the width is too large.

4) The results in Figs. 8 and 9 reveal that the PSF can be, and often is, different across bands. It is, thus, of great interest to apply band-dependent methods that can estimate the PSF for each band separately, as can be done exactly by the proposed PSF estimation solution.

This paper will also provide new insights for more general issue of downscaling remote sensing quantitative products, such as soil moisture and precipitation, etc..

Acknowledgment

This work was supported by National Natural Science Foundation of China under Grant 41971297, Fundamental Research Funds for the Central Universities under Grant 02502150021 and Tongji University under Grants 02502350047. The authors would like to thank the Editors and anonymous reviewers for their valuable and constructive comments which greatly improved this manuscript.

References

Aiazzi, B., Alparone, L., Baronti, S., Garzelli, A., Selva, M., 2006. MTF-tailored multiscale fusion of high-resolution MS and pan imagery. Photogrammetric Engineering and Remote Sensing 72, 591–596.

Atkinson, P. M., 1997. Mapping sub-pixel boundaries from remotely sensed images. Innovations in GIS 4, 166–180.

Atkinson, P. M., Pardo-Igúzquiza, E., Chica-Olmo, M., 2008. Downscaling cokriging for super-resolution mapping of continua in remotely sensed images. IEEE Transactions on Geoscience and Remote Sensing 46(2), 573–580.

Atkinson, P. M., 2009. Issues of uncertainty in super-resolution mapping and their implications for the design of an inter-comparison study. International Journal of Remote Sensing 30(20), 5293–5308.

Atkinson, P. M., 2013. Downscaling in remote sensing. International Journal of Applied Earth Observation and Geoinformation 22, 106–114.

Bevilacqua, M., Roumy, A., Guillemot, C., Morel, M-L. A., 2014. Single-image super-resolution via linear mapping of interpolated self-examples. IEEE Transactions on Image Processing 23(12), 5334–5347.

Brunsdon, C., Fotheringham, A. S., Charlton, M. E., 1996. Geographically weighted regression: A method for exploring spatial nonstationarity. Geographical Analysis 28(4), 281–298.

Chen, Y., Ge, Y., Chen, Y., Jin, Y., An, R., 2018. Subpixel land cover mapping using multiscale spatial dependence. IEEE Transactions on Geoscience and remote sensing 56(9), 5097–5106.

Drusch, M. et al., 2012. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. Remote Sensing of Environment 120, 25–36.

Foody, G. M., 2003. Geographical weighting as a further refinement to regression modelling: An example focused on the NDVI–rainfall relationship. Remote Sensing of Environment 88, 283–293.

Gao, F., Anderson, M. C., Zhang, X., Yang, Z., Alfieri, J. G., Kustas, W. P., Mueller, R., Johnson, D. M., Prueger, J. H., 2017. Toward mapping crop progress at field scales through fusion of Landsat and MODIS imagery. Remote Sensing of Environment 527(188), 9–25.

Ge, Y., Chen, Y., Stein, A., Li, S., Hu, J., 2016. Enhanced subpixel mapping with spatial distribution patterns of geographical objects. IEEE Transactions on Geoscience and Remote Sensing 54(4), 2356–2370.

Goovaerts, P., 2008. Kriging and semivariogram deconvolution in presence of irregular geographical units. Mathematical Geosciences 40(1), 101–128.

He, Z., Tang, S., Yang, J., Cao, Y., Ying, M., Cao, Y., 2011. Cascaded deep networks with multiple receptive fields for infrared image super-resolution. IEEE Transactions on Circuits and Systems for Video Technology 29(8), 2310–2322.

Huang, C., Townshend, R. G., Liang, S., Kalluri, S. N. V., DeFries, R. S., 2002. Impact of sensor's point spread function on land cover characterization: Assessment and deconvolution. Remote Sensing of Environment 80, 203–212.

Inamdar, D., Kalacska, M., Leblanc, G., Arroyo-Mora, J. P., 2020. Characterizing and mitigating sensor generated spatial correlations in airborne hyperspectral imaging data. Remote Sensing 12, 641.

Jia, S., Zhu, W., Lű, A., Yan, T., 2011. A statistical spatial downscaling algorithm of TRMM precipitation based on NDVI and DEM in the Qaidam Basin of China. Remote Sensing of Environment 115, 3069–3079.

