A Large-Scale Benchmark Data Set for Evaluating Pansharpening Performance

Overview and implementation

XIANGCHAO MENG, YIMING XIONG, FENG SHAO, HUANFENG SHEN, WEIWEI SUN, GANG YANG, QIANGQIANG YUAN, RANDI FU, AND HONGYAN ZHANG

Pansharpening aims to sharpen a low-spatial-resolution (LR) multispectral (MS) image using a high-spatial-resolution (HR) panchromatic (Pan) image to obtain the HR MS image. It has been a fundamental and active research topic in remote sensing, and pansharpening methods have been developed for nearly 40 years. While the performance evaluation of pansharpening methods is still based on a small number of individual images, data-driven pansharpening approaches are attracting increasing attention. However, few publicly available benchmark data sets for pansharpening are available, especially large-scale ones. This has been a serious limitation for the future development of pansharpening methods.

In this article, we propose a large-scale data set for pansharpening. The data set has 2,270 pairs of HR Pan/LR MS images from different kinds of remote sensing satellites, including 200 *Ikonos*, 500 *QuickBird*, 410 *Gaofen* (*GF*)*-1*, 500 *WorldView-4*, 500 *WorldView-2*, and 160 *WorldView-3* images. In addition, we construct thematic remote sensing data sets, including ones that are typical of urban areas, green vegetation, and water scenarios, with 510, 258, and 318 pairs of HR Pan/LR MS images, respectively. An unlabeled data set containing 1,184 pairs of HR Pan/LR MS images with mixed surface features is also constructed. Traditional pansharpening methods are reviewed along with the current approach, and the performance of popular pansharpening methods is evaluated and statistically analyzed based on the proposed data set.

BACKGROUND

Due to the limitations of radiation energy, data volumes, and similar elements of remote sensing satellite-imaging system, the acquired images require a fundamental tradeoff between spatial and spectral resolutions [1]–[4]. For example, most of the remote sensing satellites provide bundled LR MS images and HR Pan images, rather than an HR MS image. Fortunately, pansharpening (typically referred to as *sharpening MS by Pan*), i.e., Pan/MS fusion, can overcome this hardware limitation to obtain the HR MS image. Pansharpening methods originated during the 1980s [5]–[7]. Since 1986, the *Système Pour l'Observation de la Terre-1* (*SPOT-1*) system has provided LR MS images with HR Pan images, so pansharpening methods have had a rapid development.

Digital Object Identifier 10.1109/MGRS.2020.2976696
Date of current version: 24 April 2020

This development can be divided into three major stages motivated by several factors.

1) *1980s*: A pansharpening theory was proposed, and several popular methods were developed. Representative methods include traditional intensity-hue-saturation (IHS) fusion [6], [8] and high-pass filter (HPF) fusion [7], [9].
2) *1990s*: The concept of image fusion was critically discussed [10]–[12], and multiresolution analysis (MRA)-based pansharpening methods, especially wavelet-based ones [13]–[16], attracted growing attention. In addition, with the increase in developed methods, review and comparison of different pansharpening methods arose [11], [17], [18].
3) *Blowout stage*: This occurred after 2000, especially since Tu et al. [19] proposed a general fusion framework that made a great contribution to the improvement and development of pansharpening methods. During this period, many state-of-the-art methods based on new theories continued to spring up, such as model-based methods [20], [21] and sparse-representation-based approaches [22], [23]. Pansharpening methods based on deep learning (DL) [24]–[27] are attracting attention.

To the best of our knowledge, there are several development factors [1], including advances in remote sensing satellite sensors. These advances mainly focus on the variation of the number of spectral bands and the difference of the spectral range between MS and Pan images. Specifically, this variation has evolved from previous MS images that had only three bands and Pan images that covered only the visible spectrum (such as *SPOT-1* and *SPOT-2*) to MS with four bands and Pan covering the visible and near-infrared spectrum (such as *Ikonos*, *QuickBird*, *GF-1*) and MS with six or more bands only partly covered by the Pan image (such as *Landsat Enhanced Thematic Mapper Plus*, *WorldView-2*, and *WorldView-3*).

Second, there are relevant newly emerging theories and hot-spot mathematical research, such as sparse-representation-based and DL-based pansharpening methods. The development is further motivated by the demands of practical engineering and remote sensing applications. For example, effective pansharpening methods to obtain HR MS images for change detection [28], [29], thematic mapping [30], [31], and so forth are highly desirable. It should be noted that different applications may have contrasting requirements for more spectral fidelity and spatial enhancement. Hence, application-oriented pansharpening methods should be given more attention.

To date, large numbers of pansharpening methods have been developed, and they have been classified in several ways. In [3] and [32], existing pansharpening methods were classified into two major categories, i.e., component substitution (CS)-based methods and MRA-based ones. In addition, Zhang et al. [33] classified them into CS-, MRA-, and Bayesian-based methods, while Shen et al. [2] organized them into CS-, MRA-, sparse reconstruction (SR)-, and model-based optimization (MBO) techniques. It should be noted that the major fusion process of the Bayesian-, SR-, and MBO-based methods is built on or converted to the optimization of a variational model; hence, these approaches can be generalized into the variational optimization (VO)-based methods. Garzelli [34] first performed a comprehensive review of VO-related methods based on the super-resolution concept. Kwan et al. [35], [36] classified the existing pansharpening methods based on whether the point spread function was used or not.

In this article, based on the previous categories, pansharpening methods are classified into four groups: CS, MRA, VO, and DL. Among them, the CS- and MRA-based methods have been widely used due to their simplicity and high efficiency. The CS-based methods, in particular, seem to be favored by professional remote sensing software, such as Environment for Visualizing Images (ENVI), Earth Resources Data Analysis System Imagine, and PCI Geomatica.

1) CS-based pansharpening methods are established with the substitution of the LR component of the MS by the HR Pan to obtain the fused HR MS image. The LR component is generally obtained by spectral transformation of the MS image. The development of CS-based pansharpening methods includes two stages: 1) the traditional understanding characterized by the process of forward projection transformation–CS–inverse projection transformation and 2) the general understanding featuring a unified fusion framework without forward and backward transformation. The representative CS-based methods include the classical principal component analysis (PCA) method [17], IHS method [17], [37], and Gram-Schmidt (GS) method [38], [39].
2) MRA-based pansharpening methods are founded on the spatial decomposition of the HR Pan image by the wavelet transform or Laplacian pyramids to extract the high spatial structures, which are injected into the interpolated LR MS image to obtain the HR MS image. Similar to CS-based methods, MRA-based approaches have developed from the traditional to the general understanding. The traditional understanding is strictly based on MRA algorithms, which are characterized by the three-step process of multiresolution decomposition–fusion–image reconstruction. The general understanding is characterized by a unified fusion framework without complex three-step processing. The representative methods include the popular additive wavelet luminance proportional (AWLP) method [3], [40] and generalized Laplacian pyramid (GLP)-based methods [3], [41], [42].
3) VO-based pansharpening methods are built on variational theory, and the major process is generally based on or converted to the optimization of an energy functional. This category was proposed during the 1990s [43]; after Ballester et al. [20] proposed the variational pansharpening method called P+XS, VO-based pansharpening methods attracted ever-increasing attention and underwent rapid development. Model-based

methods [20], [21], [44] and sparse-based methods [22], [23] are two representative VO-related techniques.

