A Geostatistical Filter for Remote Sensing Image Enhancement

Qunming Wang [1,*], Xiaohua Tong [1], Peter M. Atkinson [2,3,4,5]

[1] College of Surveying and Geo-Informatics, Tongji University, 1239 Siping Road, Shanghai 200092, China
[2] Lancaster Environment Centre, Lancaster University, Lancaster LA1 4YQ, UK
[3] Geography and Environment, University of Southampton, Highfield, Southampton SO17 1BJ, UK
[4] School of Geography, Archaeology and Palaeoecology, Queen's University Belfast, BT7 1NN, Northern Ireland, UK
[5] Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

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

Abstract: In this paper, a new method was investigated to enhance remote sensing images by alleviating the point spread function (PSF) effect. The PSF effect exists ubiquitously in remotely sensed imagery. As a result, image quality is greatly affected, and this imposes a fundamental limit on the amount of information captured in remotely sensed images. A geostatistical filter was proposed to enhance image quality based on a downscaling-then-upscaling scheme. The difference between this method and previous methods is that the PSF is represented by breaking the pixel down into a series of sub-pixels, facilitating downscaling using the PSF and then upscaling using a square wave response. Thus, the sub-pixels allow disaggregation as an attempt to remove the PSF effect. Experimental results on simualted and real datasets both suggest that the proposed filter can enhance the original images by reducing the PSF effect and quantify the extent to which this is possible. The predictions using the new method outperform the original coarse PSF-contaminated imagery as well as a benchmark method. The proposed method represents a new solution to compensate for the limitations introduced by remote sensors (i.e., hardware) using computer techniques (i.e., software). The method has widespread application value, particularly for applications based on remote sensing image analysis.

Keywords: Remote sensing, Geostatistics, Image enhancement, Image filtering, Point spread function (PSF).

1 Introduction

Image enhancement is an important technique in image processing, optics, computer graphics and computer vision. It refers to enhancement of the visual appearance of an observed image (usually called degraded image) by increasing the sharpness and reducing the blur. It is also termed image sharpening in the field of image processing. The technique plays a critical role in many applications such as robotics, medical imaging and quality control (Gilboa et al. 2002), where the enhanced data can help to support more reliable decisions.

In this paper, image enhancement was investigated for remote sensing images. Due to the limitations of hardware, the character of landscapes and the uncertainties in imaging (e.g., atmospheric condition, sensor platform, etc.), blurry artifacts exist commonly in acquired remotely sensed images and have brought great challenges for post-image processing, analysis and related applications. This issue necessitates the development of more up-to-date image enhancement methods for remote sensing image enhancement or sharpening.

Note that the terminology of image sharpening in remote sensing refers to pan-sharpening in some cases, where a fine spatial resolution panchromatic band image is available for sharpening the coarse multispectral band images (i.e., the multivariate case), and the spatial resolution of the prediction is finer than the original multispectral images (Vivone et al. 2015). In this paper, the term sharpening was used to refer to the more general case where only the degraded image is available as input (i.e., the univariate case), and the prediction is at the same spatial resolution as the original degraded image.

Over the past few decades, several enhancement methods have been developed. The linear unsharp masking (UM) is one of most widely used methods and it has been incorporated into popular software such as Adobe Photoshop (Cao et al. 2011). The basic idea dehind UM is to add a weighted high-pass filtered version of the input image back to the original data and, thus, enhance the edge and detail information (Polesel et al. 2000). UM underpins the development of various further methods including adaptive UM (Polesel et al. 2000), optimal UM (OUM) (Kim and Allebach 2005), and polynomial UM (Ramponi et al. 1996). In addition, other solutions such as the fuzzy logic filter (Farbiz et al. 2000), bilateral filter (Zhang and Allebach 2008) and Munsell's scale-based scheme (Matz et al. 2006) were also developed for image enhancement.

As acknowledged widely, the point spread function (PSF) effect exists ubiquitously in remotely sensed data and it is the main factor leading to blur. The PSF effect is caused mainly by the optical system, motion of the sensor platform (in in-track and/or crosstrack direction), the detector and electronics, atmospheric effects, and image resampling (Huang et al. 2002; Kaiser and Schneider 2008). Due to the PSF effect, the signal of a pixel is contaminated by its neigbhoring pixels, rather than being a simple average of the contributions from within the spatial coverage of the pixel (Manslow and Nixon 2002; Townshend et al. 2000; Van der Meer 2012). The PSF is also termed adjacency effect in some literature (Kükenbrink et al. 2019; Wang et al. 2019; Zheng et al. 2019). As a result, the inherent contrast between pixels is reduced and blur is created. Such an effect has also been investigated for thermal infrared data (Zheng et al. 2019) and hyperspectral data (Wang et al. 2019) in recent years.

Of the existing image enhancement methods, very few consider the physical blurring process due to the PSF effect. Huang et al. (2002) proposed a deconvolution approach to reduce the PSF effect in land cover mapping. The method quantifies the influence of neighbors at the original pixel level, which is a coarse characterization of the PSF. Kaiser and Schneider (2008) proposed a spatial sub-pixel analysis-based method for joint sensor PSF estimation and image sharpening. This method, however, is designed for the special case where land cover patterns are composed of large homogeneous fields separated by straight boundaries. It is, therefore, of great interest to consider the PSF effect for more general land cover patterns to produce more accurate predictions through image enhancement.

In this paper, a geostatistical filter was proposed to perform image enhancement from the viewpoint of the physical PSF effect in the imaging process. The effect of surrounding pixels on the central pixel was investigated, as well as how to reduce the effect. The filter can also be thought of as a sub-pixel analysis method that explores the information at the sub-pixel level based on the principles of geostatistics. Different from the sub-pixel analysis-based method in Kaiser and Schneider (2008), its implementation is not restricted by the specific structure of landscapes. In comparison with the classical UM approach that treats the effect of neighbors equally and determines the parameter (i.e., weight for the high-pass filter) manually, the geostatistical filter determines the contributions of neighbors adaptively and the parameters are estimated automatically. Moreover, the geostatistical filter is conceptually suitable for *any* PSF. Crucially,

the method and its predictions can reduce the PSF-imposed limit on the amount of information captured in remote sensing images.

In our previous research, it was demonstrated that the PSF effect can impact spectral unmixing (Wang et al. 2018) and sub-pixel mapping (Wang and Atkinson 2017) greatly, and developed some solutions to reduce the PSF effect in these two specific problems. This paper, however, investigates the more *general* case and finds solutions to enhance the original multispectral images. Specifically, the predictions of Wang et al. (2018) and Wang and Atkinson (2017) are a set of class proportions of land cover and thematic land cover maps, respectively. The prediction of this paper is enhanced multispectral images (in units of surface reflectance, brightness, radiance, etc.) at the original spatial resolution. Importantly, the predictions produced using the proposed filter in this paper have more widespread application value. They can be used for more reliable post-image analysis (not only for spectral unmixing and sub-pixel mapping, but also for many other technical problems) and post-quantitative-inversion (e.g., inversion of environmental and ecological variables based on reflectance data) that need to reduce the effect imposed by the PSF.

The rest of the paper is organized as follows. In Sect. 2, the mechanism of the PSF effect on remote sensing images (pixel values) is introduced briefly and then the geostatistical filter is introduced for reducing the PSF effect and realizing image enhancement. The theoretical advantages of the method over UM is also analyzed. Section 3 provides the experimental results for both a simulated dataset and a real dataset to validate the proposed method. Further analysis, remarks and relevant further issues are discussed in Sect. 4. Section 5 concludes the paper.

2 Methods

2.1 Difference Between Observed Data and Ideal Data

Let G_{Vb} be the observed gray value (e.g., reflectance, radiance, brightness, etc.) of pixel V in band b (b=1, 2, …, B, where B is the number of bands of the observed image), and G_{vb} be the *error-free* gray value of sub-pixel v in band b. Due to the PSF effect, the pixel value of V is a convolution of the values of sub-pixels

$$G_{Vb} = G_{vb} * h_V ,\tag{1}$$

where $*$ is the convolution operator and h_V is the PSF.