Jin, Y., Ge, Y., Wang, J., Chen, Y., Heuvelink, G. B. M., Atkinson, P. M., 2018. Downscaling AMSR-2 soil moisture data with geographically weighted area-to-area regression kriging. IEEE Transactions on Geoscience and Remote Sensing 56(4), 2362–2376.

Kaiser, G., Schneider, W., 2008. Estimation of sensor point spread function by spatial subpixel analysis. International Journal of Remote Sensing 29, 2137–2155.

Kwon, H., Tai, Y., 2015. RGB-guided hyperspectral image upsampling. IEEE International Conference on Computer Vision (ICCV), Santiago, 307–315.

Kwon, Y., Kim, K. I., Tompkin, J., Kim, J. H., Theobalt, C., 2015. Efficient learning of image super-resolution and compression artifact removal with semi-local Gaussian processes. IEEE Transactions on Pattern Analysis and Machine Intelligence 37(9), 1792–1805.

Kyriakidis, P. C., 2004. A geostatistical framework for area-to-point spatial interpolation. Geographical Analysis 36(3), 259–289.

Li, X., Du, Y., Ling, F., Wu, S., Feng, Q., 2011. Using a sub-pixel mapping model to improve the accuracy of landscape pattern indices. Ecological Indicators 11(5), 1160–1170.

Li, X., Ling, F., Foody, G. M., Ge, Y., Zhang, Y., Du, Y., 2017. Generating a series of fine spatial and temporal resolution land cover maps by fusing coarse spatial resolution remotely sensed images and fine spatial resolution land cover maps. Remote Sensing of Environment 196, 293–311.

Loncan, L. et al., 2015. Hyperspectral pansharpening: A review. IEEE Geoscience and Remote Sensing Magazine 3(3), 27–46.

Manslow, J. F., Nixon, M. S., 2002. On the ambiguity induced by a remote sensor's PSF. In Uncertainty in Remote Sensing and GIS, 37–57.

Metari, S., Deschenes, F., 2007. A new convolution kernel for atmospheric point spread function applied to computer vision. 11th IEEE International Conference on Computer Vision (ICCV), Rio de Janeiro, 1–8.

Pardo-Iguzquiza, E., Rodr ́guez-Galiano, V. F., Chica-Olmo, M., Atkinson, P. M., 2011. Image fusion by spatially adaptive filtering using downscaling cokriging. ISPRS Journal of Photogrammetry and Remote Sensing 66, 337–346.

Peng, J., Loew, A., Merlin, O., Verhoest, N. E. C., 2017. A review of spatial downscaling of satellite remotely sensed soil moisture. Reviews of Geophysics 55, 341–366.

Shajkofci, H., Liebling, M., 2018. Semi-blind spatially-variant deconvolution in optical microscopy with local point spread function estimation by use of convolutional neural networks. 25th IEEE International Conference on Image Processing (ICIP), Athens, 3818–3822.

Shen, H., Huang, L., Zhang, L., Wu, P., Zeng, C., 2016. Long-term and fine-scale satellite monitoring of the urban heat island effect by the fusion of multi-temporal and multi-sensor remote sensed data: A 26-year case study of the city of Wuhan in China. Remote Sensing of Environment 172, 109–125.

Trang, V., Jaime, E. H., Kimberly, A. B., Sun, Z., Rulla, M., 2016. Tamimi and Francine Laden, spatiotemporal exposure modeling of ambient erythemal ultraviolet radiation. Environmental Health 15, 111.

Tan, B. et al., 2006. The impact of gridding artifacts on the local spatial properties of MODIS data: Implications for validation, compositing, and band-to-band registration across resolutions. Remote Sensing of Environment 105, 98–114.

Townshend, R. G., Huang, C., Kalluri, S. N. V., Defries, R. S., Liang, S., 2000. Beware of per-pixel characterization of land cover. International Journal of Remote Sensing 21, 839–843.

Van der Meer, F. D., 2012. Remote-sensing image analysis and geostatistics. International Journal of Remote Sensing 33(18), 5644–5676.