4) DL-based pansharpening methods have attracted growing attention in recent years, and they can be regarded as a new generation. The basic idea is to learn a network between the fused image and the observations based on DL theory, with all the parameters automatically learned under the supervision of training samples. To the best of our knowledge, Huang et al. [24] proposed the first DL-based pansharpening method, and several works were subsequently developed [26], [45], [46]. However, compared with CS-, MRA-, and VO-based methods, DL-based approaches generally require more data to train a robust network, and this is one of the limitations for their development.

On the whole, pansharpening methods have existed for nearly 40 years, and many studies [3], [11], [17], [47]–[60] have made great contributions to their improvement and development. However, there is a serious issue that should be given close attention, i.e., the lack of a specific large-scale benchmark data set. On the one hand, to more effectively and objectively validate the performance of pansharpening methods, experiments should be performed with large amounts of data, rather than one or two images, which may constitute subjective selection. Most existing studies assess pansharpening methods based on only a few images, due to the lack of a large-scale benchmark data set. Although [1] reviewed and discussed the performance of the pansharpening methods from a perspective of meta-analysis, which was performed by making a statistical analysis of all the large-scale studies ever published to avoid the limitation of the data sets, it was still restricted by the strict selection of satisfactory articles to validate a specific approach. On the other hand, data-driven pansharpening methods, such as DL-based ones, are attracting more attention, but few opening benchmark data sets are available, especially large-scale data ones. This has been a serious limitation for the development of pansharpening methods.

In this article, we present a large-scale benchmark data set for pansharpening, including 2,270 pairs of Pan/MS images from several remote sensing satellites, with 200 *Ikonos*, 500 *QuickBird*, 410 *GF-1*, 500 *WorldView-4*, 500 *WorldView-2*, and 160 *WorldView-3* images. Pansharpening methods are reviewed, and the performance of popular approaches is evaluated based on the proposed data set. Large-scale thematic remote sensing data sets, including urban areas, green vegetation, and water scenarios, for pansharpening are constructed, as is a large-scale unlabeled data set with mixed surface features. In summary, the major contributions of this article are as follows:

1) We construct a large-scale pansharpening data set for different remote sensing satellites and typical thematic surface features. To the best of our knowledge, it is the first large-scale data set for pansharpening, and it can provide the research community a better resource to evaluate and advance state-of-the-art pansharpening algorithms.
2) We review the pansharpening methods, from the traditional understanding to the current general understanding, including CS-, MRA-, VO-, and DL-based approaches.
3) We present a performance evaluation of pansharpening methods based on the large-scale data set, rather than one or two few images, from a perspective of statistical analysis. This can more sufficiently reflect the shortcomings and advantages of existing pansharpening methods.
4) The proposed large-scale data set to the community. It will be a helpful data set for researchers, especially those who have difficulty obtaining very-high-resolution remote sensing images in pansharpening.

REVIEW OF PANSHARPENING METHODS

COMPONENT SUBSTITUTION-BASED METHODS

CS-based pansharpening methods are founded on the substitution of an LR component of the MS image by the HR Pan image to obtain the fused HR MS image. As shown in Figure 1, the development of CS-based pansharpening methods includes two stages: the traditional understanding and the new understanding based on general formalization.

For the traditional understanding, CS-based pansharpening methods are characterized by complex three-step processing with forward projection transformation–CS–inverse projection transformation. Therefore, CS-based methods in the traditional understanding are also referred to as *projection-substitution* methods [58]. However, in this case, development and improvement of CS-based methods are limited by the rigid projection transformation [17], [39], [48], such as IHS and PCA. Compared to the traditional understanding, the general understanding featuring a unified fusion framework is simpler and more flexible. The fused image is obtained based on the injection of the high spatial structures of the HR Pan into a resampled MS image. These high spatial structures are generally obtained by the difference between the HR Pan and LR component, which is found through the linear combination of the MS spectral bands. The general understanding is represented as

$$\hat{\mathbf{M}} = \tilde{\mathbf{M}} + \mathbf{g}(\mathbf{P} - \mathbf{I}_L), \tag{1}$$

where $\hat{\mathbf{M}}$ is the fused HR MS image; $\tilde{\mathbf{M}}$ is the resampled LR MS image; $\mathbf{I}_L$ denotes the component to be substituted; $\mathbf{P}$ denotes the HR Pan image, which is generally normalized (e.g., histogram matching) with $\mathbf{I}_L$ to reduce spectral distortion; and $\mathbf{g}$ represents the injection weight of the high spatial structures.

The new understanding with the unified framework emancipates the CS-based methods from complex processing with rigid projection transformations. This not only