The task of image enhancement is identified as the prediction of the ideal pixel value of V (denoted as Q_{Vb}) for all pixels and all bands of the image. Q_{Vb} is the value not contaminated by its neighbors and it is assumed that the blur can be eliminated by restoration of the image composed of the individual Q_{Vb}. The ideal pixel value is conceptually defined as the average of all *error-free* sub-pixel values falling within V

$$Q_{Vb} = G_{vb} * h_V{}' ,\tag{2}$$

in which $h_V{}'$ is a specified square wave filter (Pardo-Iguzquiza et al. 2006; Wang et al. 2018)

$$h_V{}'(\mathbf{x}) = \begin{cases} \dfrac{1}{\tau}, & \text{if } \mathbf{x} \in V(\mathbf{x}) \\ 0, & \text{otherwise} \end{cases} ,\tag{3}$$

where $\mathbf{x}$ is the two-dimensional spatial coordinate of a sub-pixel, $V(\mathbf{x})$ is the spatial coverage of the coarse pixel V, and τ is the pixel areal ratio (in zoom) between V and v (i.e., the number of sub-pixels v within V). Eq. (2) means that the ideal pixel value is determined only by the sub-pixels within it and it is not affected by neighboring pixels.

The real PSF h_v in practical cases is not the same as the ideal square wave $h_v{'}$. The differences are generally twofold.

i)   The spatial coverage of h_V is normally larger than a pixel (as is the case in $h_v{'}$). This means that the observed pixel value G_{Vb} is affected by its surrounding pixels, and more precisely, is a convolution of the sub-pixel values not only within, but also surrounding pixel V.

ii) Different sub-pixels (both within and surrounding the center pixel) have different effects on the center pixel. Compared to more distant sub-pixels, the close (relative to the centroid of the observed pixel) sub-pixels have greater effect on G_{Vb}.

As widely acknowledged, the PSF h_V is often assumed to be a Gaussian filter (Campagnolo and Montano 2014; Huang et al. 2002; Kaiser and Schneider 2008; Manslow and Nixon 2002; Townshend et al. 2000; Van der Meer 2012; Wang et al. 2018; Wenny et al. 2015)

$$h_V(\mathbf{x}) = \begin{cases} \dfrac{1}{2\pi\sigma^2}\exp\left[-\left(\dfrac{x_1^2 + x_2^2}{2\sigma^2}\right)\right], & \text{if } \mathbf{x}\in V'(\mathbf{x}) \\ 0, & \text{otherwise} \end{cases}, \tag{4}$$

in which σ is the standard deviation (related to the width of the Gaussian PSF), $\mathbf{x} = \{x_1, x_2\}$, and $V'(\mathbf{x})$ is the spatial coverage on which the PSF h_V can exert an effect. $V'(\mathbf{x})$ is usually a local window centered at coarse pixel V and it is larger than $V(\mathbf{x})$ in Eq. (3).

The obvious difference between h_V and $h_v{'}$ leads to a substantial lack of coherence between the observed pixel value G_{Vb} and ideal value Q_{Vb}. It is clear that the prediction of Q_{Vb} in Eq. (2) is essentially the prediction of the *error-free* sub-pixel values G_{vb}. As seen from the relation in Eq. (1), G_{vb} can be predicted inversely if the real PSF h_V is available. The predicted sub-pixel values can then be convolved with the ideal square wave filter to produce Q_{Vb}. In Sect. 2.2, a geostatistical solution is proposed to predict G_{vb}.

2.2 The Proposed Geostatistical Filter

The inverse prediction problem of predicting sub-pixel values G_{vb} from pixel value G_{Vb} involves downscaling, which predicts 'point' values from observed 'areal' values. The challenge in the mathematical problem is to account for the PSF h_V explicitly. The geostatistics-based area-to-point kriging (ATPK) is exactly a downscaling approach that can account for the PSF explicitly. It is employed to predict sub-pixel values G_{vb} in this paper. ATPK is essentially an interpolation approach, where the sub-pixel values are predicted as a weighted combination of neighboring pixel values. Specifically, for a given sub-pixel v_i, the value is predicted as (Kyriakidis 2004)

$$\hat{G}_{v_ib} = \sum_{j=1}^{N} \lambda_{ij} G_{V_jb}, \text{ s.t. } \sum_{j=1}^{N} \lambda_{ij} = 1,\tag{5}$$

where λ_{ij} is the weight for neighboring pixel V_j and N is the number of neighboring pixels (including the center pixel). It is always necessary to determine an optimal size of neighborhood in advance. The size is determined according to the width of the PSF. If the width of the PSF is large, a large number of neighbors affect the center pixel and a large number of pixels in the neighborhood need to be considered in ATPK; If the width of the PSF is small, only a small number of neighbors need to be considered. The weights in Eq. (5) are calculated according to the equation below (Kyriakidis 2004)

$$\boldsymbol{\gamma}_{CC} \boldsymbol{\lambda}_i = \boldsymbol{\gamma}_{FC},\tag{6}$$

where $\boldsymbol{\gamma}_{CC}$ is a $(N+1) \times (N+1)$ matrix composed of area-to-area semivariograms, $\boldsymbol{\gamma}_{FC}$ is a $(N+1) \times 1$ vector composed of point-to-area semivariograms, and $\boldsymbol{\lambda}_i$ is a $(N+1) \times 1$ vector composed of the weights. The calculation of $\boldsymbol{\gamma}_{CC}$ and $\boldsymbol{\gamma}_{FC}$ accounts for the PSF h_V in the scale transformation. Note that in this paper, the pixel gray values are used for semivariogram modeling. This is different from the cases in spectral unmixing (Wang et al. 2018) and sub-pixel mapping (Wang and Atkinson 2017), where land cover proportions (with values between 0 and 1 for each class) and land cover class labels (with value of 0 or 1 for each class) are used for semivariogram modeling, respectively. Readers may refer to Wang et al. (2015, 2016) for more detail on the calculation of the semivariograms.

A crucial property of ATPK is perfect coherence: when the ATPK predictions $\hat{G}_{vb}$ are upscaled (convolved with the same PSF h_V) back to the spatial resolution of pixel V, the value is exactly the same as the original data G_{Vb} (Kyriakidis 2004; Kyriakidis and Yoo 2005)

$$G_{Vb} = \hat{G}_{vb} * h_V.\tag{7}$$

Thus, Eq. (7) demonstrates mathematically that the ATPK prediction $\hat{G}_{vb}$ is a reliable solution for predicting the pixel value G_{Vb} identified in Eq. (1), as it strictly satisfies Eq. (1).

The ideal pixel value Q_{Vb} of the enhanced image is finally predicted as

$$\hat{Q}_{Vb} = \hat{G}_{vb} * h_V'.\tag{8}$$

The proposed method is compared to the classical UM approach. In UM, the Laplacian operator is used widely and the prediction, denoted as, $\hat{U}_{Vb}$, is calculated as

$$\begin{aligned}
\hat{U}_{Vb} &= G_{Vb} + w\left(G_{Vb} - \frac{1}{N'-1} \sum_{k=1}^{N'-1} G_{V_kb} \right) \\
&= (w+1)G_{Vb} - \frac{w}{N'-1} \sum_{k=1}^{N'-1} G_{V_kb} \\
&= \sum_{j=1}^{N'} \beta_j G_{V_jb}
\end{aligned}\tag{9}$$

where

$$\beta_j = \begin{cases} w+1, & \text{if } V_j \text{ is the center pixel} \\ \dfrac{w}{N'-1}, & \text{otherwise} \end{cases} . \tag{10}$$

In Eqs. (10) and (11), $w \in (0,1]$ is the weight for the high-pass filtered version and N' is the number of neighboring pixels (including the center pixel). For the Laplacian kernel, the 4-neighborhood or 8-neighborhood is always used.

With respect to the proposed ATPK-based geostatistical filter, the prediction in Eq. (8) can be further transformed as follows

$$\begin{aligned}
\hat{Q}_{Vb} &= \hat{G}_{vb} * h_V' \\
&= \frac{1}{\tau} \sum_{i=1}^{\tau} \left(\hat{G}_{v_i b} \right) \\
&= \frac{1}{\tau} \sum_{i=1}^{\tau} \left(\sum_{j=1}^{N} \lambda_{ij} G_{V_j b} \right), \\
&= \sum_{j=1}^{N} \left(\frac{1}{\tau} \sum_{i=1}^{\tau} \lambda_{ij} \right) G_{V_j b} \\
&= \sum_{j=1}^{N} \alpha_j G_{V_j b}
\end{aligned} \tag{11}$$

where $\alpha_j = \sum_{i=1}^{\tau} \lambda_{ij} / \tau$ is the weight for the neighboring pixel V_j. This means that the final prediction of the proposed method can be viewed as a linear weighted combination of the neighboring pixel values.