Vivone, G., Alparone, L., Chanussot, J., Mura, M. D., Garzelli, A., Licciardi, G. A., Restaino, R., Wald, L., 2015. A critical comparison among pansharpening algorithms. IEEE Transactions on Geoscience and Remote Sensing 53(5), 2565–2586.

Vivone, G., Simões, M., Mura, M. D., Restaino, R., Bioucas-Dias, J. M., Licciardi, G. A., Chanussot, J., 2015. Pansharpening based on semiblind deconvolution. IEEE Transactions on Geoscience and Remote Sensing 53(4), 1997–2010.

Vivone, G., Addesso, P., Restaino, R., Mura, M. D., Chanussot, J., 2019. Pansharpening based on deconvolution for multiband filter estimation. IEEE Transactions on Geoscience and Remote Sensing 57(1), 540–553.

Voynov, O., Artemov, A., Egiazarian, V., Notchenko, A., Bobrovskikh, G., Burnaev, E., Zorin, D., 2019. Perceptual deep depth super-resolution. IEEE International Conference on Computer Vision (ICCV), Seoul, 5652–5662.

Wang, Q., Zhang, C., Atkinson, P. M, 2020a. Sub-pixel mapping with point constraints. Remote Sensing of Environment 244, 111817.

Wang, J., Schmitz, O., Lu, Meng., Karssenberg, D., 2020b. Thermal unmixing based downscaling for fine resolution diurnal land surface temperature analysis. ISPRS Journal of Photogrammetry and Remote Sensing 161, 76–89.

Wang, Q., Shi, W., Atkinson, P. M., Zhao, Y., 2015. Downscaling MODIS images with area-to-point regression kriging. Remote Sensing of Environment 166, 191–204.

Wang, Q., Atkinson, P. M., 2017. The effect of the point spread function on sub-pixel mapping. Remote Sensing of Environment 193, 127–137.

Wang, Q., Shi, W., Atkinson, P. M., 2020c. Information loss-guided multi-resolution image fusion. IEEE Transactions on Geoscience and Remote Sensing 58(1), 45–57.

Wardrop, N. A., Jochem, W. C., Bird, T. J., Chamberlain, H. R., Clarke, D., Kerr, D., Bengtsson, L., Juran, S., Seaman, V., Tatem, A. J., 2018. Spatially disaggregated population estimates in the absence of national population and housing census data. Proceedings of the National Academy of Sciences of the United States of America 115(14), 3529–3537.

Xu, S., Wu, C., Wang, L., Gonsamo, A., Shen, Y., Niu, Z., 2015. A new satellite-based monthly precipitation downscaling algorithm with non-stationary relationship between precipitation and land surface characteristics. Remote Sensing of Environment 162, 119–140.

Yuan, Q., Shen, H., Li, T., Li, Z., Li, S., Jiang, Y., Xu, H., Tan, W., Yang, Q., Wang, J., Gao, J., Zhang, L., 2020. Deep learning in environmental remote sensing: Achievements and challenges. Remote Sensing of Environment 241, 111716.

Zhang, Q., Yuan, Q., Li, J., Li, Z., Shen, H., Zhang, L., 2020a. Thick cloud and cloud shadow removal in multitemporal images using progressively spatio-temporal patch group deep learning. ISPRS Journal of Photogrammetry and Remote Sensing 162, 148–160.

Zhang, Q., Yuan, Q., Li, J., Sun, F., Zhang, L., 2020b. Deep spatio-spectral Bayesian posterior for hyperspectral image non-i.i.d. noise removal. ISPRS Journal of Photogrammetry and Remote Sensing 164, 125–137.

Zhang, X., Wang, J., Gao, F., Liu, Y., Schaa, C., Friedl, M., Yu, Y., Jayavelu, S., Gray, J., Liu, L., Yan, D., Henebry, G. M., 2017. Exploration of scaling effects on coarse resolution land surface phenology. Remote Sensing of Environment 190, 318–330.

Zhao, J., Zhong, Y., Wu, Y., Zhang, L., Shu, H., 2015. Sub-pixel mapping based on Conditional Random Fields for hyperspectral remote sensing imagery. IEEE Journal of Selected Topics in Signal Processing 9(6), 1049–1060.