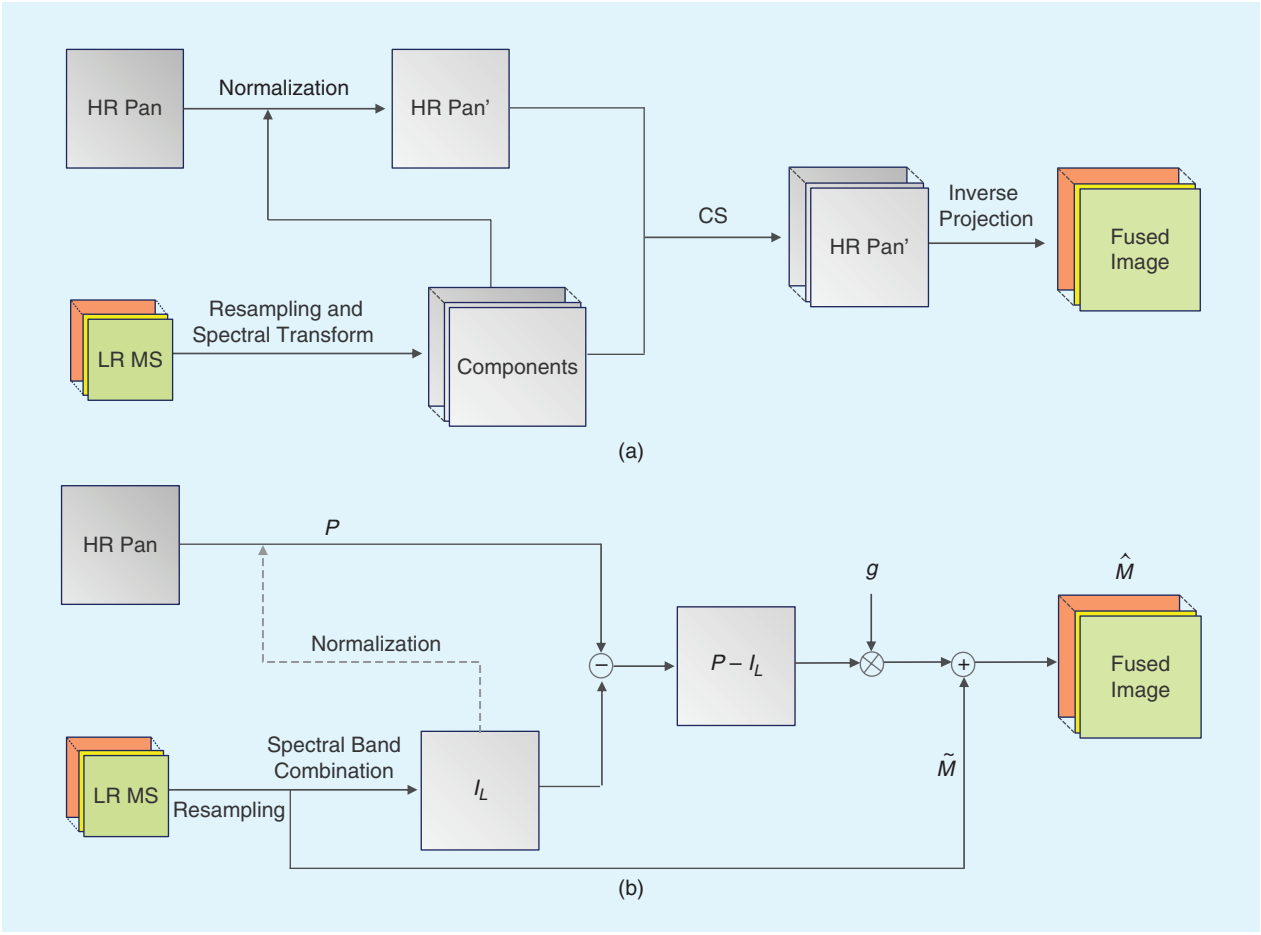

FIGURE 1. CS-based pansharpening methods. (a) The traditional understanding of the CS-based pansharpening methods. (b) The new understanding of these methods based on the general formalization.

promotes a faster implementation but opens new horizons for improvement and development. Correspondingly, a number of state-of-the-art CS-based pansharpening methods were developed on the unified framework, such as the adaptive GS (GSA) [38], band-dependent spatial detail (BDSD) [61], and partial-replacement adaptive CS (PRACS) approaches [62].

MULTIRESOLUTION ANALYSIS-BASED METHODS

MRA-based pansharpening methods are founded on the injection of the high frequency, which is extracted by spatial filters (such as the wavelet transform [14] and Laplacian pyramids [63]) from the HR Pan image into the resampled LR MS image to obtain the fused HR MS image. As shown in Figure 2, similar to CS-based methods, MRA-based approaches have been developed from the traditional understanding to the new general understanding.

The traditional understanding of the MRA-based methods [64] features complex three-step processing with multiresolution decomposition–fusion–image reconstruction. It is extended to the Amélioration de la Résolution Spatiale par Injection de Structures [65] to highlight that the purpose is to preserve the whole content of the LR MS image and add the high–spatial–structure information of the HR Pan image [3], [48], [66]. MRA-based methods are further extended to the current understanding with a unified fusion framework [3], [59], [64], represented as

$$\hat{\mathbf{M}} = \tilde{\mathbf{M}} + \mathbf{g}(\mathbf{P} - \mathbf{P}_L), \tag{2}$$

where $\mathbf{P}_L$ is the low-pass version of the HR Pan image. It is shown that the new understanding liberates MRA-based pansharpening methods from complex three-step processing and mainly focuses on effective spatial-structure extraction and injection-weight calculation. Correspondingly, a number of state-of-the-art pansharpening methods have been proposed, such as the popular additive à trous wavelet transform (ATWT) [3], [15], [67], AWLP [40], [49], and modulation transfer function (MTF)-GLP [41], [42].

On the whole, the CS- and MRA-based methods under the general understanding first extract the high spatial structures of the HR Pan image and inject them into the resampled LR MS image to obtain the fused HR MS image. In comparison, the difference between them mainly depends on how to extract the high spatial structures from

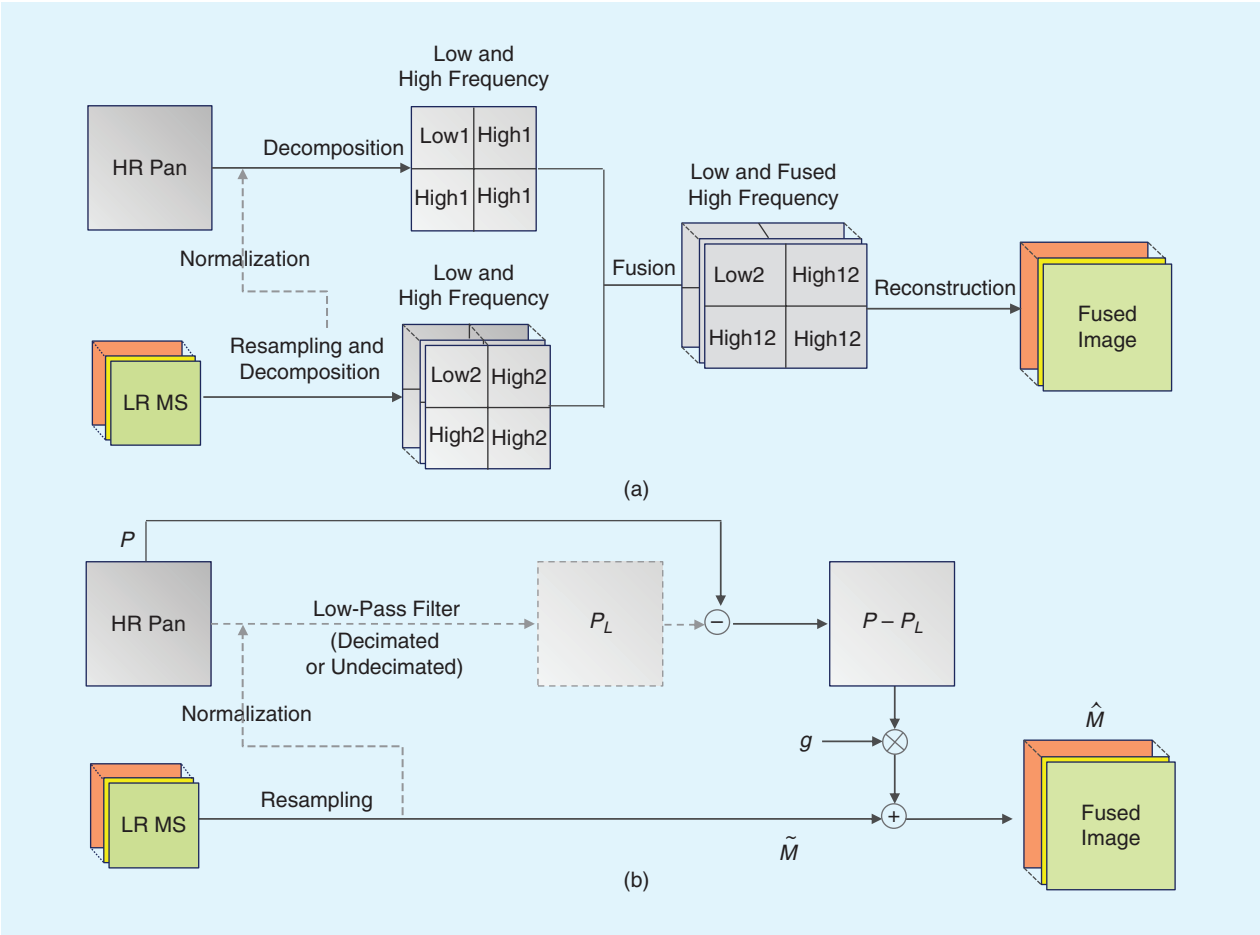

FIGURE 2. MRA-based pansharpening methods. (a) The traditional understanding of the MRA-based pansharpening methods. (b) The new understanding of these methods based on the general formalization.