As seen from Eq. (11), the mathematical model of the proposed filter has the same format as the UM method, but the weights are totally different. Usually, the parameter w in UM is determined manually or empirically. In the proposed method, the parameter λ_{ij} is determined automatically according to the Kriging matrix in Eq. (6), where the semivariograms are computed based on the spatial structure information (in terms of spatial variation) and different neighbors exert a different influence on the center pixel as a result. Moreover, sub-pixel level information is explored in the proposed filter, which is not the case in the UM method. Regarding the generalization capability, physically there is no restriction on the land cover patterns for the proposed filter. That is, it is suitable for any land cover patterns. Mathematically, its implementation is not affected by the specific form of PSF and any PSF can be used readily once the PSF is known or predicted.

Our previous work for spectral unmixing (Wang et al. 2018) and sub-pixel mapping (Wang and Atkinson 2017) was carried out based on the mathematical relation between coarse and sub-pixel proportions (or sub-pixel class labels). This paper investigates a more general mathematical model. The previous mathematical models can be viewed as special cases of the model in Eq. (1) in this paper, where the more general terms of coarse pixel gray value G_{Vb} and fine sub-pixel value G_{vb} are used. The predictions of this paper can be used to reduce the PSF effect in spectral unmixing and sub-pixel mapping (i.e., the two techniques can be performed on the filtered data directly instead of on the original observed data). Moreover, the filtered data have value in any application where there is a need to reduce the PSF effect.

3 Experiments

Two datasets, including a simulated PSF-contaminated, blurry dataset and a real dataset, were used to validate the proposed geostatistical filter-based image enhancemnet method. In the first experiment, a blurry dataset was simulated by convolving a 30 m Landsat 8 image with a Gaussian PSF. The geostatistical filtering method was then applied to enhance these data simulated with various standard deviations and spatial resolutions. With such a design, one can concentrate solely on the recovery capability of the proposed method and, more importantly, the ideal images (i.e., images that are not contaminated by the PSF) are known perfectly as a reference for evaluation. The root mean square error (RMSE) and correlation coefficient (CC) were used to provide a quantitative evaluation of the predictions. In the second experiment, the method was applied to enhance a real dataset and the evaluation was performed by visual inspection and comparison of empirical semivariograms before and after filtering.

3.1 Experiment 1: Simulated Data

The selected Landsat 8 surface reflectance image covers an area of 18 km by 18 km (corresponding to 600 by 600 pixels) in Coleambally, Australia. The image was acquired on 6 July 2013. Bands 2-4 were considered in the experiment and the false color composite is shown in Fig. 1.
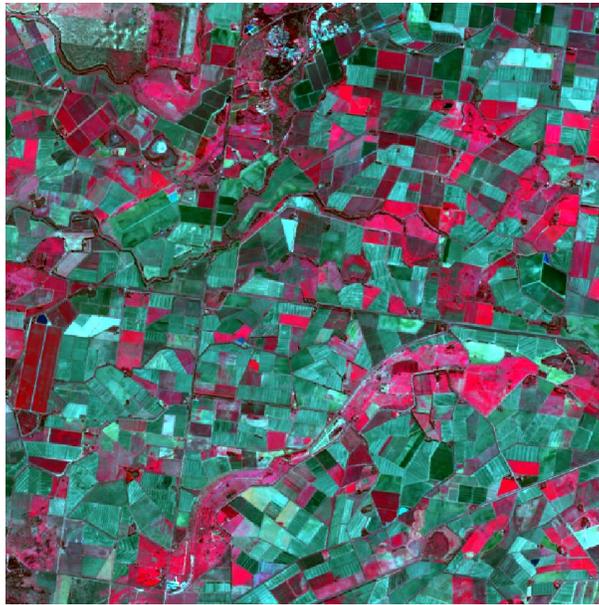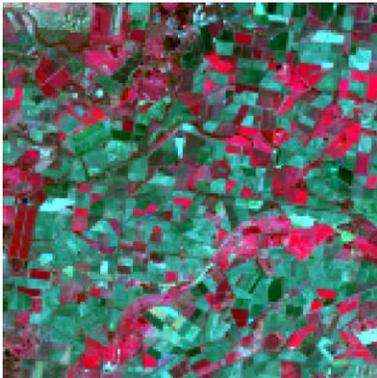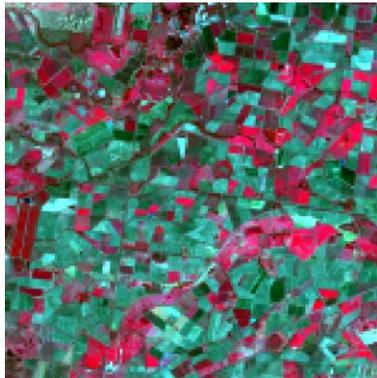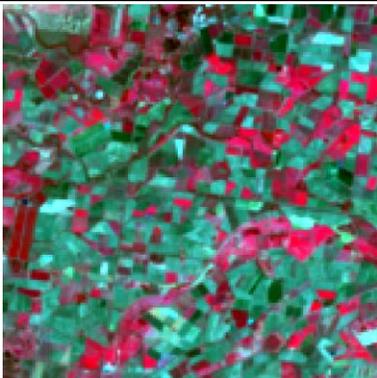

Fig. 1 The Landsat 8 image used in experiment 1 (bands 4, 3 and 2 as RGB)

3.1.1 Data Simulated with Different PSF Parameters

The 30 m Landsat image in Fig. 1 was upscaled with a factor of 4 and a Gaussian PSF. The standard deviation of the PSF (i.e., σ) was set to 0.3, 0.5, 0.7 and 0.9, creating four different PSF-contaminated images with a spatial resolution of 120 m (containing 150 by 150 pixels), as shown in the first column of Fig. 2. The last column shows the 120 m reference image (i.e., the ideal image that is not contamitaed by the PSF) that was produced by upscaling the 30 m data with the ideal square wave filter. Comparing the blurred images with the ideal image, it is seen that as σ

increases, the image becomes increasingly blurred and more spatial detail is lost. For example, the boundaries between classes appear to be more ambiguous and the linear features are almost completely lost in the image contaminated with $\sigma=0.9$.

The proposed ATPK-based geostatistical filter was used to enhance the four blurred images, and the filtered images are shown in the second column of Fig. 2. It is clear how the original images are enhanced and how the prediction of the proposed method is similar to the 120 m reference in all four cases. Even for the case of $\sigma=0.9$ where the blurring was the most serious, the filter can still reproduce most of the spatial details. The results suggest that the proposed filter is applicable for PSFs with various widths.