the HR Pan image. For CS-based methods, the high spatial structures are extracted based on the subtraction of the LR component (obtained by the spectral band combination of the MS image) from the HR Pan image. For MRA-based methods, the high spatial structures are extracted from the HR Pan image alone, based on spatial filters. Therefore, the CS-based methods under the general understanding are also called *spectral methods*, and the MRA-based methods are referred to as *spatial methods* [49].

VARIATIONAL OPTIMIZATION-BASED METHODS

VO-based methods are established on variational theory. To the best of our knowledge, the idea of the VO-based methods was proposed during the 1990s [43]; since Ballester et al. [20] proposed the method called *P+XS*, they have attracted ever-increasing attention and experienced rapid development. A schematic of VO-based methods is presented in Figure 3. It is shown that VO-based methods feature two major parts: 1) the construction of an energy functional and 2) the optimization solution of the constructed energy functional. For the construction of the energy functional, the methods based on the observation model [20], [43], [44], [68]–[71] and sparse-representation theory [22], [23], [72], [73] are the most popular. The model-based methods regard the fusion process as a typical ill-posed inverse problem, and the energy functional is generally constructed from the observation model by considering the imaging process of the observations [21], [74], [75], represented as follows:

$$E(\mathbf{x}) = f_{\text{spectral}}(\mathbf{x}, \text{LR MS}) + f_{\text{spatial}}(\mathbf{x}, \text{HR Pan}) + f_{\text{prior}}(\mathbf{x}), \quad (3)$$

where $\mathbf{x}$ denotes the ideal fused image.

As shown in (3), the energy functional generally contains three terms: spectral fidelity, spatial enhancement, and the prior. The spectral fidelity term is constructed from the relationship between the ideal HR MS and the LR MS image [2], [44], [74]–[76] to preserve the spectral information. The spatial-enhancement term is constructed based on the relationship between the ideal HR MS and the HR Pan. There are generally two ways: 1) the spatial structure similarity between the HR MS and the HR Pan [77]–[79] and 2) the spectral band combination relationship between the narrow HR MS bands and the wide HR Pan band [44], [74], [80]. The prior term is constructed to impose constraints on the ideal fused image, such as the popular

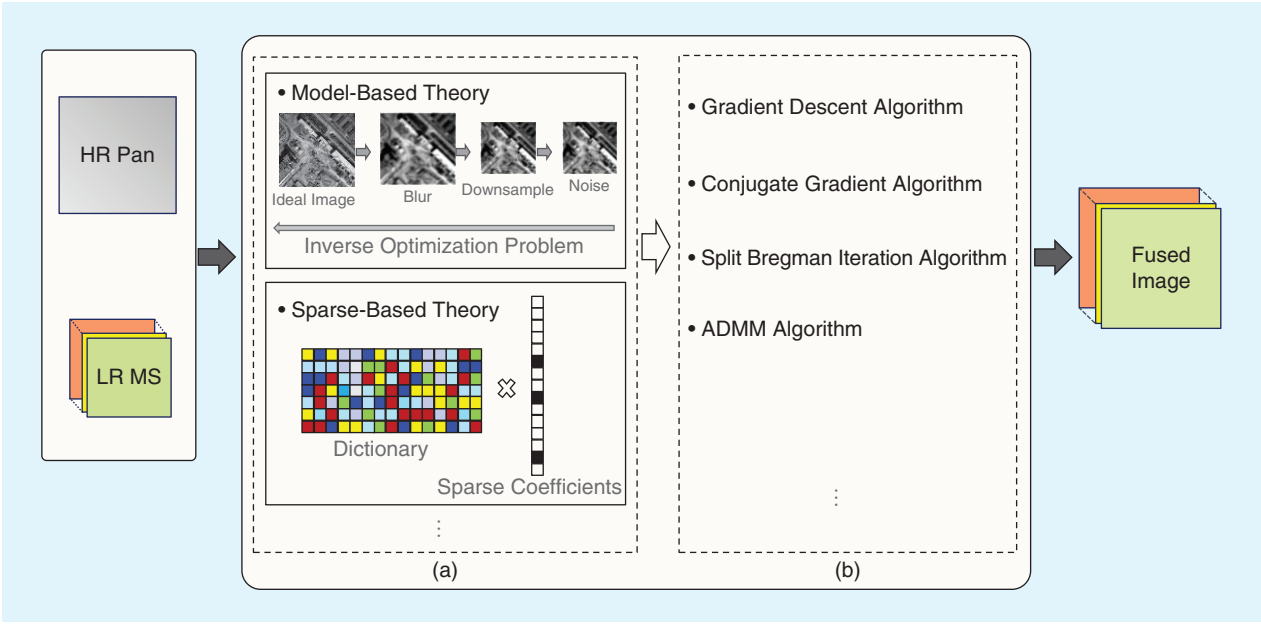

FIGURE 3. VO-based pansharpening methods. (a) The energy-functional construction. (b) The optimization solution. ADMM: alternating-direction method of multipliers.

Huber-Markov prior [44] and the total variation prior [71]. The sparse-based methods are founded on the sparse-representation theory [81], [82], and the first work was proposed by Li and Yang [22]. The sparse-based methods are mainly characterized by dictionary learning, and this generally includes offline dictionary learning [22], [73], which relies on an external database, and online dictionary learning [23], [72], [83] directly from the images to be fused.

The optimization method is generally based on iterative optimization algorithms [2], [84], such as the conjugate gradient [2] and alternating-direction method of multipliers [76]. However, these iterative optimization methods are generally time consuming, especially for images in large dimensions, and this seriously hinders the engineering application of VO-based pansharpening methods.

DEEP-LEARNING-BASED METHODS

In addition to the previous three categories, DL-based pansharpening methods have attracted growing attention during recent years. They are based on the hot-spot DL theory [85]–[87], featuring deep neural networks (DNNs) with more than two layers [87]. DL-based pansharpening methods assume that the relationship between the observations and the ideal fused HR MS image is complex and nonlinear. State-of-the-art DNNs with significant nonlinear representational power are then utilized to learn the network between the ideal fused image and the observations, with all parameters automatically being learned under the supervision of large-scale training samples. A schematic of typical DL-based pansharpening methods is given in Figure 4.

As shown in Figure 4, typical DL-based pansharpening methods generally include two-step processing: 1) network training and 2) obtaining the fused image based on the pretrained network. Among them, network training is relatively more important, and it is performed to learn the transformation relationship from the observations to the ideal fused HR MS images. It should be noted that practical ideal HR MS images generally do not exist. Therefore, training samples are typically obtained based on Wald's protocol [88]; i.e., the original MS and Pan images are spatially degraded by the spatial-resolution ratio r between the HR Pan and LR MS images, and the original MS images are then regarded as the target HR MS images. During network training, convolutional neural networks (CNNs) are generally utilized, and a number of parameters $\{W, b\}$ for all the convolutional layers should be learned from the training samples. This is realized by the minimization of the loss function $\mathrm{Diff}(\mathrm{Ref}_{\mathrm{MS}}, F_{\mathrm{MS}})$, which is constructed based on the difference between the output of the network F_{MS} and the target label image $\mathrm{Ref}_{\mathrm{MS}}$.

To the best of our knowledge, Huang et al. [24] proposed the first DL-based pansharpening method. They constructed the fused HR MS image from the LR MS image based on the pretrained network, which was learned from the HR/LR Pan images by the DNN. Masi et al. [26] proposed a DL method based on a CNN borrowed from the super-resolution CNN proposed in [89]. For network training, the HR Pan image was regarded as a new band to be overlaid on the resampled LR MS image to train the network. It was further improved with the assistance of typical nonlinear radiometric indices calculated from the MS image, such as the normalized-difference water index and normalized-difference vegetation index. Zhong et al. [46] proposed performing super-resolution of the LR MS image based first on a CNN; the

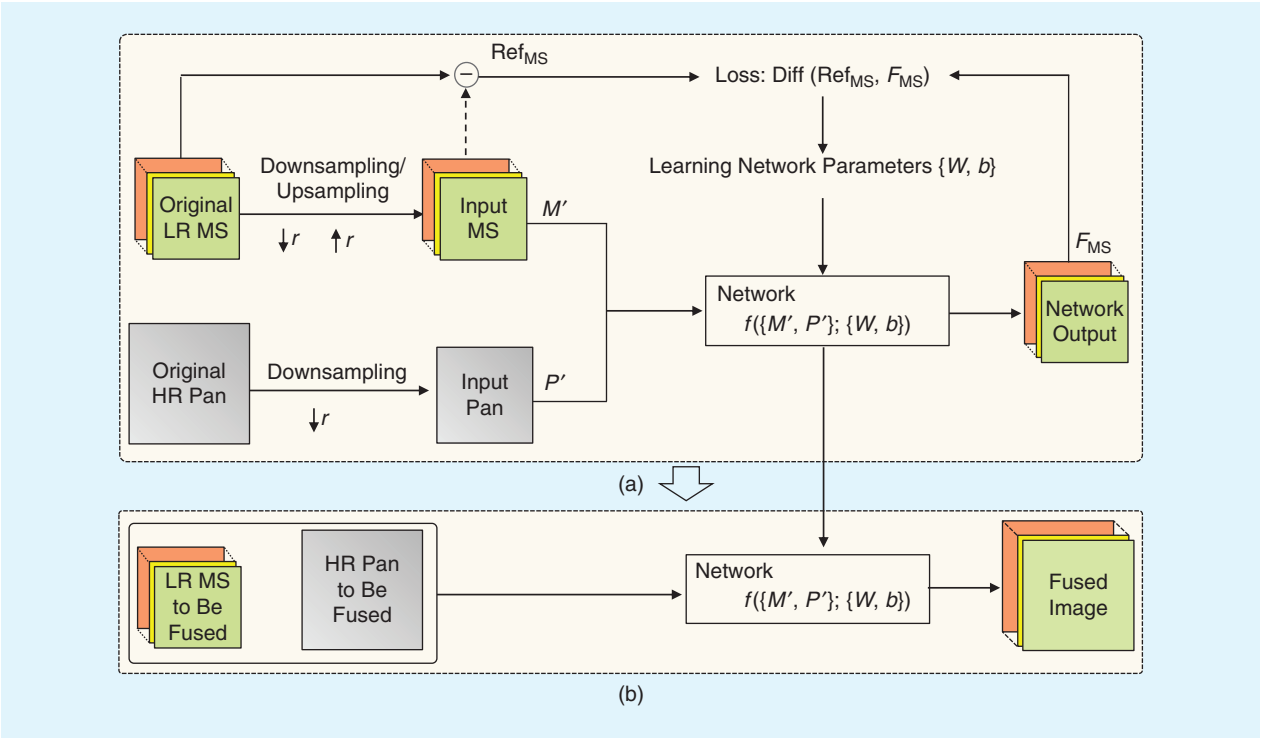

FIGURE 4. Typical DL-based pansharpening methods. (a) Networking training based on the Wald protocol. (b) Obtaining the fused image based on the pretrained network.

enhanced MS image and HR Pan image were further fused based on the GS pansharpening method.

The preceding CNN-based pansharpening methods are based on simple and flat networks with only three convolutional layers. This relatively shallow architecture is generally regarded as insufficient to take full advantage of the strong nonlinear representative ability of DL models. Therefore, most studies [27], [45], [90]–[94] advocated deeper networks to improve performance. Accordingly, Yuan et al. [27], [45], Rao et al. [92], Yang et al. [94], and Li et al. [90] proposed DL-based pansharpening methods founded on the residual learning network [95]. In addition, Scarpa et al. [96] proposed a robust target-adaptive CNN-based pansharpening method by considering the presence of a mismatch with respect to the training set and across different sensors.

PROPOSED DATA SET FOR PANSHARPENING

To advance the state of the art in pansharpening, a large-scale data set is constructed in this article. (The proposed data set is available at http://www.escience.cn/people/fshao/database.html.) It consists of 2,270 pairs of HR Pan and LR MS images. For the data set construction, geometrical registration was performed, and the HR Pan and LR MS images were cropped for convenient usage. In addition, all the images in the data set were rearranged and grouped based on remote sensing satellite types and thematic surface features, respectively. Samples of the proposed data set are shown in Figure 5, and details are displayed in Table 1.

DATA SET GROUPED BY REMOTE SENSING SATELLITES

The proposed data set was grouped by remote sensing satellite types, consisting of 200 pairs of *Ikonos* Pan/MS images, 500 pairs of *QuickBird* Pan/MS images, 410 pairs of *GF-1* Pan/MS images, 500 pairs of *WorldView-4* Pan/MS images, 500 pairs of *WorldView-2* Pan/MS images, and 160 pairs of *WorldView-3* Pan/MS images. The Pan images have a spatial dimension of 1,024 × 1,024, and the MS images have a spatial dimension of 256 × 256. In the spectral dimension, the MS images from *Ikonos*, *QuickBird*, *GF-1*, and *WorldView-4* have four bands, and the *WorldView-2* and *WorldView-3* MS images have eight.