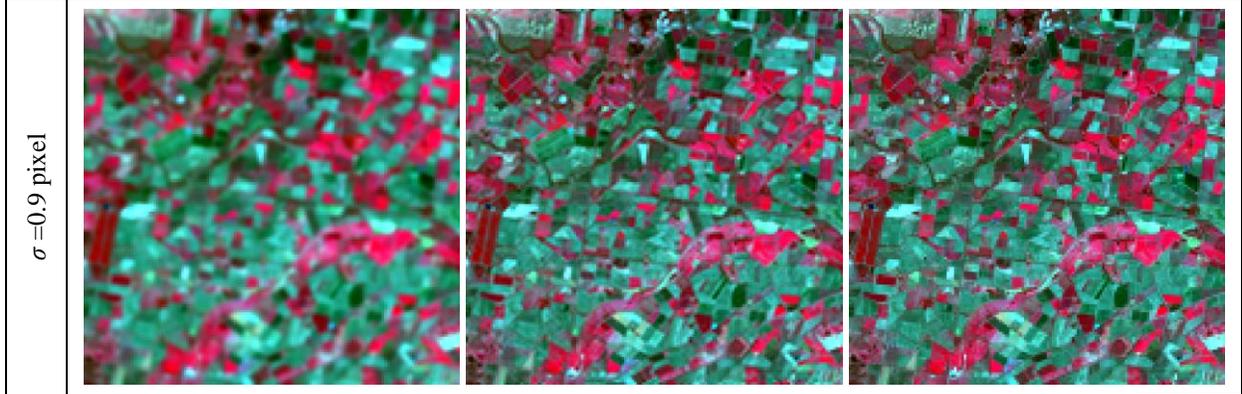|	Original data	Enhanced data	Reference
$\sigma=0.3$ pixel			
$\sigma=0.5$ pixel			
$\sigma=0.7$ pixel			
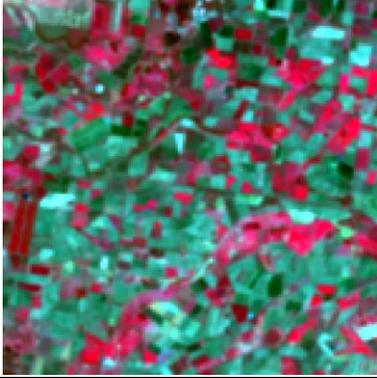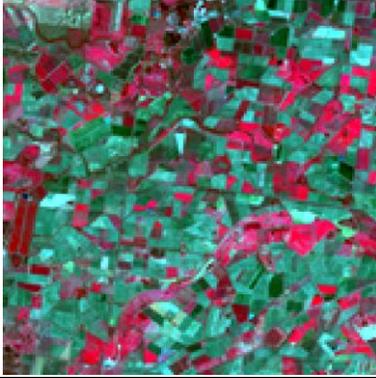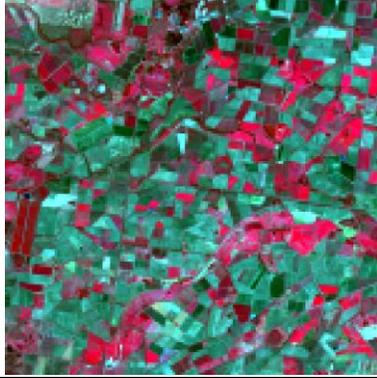
Fig. 2 The results (at 120 m) of the ATPK-based geostatistical filter in relation to different PSF parameters

Table 1 provides a quantitative assessment of the results of predicting the reference using the original 120 m PSF-contaminated images and using the proposed ATPK-based geostatistical filter. Generally, two observations can be made by checking the results.

First, as σ increases, the accuracies of both the original and filtered images decrease. More precisely, from σ=0.3 to σ=0.9 the CC of the original data decreases from 0.9990 to 0.9443, while the CC of the filtered data decreases from 0.9990 to 0.9869. This is because as σ increases, more sub-pixel neighbors contaminate the center pixel, leading to blurred images with a greater losss of information. Meanwhile, the uncertainty in reproduction increases as the inverse prediction problem becomes more complex and when more sub-pixel neighbors are involved.

Second, the filtered results have consistently greater accuracy than the original data. The advantages are more obvious for large σ. For example, the RMSE of the filtered result is 0.0001 (in units of reflectance) smaller than that of the original data for σ=0.3, but the gains increase steadily from 0.0034 to 0.0069 when σ increases from 0.5 to 0.9. For the RMSE, the corresponding reduction in remaining error (RRE) (Wang et al. 2015) of σ=0.5, 0.7 and 0.9 are 6.67%, 59.65%, 59.38% and 55.20%, respectively.

Table 1 Quantitative assessment of the predictions of the proposed filter in relation to different PSF parameters (the standard deviation σ is relative to a 120 m pixel)

	CC				RMSE (in units of reflectance)			
	σ=0.3	σ=0.5	σ=0.7	σ=0.9	σ=0.3	σ=0.5	σ=0.7	σ=0.9
Original	0.9990	0.9884	0.9672	0.9443	0.0015	0.0057	0.0096	0.0125
Filtered	0.9990	0.9976	0.9933	0.9869	0.0014	0.0023	0.0039	0.0056

3.1.2 Data Simulated with Different Spatial Resolutions

The Landsat image in Fig. 1 was upscaled with a Gaussian PSF (σ=0.5 pixel) and three different factors of 4, 8 and 12, creating three PSF-contaminated images with pixel sizes of 120 m, 240 m and 360 m. The results before and after filtering are shown in the first two columns of Fig. 3. The three reference images are created by upscaling Fig. 1 with the corresponding factor and ideal square wave filter. For blurred data at different spatial resolutions, the proposed method consistently can enhance the images and the predictions are visually closer to the reference.

The quantitative assessment results are shown in Table 2. The effect of the PSF is more obvious for data at coarser spatial resolution and the inconsistency with the reference is larger. The filtered predictions have a greater CC and smaller RMSE in relation to the reference, suggesting that the

data are enhanced using the filter. For example, the CC values of the filtered predictions are 0.0092, 0.0135, 0.0187 larger than those of the original data at 120 m, 240 m and 360 m, respectively.

Table 2 Quantitative assessment of the predictions of the proposed filter in relation to different spatial resolutions