DATA SET GROUPED BY THEMATIC SURFACE FEATURES

The proposed data set was regrouped by typical thematic surface features. They include the urban thematic data set with 510 pairs of Pan/MS images, green vegetation thematic data set with 258 pairs, and water-scenario thematic data set with 318 pairs. In addition, a large-scale unlabeled data set containing 1,184 pairs of Pan/MS images with mixed surface features was constructed.

1) *Urban data set*: The images in this data set are mainly located in urban regions, and buildings are the focus. In addition, trees, roads, and so forth may exist, which is inevitable for remote sensing images of city areas. The HR Pan images have a dimension of 1,024 × 1,024 × 1, and the LR MS images have a dimension of 256 × 256

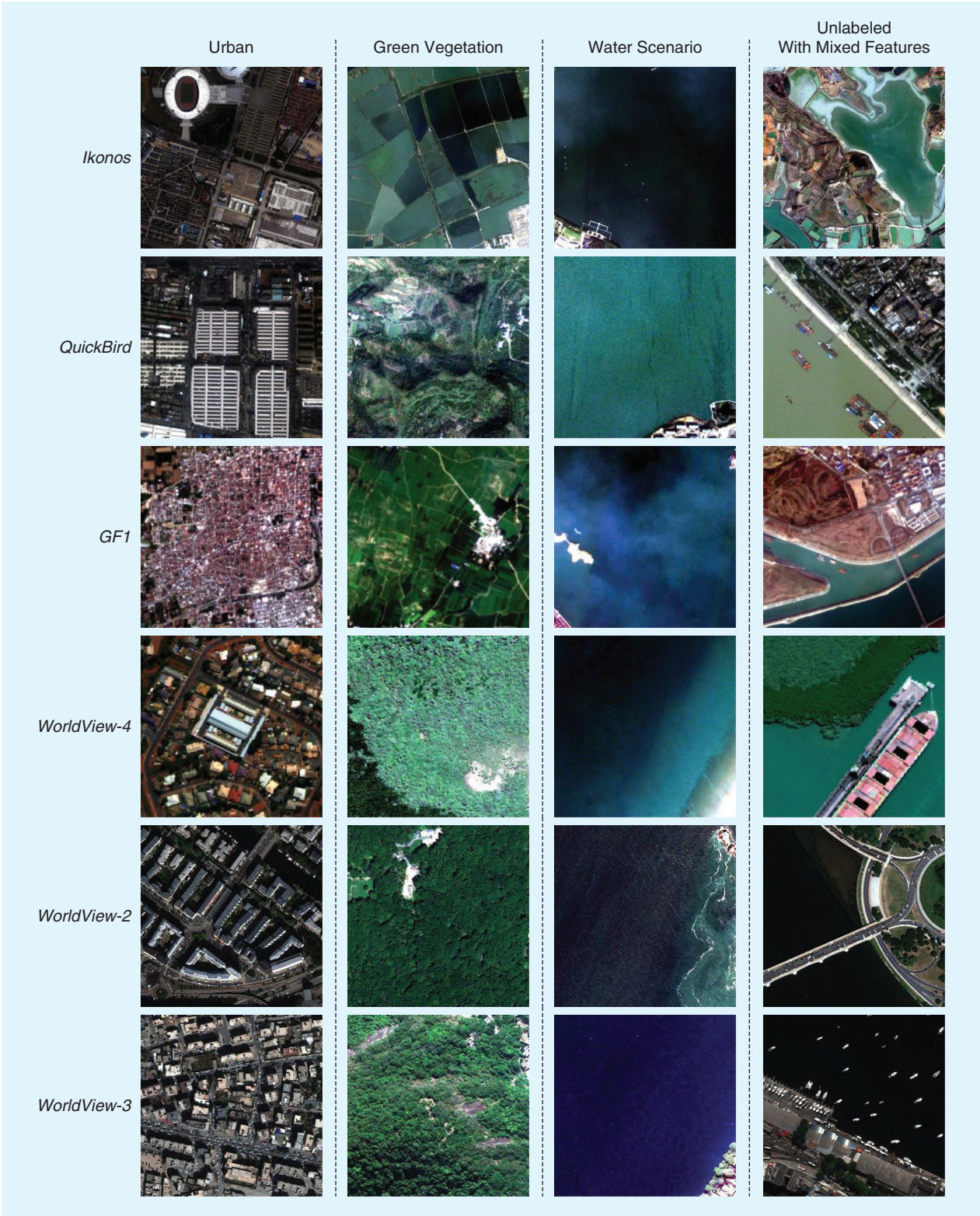

FIGURE 5. Samples of the proposed data set.

$\times$ 4 (*Ikonos*, *QuickBird*, *GF-1*, and *WorldView-4*) and 256 $\times$ 256 $\times$ 8 (*WorldView-2* and *WorldView-3*).

2) *Green vegetation data set*: The images in this data set take green vegetation as the principal part. They are generally located in suburbs, mountainous areas, and rural locations, such as farmland. The HR Pan images have a dimension of 1,024 $\times$ 1,024 $\times$ 1, and the LR MS images have a dimension of 256 $\times$ 256 $\times$ 4 (*Ikonos*, *QuickBird*, *GF-1*, and *WorldView-4*) and 256 $\times$ 256 $\times$ 8 (*WorldView-2* and *WorldView-3*).

3) *Water-scenario data set*: This data set is mainly constructed with water-surface features as the main body. For these images, boats and other ground features may exist. Like the urban and green vegetation data sets, the HR Pan images have a dimension of 1,024 × 1,024 × 1, and the LR MS images have a dimension of 256 × 256 × 4 (*Ikonos*, *QuickBird*, *GF-1*, and *WorldView-4*) and 256 × 256 × 8 (*WorldView-2* and *WorldView-3*).
4) *Unlabeled data set*: The images in this data set consist of different surface features, of which two or more different kinds are generally the main parts. Compared with the three previous thematic data sets, the images in the unlabeled data set have more complex ground features, such as the mixed attributes of large areas of vegetation and water as well as locations that have buildings, water, and soil. The HR Pan images have a dimension of 1,024 × 1,024 × 1, and the LR MS images have a dimension of 256 × 256 × 4 (*Ikonos*, *QuickBird*, *GF-1*, and *WorldView-4*) and 256 × 256 × 8 (*WorldView-2* and *WorldView-3*).

EXPERIMENTS

STATISTICAL ANALYSIS

In this article, the performance of the pansharpening methods is evaluated and statistically analyzed based on the proposed large-scale data sets for different remote sensing satellites and typical thematic remote sensing images. In the experiments, large numbers of results for each satellite data set and thematic data set were statistically analyzed rather than depending on one or two images. Therefore, a 95% confidence interval was applied to provide reliable statistical results, represented as $O_u \pm 1.96\rho(SE)$, with $\rho(SE) = \sqrt{O_{\text{var}}}/\sqrt{N}$ denoting the standard error. O_u is the mean value, while O_{var} denotes the variance, and N indicates the amount of data for analysis. The mean value reflects the overall accuracy, and the variance value denotes the robustness of the pansharpening methods. To ensure the reliability of the statistical results, the maximum and minimum values were removed before analysis. This is feasible for the statistical analysis of large numbers of data [97].