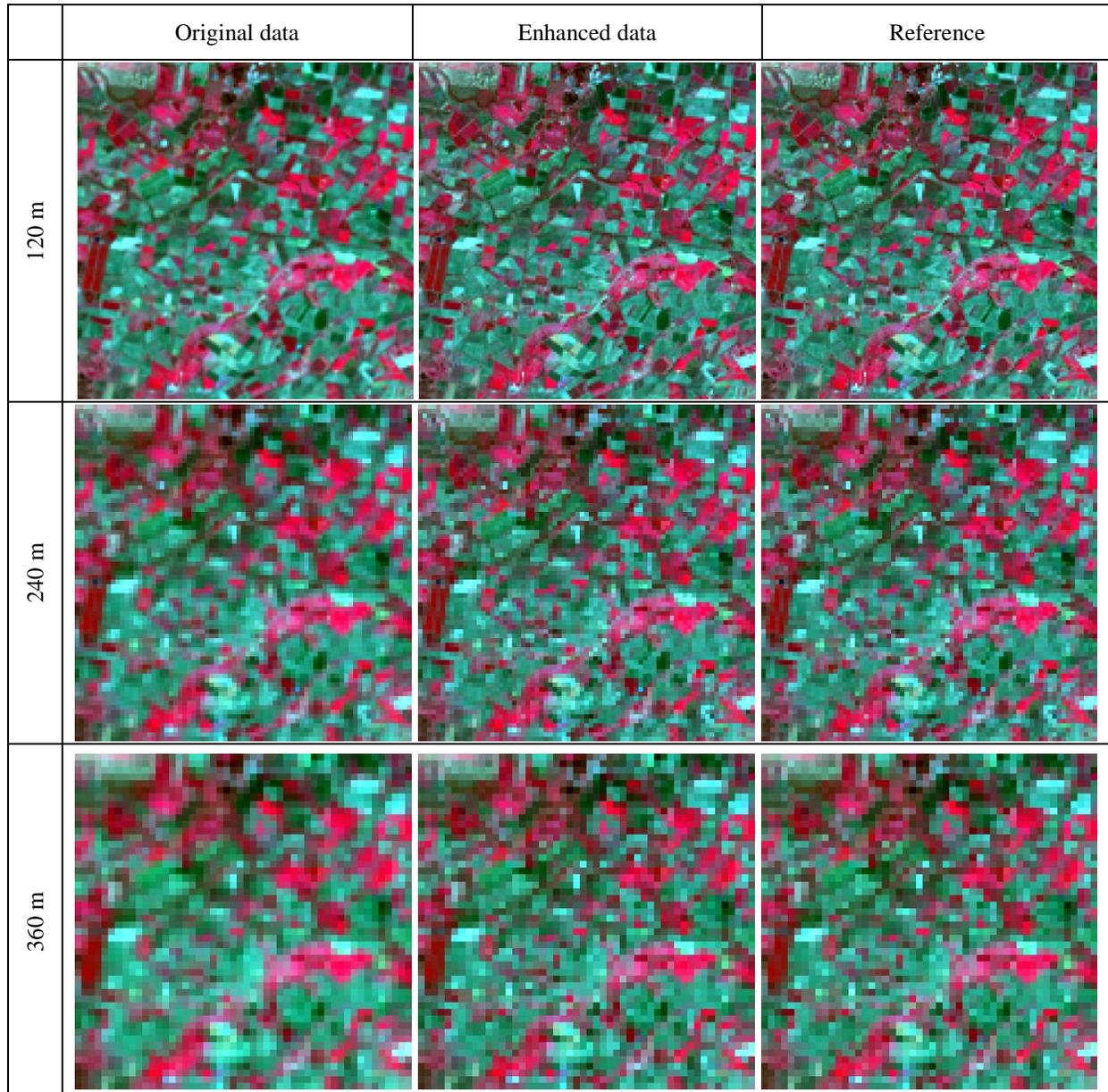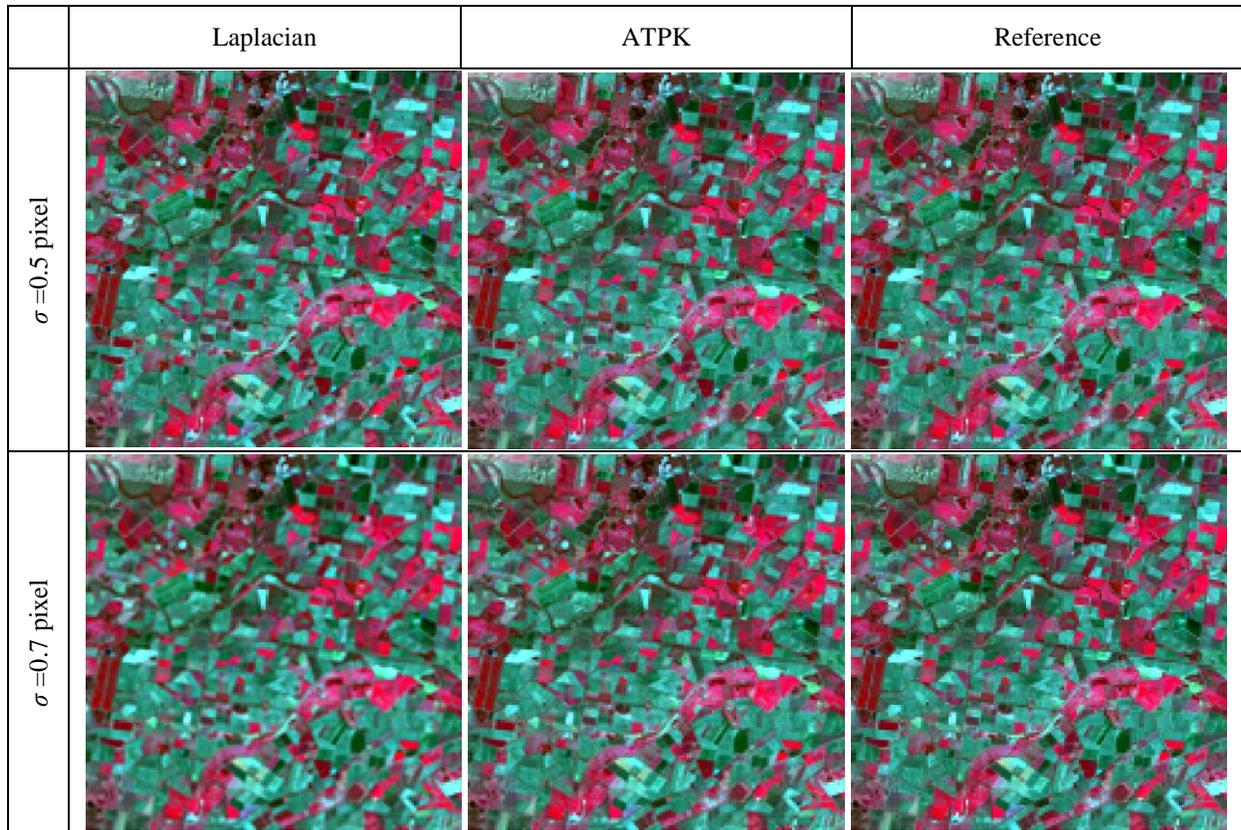|	CC			RMSE (in units of reflectance)		
	120 m	240 m	360 m	120 m	240 m	360 m
Original	0.9884	0.9826	0.9756	0.0057	0.0068	0.0077
Filtered	0.9976	0.9961	0.9943	0.0023	0.0028	0.0032

Fig. 3 The results ($\bar{\sigma}$=0.5 pixel) of the ATPK-based geostatistical filter in relation to images with different spatial resolutions

3.1.3 Comparison with Other Methods

The proposed method was compared with the classical UM method (using the popular Laplacian operator). The three cases of σ=0.5, 0.7 and 0.9 (upscaled with a factor of 4) were considered. The weight w in the Laplacian-based filter was set to 0.2, 0.4 and 0.7 for σ=0.5, 0.7 and 0.9, respectively, which was found to produce the greatest accuracy for the method after repeated testing. The results are shown in the first column of Fig. 4. Comparing to the original data in the first column of Fig. 2, the Laplacian-based method can produce visually clearer images. However, the predictions are still ambiguous to some extent, which is particularly obvious for the case of σ=0.9. Checking the predictions of the proposed method, it can be observed that they are visually clearer than those of the Laplacian-based method (the difference is more obvious for larger σ). The quantitative assessment results in Table 3 and the scatter plots in Fig. 5 also reveal that the proposed method is more accurate than the Laplacian-based method. To highlight the improvement of the proposed ATPK-based method over the benchmark methods, the RRE is also shown in Table 3. It is seen that the RREs are all above 35% for CC and around 30% for RMSE, suggesting that the increase in accuracy is obvious.
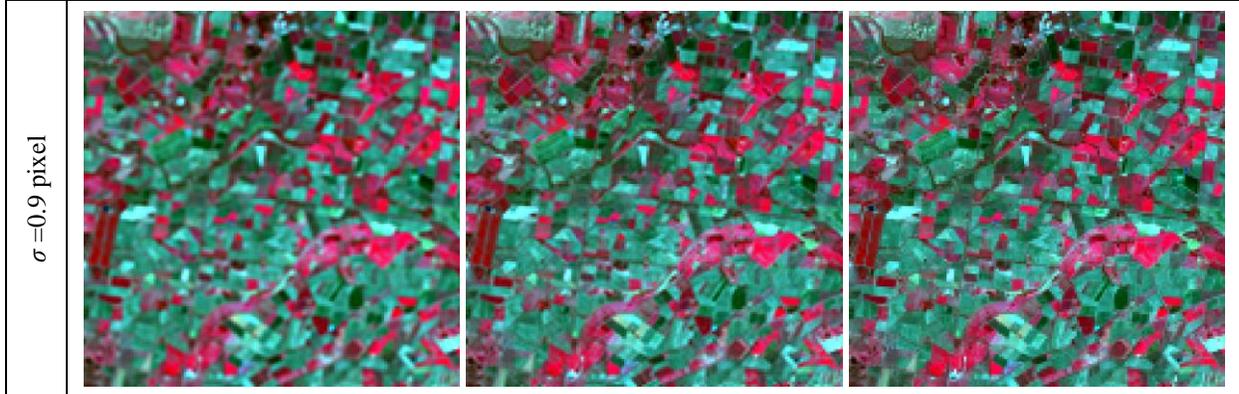
Fig. 4 The results (at 120 m) produced with different methods

Table 3 Quantitative assessment of the predictions produced with different methods (the standard deviation σ is relative to a 120 m pixel)

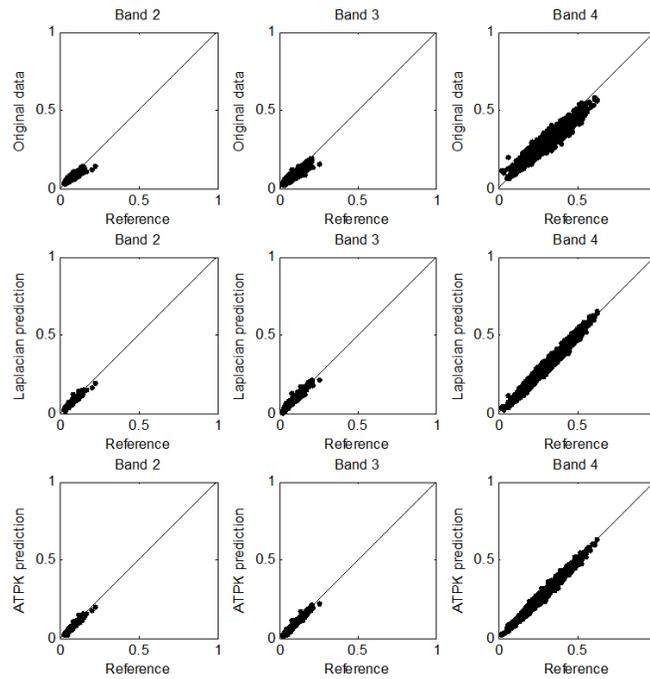|	CC			RMSE (in units of reflectance)		
	σ=0.5	σ=0.7	σ=0.9	σ=0.5	σ=0.7	σ=0.9
Original (RRE)	0.9884 (79.31%)	0.9672 (79.57%)	0.9443 (76.48%)	0.0057 (59.65%)	0.0096 (59.38%)	0.0125 (55.20%)
Laplacian (RRE)	0.9963 (35.14%)	0.9882 (43.22%)	0.9762 (44.96%)	0.0034 (32.35%)	0.0055 (29.09%)	0.0079 (29.11%)
ATPK	0.9976	0.9933	0.9869	0.0023	0.0039	0.0056