EXPERIMENT SETTINGS

To effectively verify the performance of the pansharpening methods, experiments were performed using the data set of different remote sensing satellite images, including four- and eight-band images; the data set of thematic images with the typical urban area, green vegetation, and water scenario; and the unlabeled data set with mixed features. Reduced- and full-resolution experiments were performed. The reduced-resolution experiment was based on Wald's protocol [88]; i.e., the original Pan and MS images were spatially degraded through low-pass filtering using Gaussian-MTF blurring and then downsampled by the spatial-resolution ratio between them. The original MS image was regarded as the reference image for evaluation. The popular quantitative evaluation indices of the *Erreur Relative Globale Adimensionnelle de Synthèse* (ERGAS) [88], spectra mapper angle (SAM) [98], and $Q2^n$ index [99] were employed. In the full-resolution experiment, due to the lack of a reference image, the popular nonreference-quality evaluation index of the quality with no reference (QNR) [100] was used. The quantitative-evaluation indices are shown in Table 2.

As shown in Table 2, $\mathbf{z}$ denotes the reference image, and $\hat{\mathbf{x}}$ denotes the fused image to be evaluated, while r is the spatial-resolution ratio, L_1L_2 is the spatial dimension, and B is the total number of MS spectral bands. The symbol σ denotes the standard variance, and μ indicates the mean value. $\sigma_{z\hat{x}}$ denotes the covariance of $\mathbf{z}$ and $\hat{\mathbf{x}}$, D_λ in the QNR assesses the spectral distortion, and D_s in the QNR assesses the spatial distortion, while $\mathbf{P}_{LP}$ denotes the low-resolution Pan image. The universal image-quality index proposed in [101] calculates the similarity between two images. The parameters of p, q, α, and β are generally set to one.

In the experiments, the performance of 19 pansharpening methods, as shown in Table 3, was evaluated. For the DL-based multiscale, multidepth CNN (MSDCNN) method, Wald's protocol was used for the network training, and the default parameters were applied [45]. Images that were 70% randomly selected were used to train the network, and the residual 30% of the images was utilized to validate the method for each data set. To ensure objectivity and equality, the same images were applied to the validation of the other 18 pansharpening methods.

EXPERIMENTAL RESULTS BASED ON DATA SETS FOR DIFFERENT SATELLITES

The experiment results based on data sets for different remote sensing satellites, i.e., *Ikonos*, *QuickBird*, *GF-1*, *WorldView-4*, *WorldView-2*, and *WorldView-3*, are shown in Figures 6–11 and Tables 4–9, respectively. For each remote sensing satellite, the statistical results based on reduced- and full-resolution experiments are shown. For each figure, the results with the ERGAS, SAM, $Q2^n$ index in the reduced-resolution experiment are shown in parts (a)–(c), respectively, and the results with D_λ, D_s, and the QNR in the full-resolution experiment are shown in parts (d)–(f), respectively. For Figures 6–11, the horizontal axis represents different pansharpening methods, and the vertical axis denotes the statistical quantitative-evaluation results by a mean value and the error bars for the 95% confidence interval. In addition, the top five results for each figure are marked with a yellow box. Tables 4–9 provide details. The top five results for each quantitative-evaluation index are marked in bold, respectively. In addition, it should be noted that, if three indices of a method are marked in bold, the intention is to indicate excellent performance.

As shown in Figure 6 and Table 4, the BDSD, GSA, MTF-GLP, MTF-GLP-high-pass modulation (HPM)-post processing (PP), MTF-GLP-context-based decision (CBD), and MSDCNN methods show satisfactory performance in

TABLE 1. DETAILS OF THE PROPOSED DATA SET.

SATELLITE SENSORS		SPATIAL DIMENSION	SPECTRAL DIMENSION	DIMENSION SIZE	NUMBER OF THEMATIC SCENE IMAGES				DATA VOLUME
					URBAN	GREEN VEGETATION	WATER SCENARIO	UNLABELED WITH MIXED FEATURES	
Ikonos	Pan	1 m	One band	1,024 × 1,024	60	8	6	126	200
	MS	4 m	Four band	256 × 256 × 4					
QuickBird	Pan	0.61 m	One band	1,024 × 1,024	150	20	42	288	500
	MS	2.44 m	Four band	256 × 256 × 4					
GF1	Pan	2 m	One band	1,024 × 1,024	5	90	10	305	410
	MS	8 m	Four band	256 × 256 × 4					
WorldView-4	Pan	0.31 m	One band	1, 024 × 1,024	90	85	95	230	500
	MS	1.24 m	Four band	256 × 256 × 4					
WorldView-2	Pan	0.5 m	One band	1,024 × 1,024	150	35	145	170	500
	MS	2 m	Eight band	256 × 256 × 8					
WorldView-3	Pan	0.31 m	One band	1,024 × 1,024	55	20	20	65	160
	MS	1.24 m	Eight band	256 × 256 × 8					

TABLE 2. THE QUANTITATIVE-EVALUATION INDICES.

EXPERIMENT GROUP	EVALUATION INDEX	DEFINITION	MEANING
Reduced-resolution experiment	ERGAS	$\text{ERGAS} = \frac{100}{r}\sqrt{\frac{1}{B}\sum_{b=1}^{B}\frac{\|\mathbf{z}_b - \hat{\mathbf{x}}_b\|^2 / L_1 L_2}{\mu_{\mathbf{z}_b}^2}}$	The smaller the better
	SAM	$\text{SAM} = \frac{1}{L_1 L_2}\sum_{i=1}^{L_1 L_2}\arccos\frac{\sum_{b=1}^{B}(\mathbf{z}_{i,b}\cdot\hat{\mathbf{x}}_{i,b})}{\sqrt{\sum_{b=1}^{B}\mathbf{z}_{i,b}^2}\sqrt{\sum_{b=1}^{B}\hat{\mathbf{x}}_{i,b}^2}}$	The smaller the better
	$Q2^n$ index	$Q2^n{}_{D\times D} = \frac{\sigma_{\mathbf{z}\hat{\mathbf{x}}}}{\sigma_{\mathbf{z}}\sigma_{\hat{\mathbf{x}}}}\cdot\frac{2\sigma_{\mathbf{z}}\sigma_{\hat{\mathbf{x}}}}{\sigma_{\mathbf{z}}^2+\sigma_{\hat{\mathbf{x}}}^2}\cdot\frac{2\|\mu_{\mathbf{z}}\|\cdot\|\mu_{\hat{\mathbf{x}}}\|}{\|\mu_{\mathbf{z}}\|^2+\|\mu_{\hat{\mathbf{x}}}\|^2}$ $Q2^n - \text{index} = \text{E}\left[\left\| Q2^n{}_{D\times D}\right\|\right]$	The bigger the better
Full-resolution experiment	D_λ	$D_\lambda = \sqrt[p]{\frac{1}{B(B-1)}\sum_{b=1}^{B}\sum_{j=1,j\neq b}^{B}\|UIQI(\hat{\mathbf{x}}_b,\hat{\mathbf{x}}_j) - UIQI(MS_b,MS_j)\|^p}$	The smaller the better
	D_s	$D_s = \sqrt[q]{\frac{1}{B}\sum_{b=1}^{B}\|UIQI(\hat{\mathbf{x}}_b,P) - UIQI(\text{MS}_b,P_{LP})\|^q}$	The smaller the better
	QNR	$\text{QNR} = (1-D_\lambda)^\alpha(1-D_s)^\beta$	The bigger the better