Fig. 5 Scatter plots (σ=0.7 pixel) between the predictions and reference data (at a spatial resolution of 120 m)

3.2 Experiment 2: Real Data

The proposed geostatistical filter was applied to enhance a real Landsat 7 ETM+ dataset in this experiment. The Landsat image (in units of digital number (DN)) was supplied by the Government of Canada through Natural Resources Canada, Earth Sciences Sector, Canada Centre for Remote

Sensing, which was used for pan-sharpening in our previous research (Wang et al. 2016). The image covers a 9 km by 9 km area (containing 300 by 300 pixels) in Alberta in Canada. The green, red, and near-infrared bands (i.e., bands 2, 3, and 4) with a spatial resolution of 30 m were considered in this experiment, as shown in the false color composite in Fig. 6.

Fig. 6 The Landsat 7 image (bands 4, 3, and 2 as RGB) selected for Experiment 2

The standard deviation of the Gaussian PSF σ was tuned from 0.1 to 0.9 with an interval of 0.1, and the proposed filter was performed for each parameter. The zoom factor in ATPK-based downscaling was empirically set to 4. The nine predictions are shown in Fig. 7. It was found that as σ increases, the prediction becomes clearer. When σ exceeds a certain value, however, the prediction contains more spurs and noise with further increases of σ. The optimal σ was determined finally as 0.4 after visual comparison of the nine predictions.

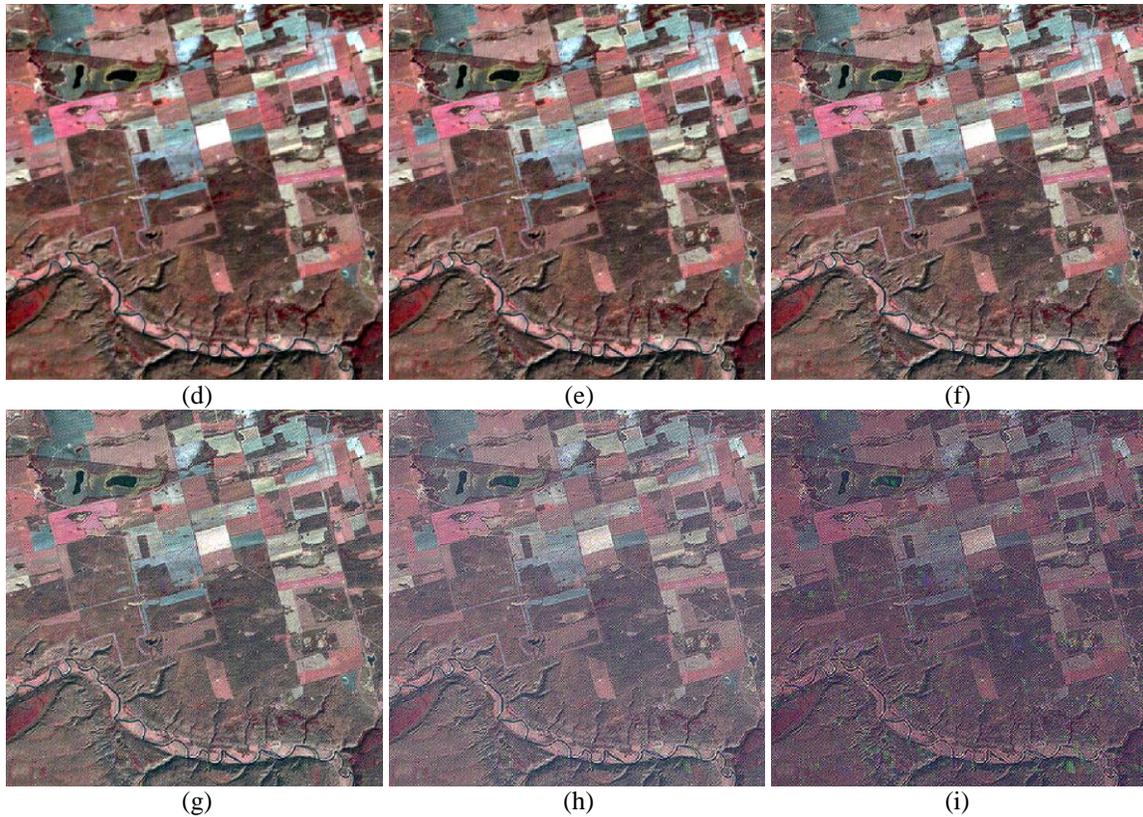(a)   (b)   (c)
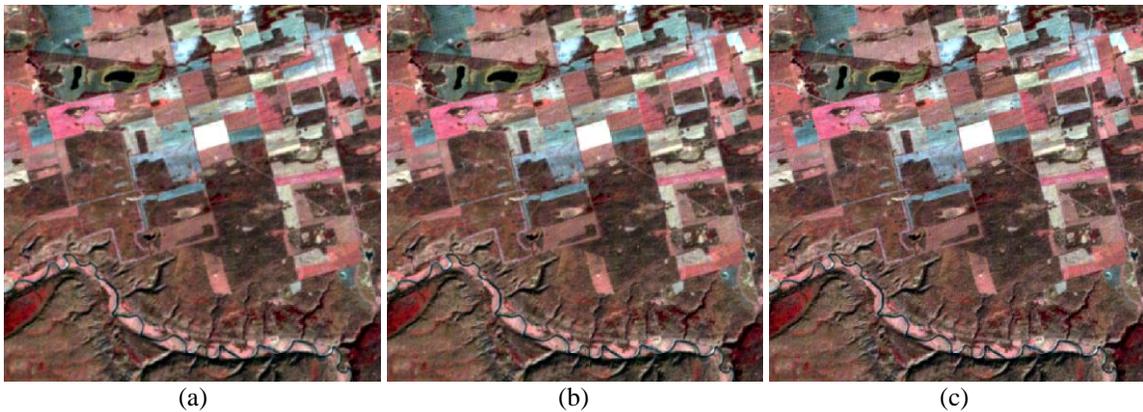
Fig. 7 The filtered results produced using the Gaussian PSF with nine different standard deviations. (a)-(i) σ=0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9

Figure 8 shows the differences between the filtered predictions and the original data based on the optimal parameter σ=0.4. It is seen clearly that the differences are located mainly on the boundaries between classes (see, for example, the elongated features for band 4 and rectangle features for all three bands), where mixed pixels exist and the PSF effect is prominent. The semivariogram was used to evaluate the prediction. Figure 9 shows the semivariograms (without any model fitting) of the original and filtered data. It can be seen that the filtered prediction exhibits greater spatial variation than the original data, suggesting that the filtered prediction increases the contrast between pixels (i.e., increases the sharpness). Table 4 shows the semivariogram (at a lag of one pixel) of the original and filtered data. The increases at a lag of one pixel for the three bands are 1.42, 3.13 and 3.53, respectively.

Table 4 The semivariogram (at a lag of one pixel, in units of DN^2) of the original and filtered data

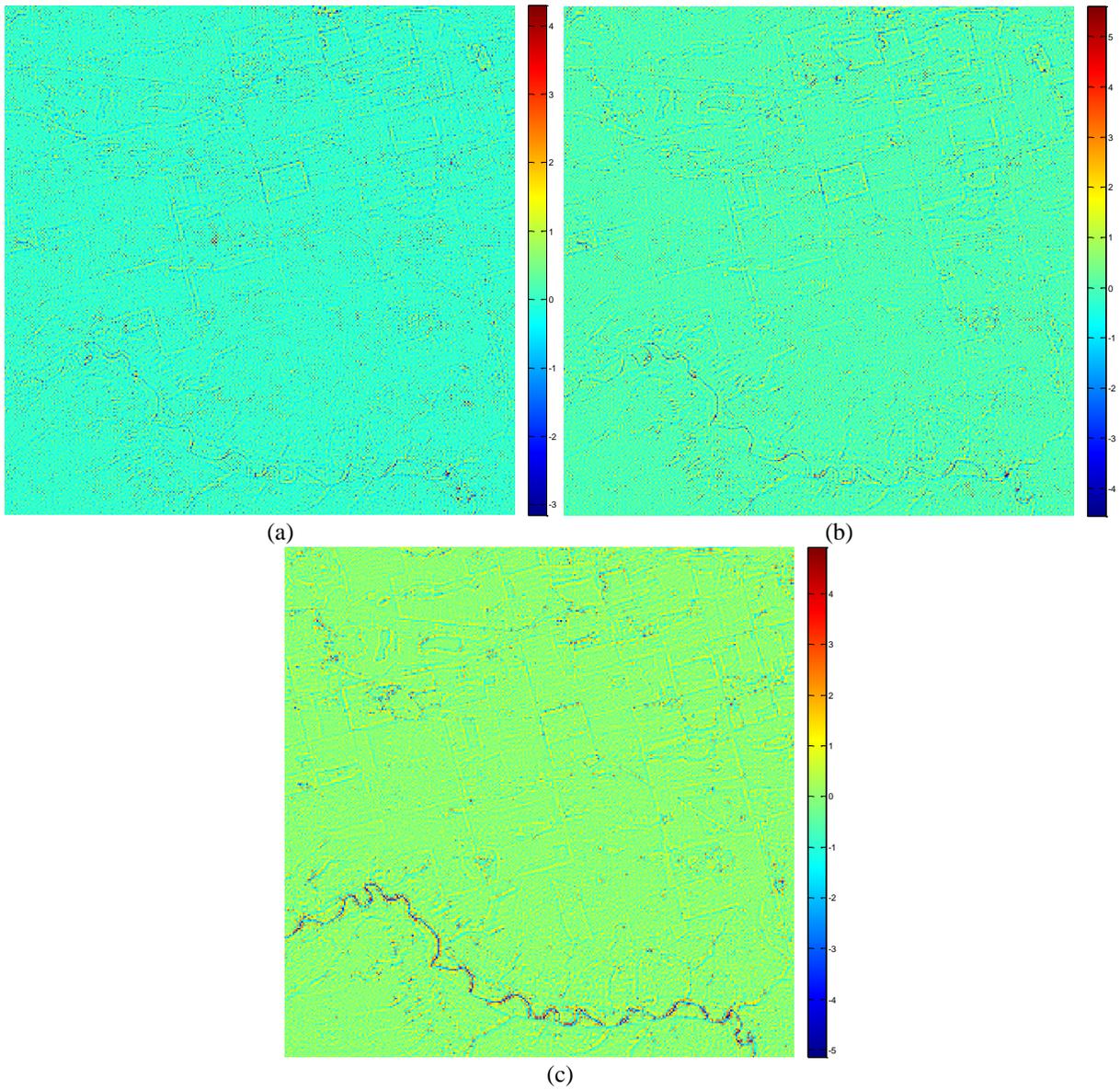|	Band 2	Band 3	Band 4
Original	2.34	5.92	9.22
Filtered	3.76	9.05	12.75

Fig. 8 The difference between the original and filtered data (σ=0.4). (a) Band 2. (b) Band 3. (c) Band 4
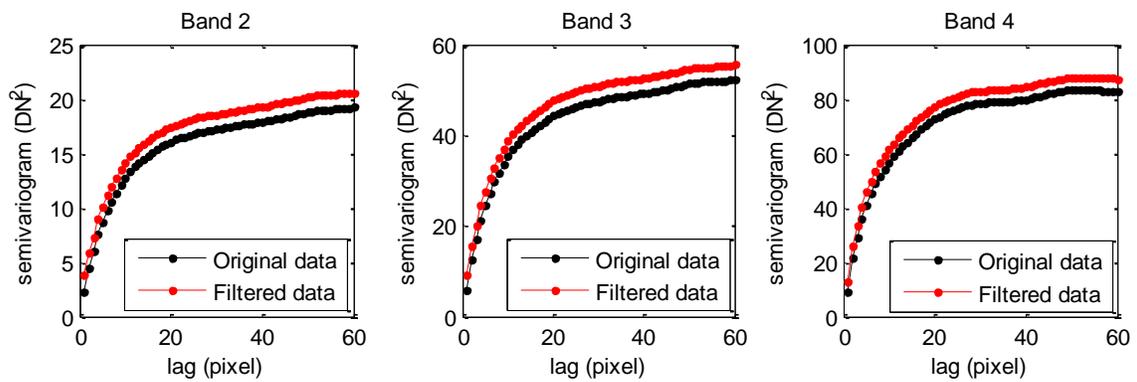

Fig. 9 The semivariograms of the original and filtered data

4 Discussion

Very few image enhancement methods consider how the neighbors affect the center pixel from the physical viewpoint of the PSF effect. This paper analyzes the mathematical relation between the signal of the center pixel and its neighbors. Such a relation is described at the sub-pixel level, and this focus on sub-pixels has seldom been considered in existing methods such as the classical UM method. The proposed downscaling-then-upscaling scheme can account explicitly for the form of the PSF of the sensor, and it is conceptually suitable for any PSF. Furthermore, the proposed method is applicable to any images and is expected to be applied to enhance images in other domains (e.g., medical imaging or computer vision) in the future.

The produced enhanced data have widespread application value, including *all* applications based on remote sensing image analysis. A crucial benefit is that after filtering, the boundaries between land cover and land use classes can be more clearly observed. This is potentially of great help for image interpretation, especially for images of heterogeneous landscapes which contain mixed pixels. Land cover classification and spectral unmixing are two of the most widely studied goals in remote sensing. As mentioned earlier in the Introduction, this paper proposed a general solution to enhance original remote sensing images contaminated by the PSF effect and, as such, the predictions can be used in many potential applications, not only spectral unmixing (Wang et al. 2018) or sub-pixel mapping (Wang and Atkinson 2017). In addition, compared to the solution to enhance spectral unmixing in Wang et al. (2018), another theoretical advantage of the solution in this paper is relaxing the assumption that the endmembers are scale-free (i.e., the same endmembers can be considered for spectra at different spatial resolutions in Wang et al. (2018)). Such an advantage may lead to more reliable spectral unmixing predictions, especially when the endmembers actually vary greatly across different scales and impose great uncertainty in the model developed, as in Wang et al. (2018).

In this paper, the ATPK approach was considered for the inverse prediction problem of predicting *error-free* sub-pixel values from observed pixel values. It should be noted that the inverse prediction problem is ill-posed, especially when the size of the PSF width is large and more neighbors impact the center pixel. This means that multiple solutions can meet the constraint in Eq. (1). For example, the implementation of ATPK involves the estimation of the point-to-point semivariogram (Wang et al. 2015, 2016). Different point-to-point semivariograms will finally lead to different predictions of sub-pixel values, but all of them strictly meet the constraint in Eq. (1) (Kyriakidis 2004; Kyriakidis and Yoo 2005). On the other hand, as acknowledged widely, downscaling becomes more challenging as the spatial heterogeneity of pixel values increases and downscaling might not even be pursued if the spatial heterogeneity is very high. Thus, the reliability of downscaling and the proposed geostatistical filter varies spatially depending on the local spatial heterogeneity. To reduce the uncertainty in downscaling-based sub-pixel value prediction, and the proposed geostatistical filter, it is necessary to seek more auxiliary information about the land cover patterns.

When characterizing the mathematical relation between the pixel and its neighbors due to the PSF effect, an interesting issue is the definition of the 'point'. Physically, the point is infinitely small when the sensor acquires signals from the Earth's surface. In the mathematical model in this paper (Eq. (1)), and in geostatistics generally, the 'point' is defined as the smaller unit (i.e., sub-pixel). In theory, to achieve the closest approximation of the PSF and coarse pixel space through sampling, the spatial size of the sub-pixel needs to be defined to be as small as possible, that is, approaching a point. However, the zoom factor will be unacceptably large in this case, and this will greatly increase the computational burden of ATPK-based downscaling. On the other hand,

to save computational cost, the spatial size of the sub-pixel needs to be defined to be as large as possible. However, the effect of neighbors will then be quantified less well due to coarse approximation of the true PSF, and the reliability of the model will be reduced. According to the experimental test in this paper, the largest spatial size of a sub-pixel was set to 1/16 of the coarse pixel (i.e., the minimum zoom factor was 4), which will result in an acceptable balance between the computational cost and reliability of the model. For example, in Experiment 1, the sub-pixel with a spatial resolution of 30 m was treated as a 'point' when enhancing the 120 m original PSF-contaminated data. However, there is still a need to develop methods to predict reliably the optimal spatial size of the 'point' in more cases.

Since the 'point' is geostatistically defined as a sub-pixel at a certain spatial resolution, the PSF in this case is actually changed to a set of sub-pixels in scale transformation (between pixel and sub-pixel), rather than the true PSF of the sensor (i.e., original measurement). There is a great need to develop methods to estimate the PSF, including the specific form of the function (not only Gaussian in this paper, but other forms (Tan et al. 2006)) and related parameters. Generally, two cases for PSF estimation can be expected.

    i)   If images are available at both of the spatial resolutions of the pixel and sub-pixel, a transfer function between the sub-pixel and pixel images, which approximates the PSF if the difference between the resolutions is large, may be estimated using the training image-pair. In using such an image-pair to estimate the PSF one consideration would be the temporal distance between the two images, and the spatial and temporal distance between the training data pair and the study area. The two images and two types of data need to be as close (both spatially and temporally) as possible for more reliable PSF estimation. Another consideration is the inconsistency (e.g., uncertainties in geometric registration and atmospheric correction) between the data at different spatial resolutions in the training image-pair.

    ii)   If only the observed blurred image is available, the PSF may be estimated by visual inspection of the predictions for all possible PSF candidates, as was done in Experiment 2 where the optimal parameter for the Gaussian PSF was determined by comparison of the nine predicted images. The benefit of the artificial visual interpretation process is that human perception, by definition, is more intelligent than a mathematically defined metric. The limitation, however, is that such a process is laborious, particularly when the number of data increases, and the risk of over-dependence on the experts' knowledge. Alternatively, the PSF can be estimated mathematically by focusing on some representative features in the images (e.g., linear features or straight boundaries) and investigating how the features can be degraded from features known prior.

In this paper, the semivariograms were characterized using a global model. That is, a single semivariogram was calculated for the entire study scene. This is reasonable for the image in the experiments, as the spatial coverage is relatively small and the choice of a stationary model is reasonable. For large application areas, however, the spatial structure may vary greatly spatially and a non-stationary model might be more appropriate. In this case, the uncertainty in semivariogram modeling and, further, in ATPK-based downscaling, as well as the proposed filter, would increase if a global model is fitted. As a solution to applying the proposed filter to large areas, one could develop local models, such as by dividing a large area into sub-areas, or by constructing a local window for semivariogram modeling for each pixel, if the expensive computational cost is tolerable.

5 Conclusion

This paper proposed a geostatistical filter to enhance remote sensing images by removing or reducing the PSF effect. The PSF effect is ubiquitous in remotely sensed images and it can be regarded as a nuisance; it adds a smoothing effect to the data that blurs the visual appearance of images and makes the data more difficult to use in relation to predicting scene components such as land cover and land use objects.

The new method explores the effect of neighbors on the pixel signal (i.e., gray value) *at the sub-pixel level* such that each pixel value is considered as a convolution of the *sub-pixel* values within the pixel and in the neighbors of the pixel. On this basis, a novel geostatistical filter method was proposed that involves downscaling-then-upscaling. The method first downscales the original image to sub-pixels while accounting for the PSF effect, and then upscales the downscaling predictions with a square wave filter (i.e., involving only sub-pixels within the pixel) to produce a predicted image without the PSF effect. The effectiveness of the method was demonstrated using two datesets.

References

Campagnolo ML, Montano EL (2014) Estimation of effective resolution for daily MODIS gridded surface reflectance products. IEEE Transactions on Geoscience and Remote Sensing 52: 5622–5632

Cao G, Zhao Y, Ni R, Kot AC (2011) Unsharp masking sharpening detection via overshoot artifacts analysis. IEEE Signal Processing Letters 18: 603–606

Gilboa G, Sochen N, Zeevi, YY (2002) Forward-and-backward diffusion processes for adaptive image enhancement and denoising. IEEE Transactions on Image Processing 11: 689–703

Farbiz F, Menhaj M B, Motamedi SA, Hagan MT (2000) A new fuzzy logic filter for image enhancement. IEEE Transactions on Systems, Man, and Cybernetics–Part B: Cybernetics 30: 110–119

Huang C, Townshend RG, Liang S, Kalluri SNV, DeFries RS (2002) Impact of sensor's point spread function on land cover characterization: assessment and deconvolution. Remote Sensing of Environment 80: 203–212

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

Kim S, Allebach JP (2005) Optimal unsharp mask for image sharpening and noise removal. Journal of Electronic Imaging 14: 023007–1

Kükenbrink D, Hueni A, Schneider FD, Damm A, Gastellu-Etchegorry J-P, Schaepman ME, Morsdorf F (2019) Mapping the irradiance field of a single tree: quantifying vegetation-induced adjacency effects. IEEE Transactions on Geoscience and Remote Sensing 57: 4994–5011

Kyriakidis PC (2004) A geostatistical framework for area-to-point spatial interpolation. Geographical Analysis 36: 259–289

Kyriakidis P, Yoo E-H (2005) Geostatistical prediction and simulation of point values from areal data. Geographical Analysis 37: 124–151

Manslow JF, Nixon MS (2002) On the ambiguity induced by a remote sensor's PSF. In Uncertainty in Remote Sensing and GIS: 37–57

Matz SC, de Figueiredo RJP (2006) A nonlinear image contrast sharpening approach based on Munsell's scale. IEEE Transactions on Image Processing 15: 900–909

Pardo-Iguzquiza E, Chica-Olmo M, Atkinson PM (2006) Downscaling cokriging for image sharpening. Remote Sensing of Environment 102: 86–98

Polesel A, Ramponi G, Mathews VJ (2000) Image enhancement via adaptive unsharp Masking. IEEE Transactions on Image Processing 9: 505–510

Ramponi G, Strobel N, Mitra SK, Yu T-H (1996) Nonlinear unsharp masking methods for image contrast enhancement. Journal of Electronic Imaging 5: 353–366

Tan B, Woodcock CE, Hu J, Zhang P, Ozdogan M, Huang D, Yang W, Knyazikhin Y, Myneni RB (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 RG, Huang C, Kalluri SNV, Defries RS, Liang S (2000) Beware of per-pixel characterization of land cover. International Journal of Remote Sensing 21: 839–843

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

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

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

Wang Q, Shi W, Atkinson PM (2016) Area-to-point regression kriging for pan-sharpening. ISPRS Journal of Photogrammetry and Remote Sensing 114: 151–165

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

Wang Q, Shi W, Atkinson PM (2018) Enhancing spectral unmixing by considering the point spread function effect. Spatial Statistics 28: 271–283

Wang X, Zhong Y, Zhang L, Xu Y (2019) Blind hyperspectral unmixing considering the adjacency effect. IEEE Transactions on Geoscience and Remote Sensing, in press, DOI: 10.1109/TGRS.2019.2907567

Wenny BN, Helder D, Hong J, Leigh L, Thome KJ, Reuter D (2015) Pre- and post-launch spatial quality of the Landsat 8 Thermal Infrared Sensor. Remote Sensing 7: 1962–1980

Zhang B, Allebach JP (2008) Adaptive bilateral filter for sharpness enhancement and noise removal. IEEE Transactions on Image Processing 17: 664–678

Zheng X, Li Z, Zhang X, Shang G (2019) Quantification of the adjacency effect on measurements in the thermal infrared region. IEEE Transactions on Geoscience and Remote Sensing, in press, DOI: 10.1109/TGRS.2019.2928525