the *Ikonos* experiment. Among them, the BDSD method has excellent results in the reduced- and full-resolution experiments. Although the GSA, MTF-GLP, and MTF-GLP-HPM-PP methods show satisfactory performance in the reduced-resolution experiment, their performance in the full-resolution experiment is slightly poorer. The performance of the DL-based MSDCNN method in the reduced- and full-resolution experiments is robust.

In the *QuickBird* experiment, the BDSD, PRACS, MTF-GLP, MTF-GLP-HPM-PP, MTF-GLP-HPM, MTF-GLP -CBD, and MSDCNN methods show overall satisfactory performance. Among them, the MTF-GLP and MTF-GLP-HPM-PP methods show good performance in the reduced-resolution experiment and relatively poorer performance in the full-resolution experiment. Conversely, the MTF-GLP-HPM and MTF-GLP-CBD methods show good performance in the full-resolution experiment and poor performance in the reduced-resolution experiment. The BDSD and MSDCNN methods show excellent and robust performance in the reduced- and full-resolution experiments.

For the experiment results based on *GF-1* images, the BDSD, ATWT-model 2 (M2), ATWT-model 3 (M3), MTF-GLP-HPM, MTF-GLP-CBD, and MSDCNN methods show overall satisfactory performance. While the best performance is shown for the MTF-GLP-CBD method in the full-resolution experiment, the approach had a relatively poorer performance in the reduced-resolution experiment, especially in the ERGAS and SAM quantitative-evaluation indices. The DL-based MSDCNN method shows the best performance and has competitive advantages over the other pansharpening methods in the reduced-resolution experiment. In addition, it shows robust performance in the full-resolution experiment.

In the *WorldView-4* experiment, the BDSD, PRACS, MTF-GLP, MTF-GLP-HPM, MTF-GLP-CBD, and MSDCNN

TABLE 3. THE PANSHARPENING METHODS IN THE EXPERIMENTS.

METHOD	BASIC MEANING AND PARAMETER SETTINGS
Generalized IHS [3], [19]	A generalization of the traditional IHS pansharpening method [17] for the fusion of more than three MS bands; it is also called the *fast generalized IHS fusion method*
Brovey [3], [102]	A CS-based method under the general understanding; it has been commercialized in professional software, such as ENVI
PCA [3], [17]	A typical CS-based method founded on the PCA transformation; it has been commercialized in professional software, such as ENVI
BDSD [3], [61]	A data-dependent, self-adaptive CS-based approach featuring two advantages: 1) the optimal solution of the spatial-detail extraction and injection based on the minimum mean square error and 2) parameter estimation based on local regions
GS [3], [39]	A typical CS-based method built on the GS transformation; it has been commercialized in professional software, such as ENVI
GSA [3], [38]	An improved GS method in terms of the adaptive calculation of the spectral band combination weight based on a linear-regression algorithm
PRACS [3], [62]	A CS-based method founded on the partial replacement of the intensity component; the Pan is simulated based on the weighted average of the Pan and MS bands instead of directly using the original Pan to replace the intensity component
HPF [3], [17]	A typical MRA-based method; in the experiments, the default parameter with a 5 × 5 box filter was used
Smoothing filter-based intensity modulation (SFIM) [3], [103], [104]	An enhanced HPF method based on the spatial-detail injection scheme of HPM, which is calculated based on the ratio of the resampled MS image and the low-pass filtering Pan image
Indusion [3], [105]	Indusion combines induction and fusion; it is a fusion method based on the induction-scaling technique and profits from multiple equalization steps to improve performance
MTF-GLP [3], [41]	A popular MRA-based method founded on the GLP with an MTF filter; the unitary detail-injection model is applied
MTF-GLP-HPM [3], [106]	The MTF-GLP with the spatial-detail injection scheme of the HPM
MTF-GLP-HPM-PP [3], [107]	The MTF-GLP-HPM followed by a PP phase [107] to correct the noise at the edge regions
MTF-GLP-CBD [3], [50]	The MTF-GLP with the spatial-detail injection scheme of the CBD, which is locally optimized by patching the image in nonoverlapping zones
ATWT [3], [15], [67]	An MRA-based method built on the ATWT with a unitary injection model
ATWT-M2 [3], [65]	The ATWT with model 2 proposed in [65], which is based on moment matching to adjust the wavelet coefficients of the Pan image
ATWT-M3 [3], [65]	The ATWT with model 3 proposed in [65], which is based on least-square fitting to adjust the wavelet coefficients of the Pan image
AWLP [3], [40]	A generalization of the AWL pansharpening method [15] in terms of a proportional spatial-detail injection weight, relying on the original MS-band radiance proportionality
MSDCNN [45]	A DL-based method founded on an MSDCNN

methods show overall satisfactory performance. Among them, the MSDCNN method has excellent experiment results in the reduced- and full resolution experiments, especially in the quantitative-evaluation indices of the ERGAS, SAM, D_s, and QNR. The MTF-GLP-HPM method shows good performance in the full-resolution experiment and relatively poorer performance in the reduced-resolution experiment.

For the eight-band *WorldView-2* and *WorldView-3* satellite images, the overall performance is relatively poorer than the four-band *Ikonos*, *QuickBird*, *GF-1*, and *WorldView-4* satellite images. This can be seen in most of the quantitative-evaluation indices, such as the ERGAS, SAM, and $Q2^n$. On the one hand, the PRACS, MTF-GLP, MTF-GLP-HPM, MTF-GLP-CBD, and MSDCNN methods show overall satisfactory performance for the *WorldView-2* and *WorldView-3* images. In addition, the BDSD method shows a good result in the *WorldView-2* experiment. On the other hand, the PRACS, MTF-GLP, and MTF-GLP-HPM methods show inconsistent performance in the reduced- and full-resolution experiments for the *WorldView-2* and *WorldView-3* images. Particularly, the PRACS and MTF-GLP-HPM methods have excellent performance in terms of the full-resolution quantitative-evaluation indices and relatively poorer performance in the reduced-resolution experiment. Conversely, the MTF-GLP method shows a satisfactory result in the reduced-resolution experiment and relatively poorer performance in the full-resolution experiment. The MTF-GLP-CBD method shows competitive results in the reduced- and full-resolution experiments for the *WorldView-2* and *WorldView-3* images.

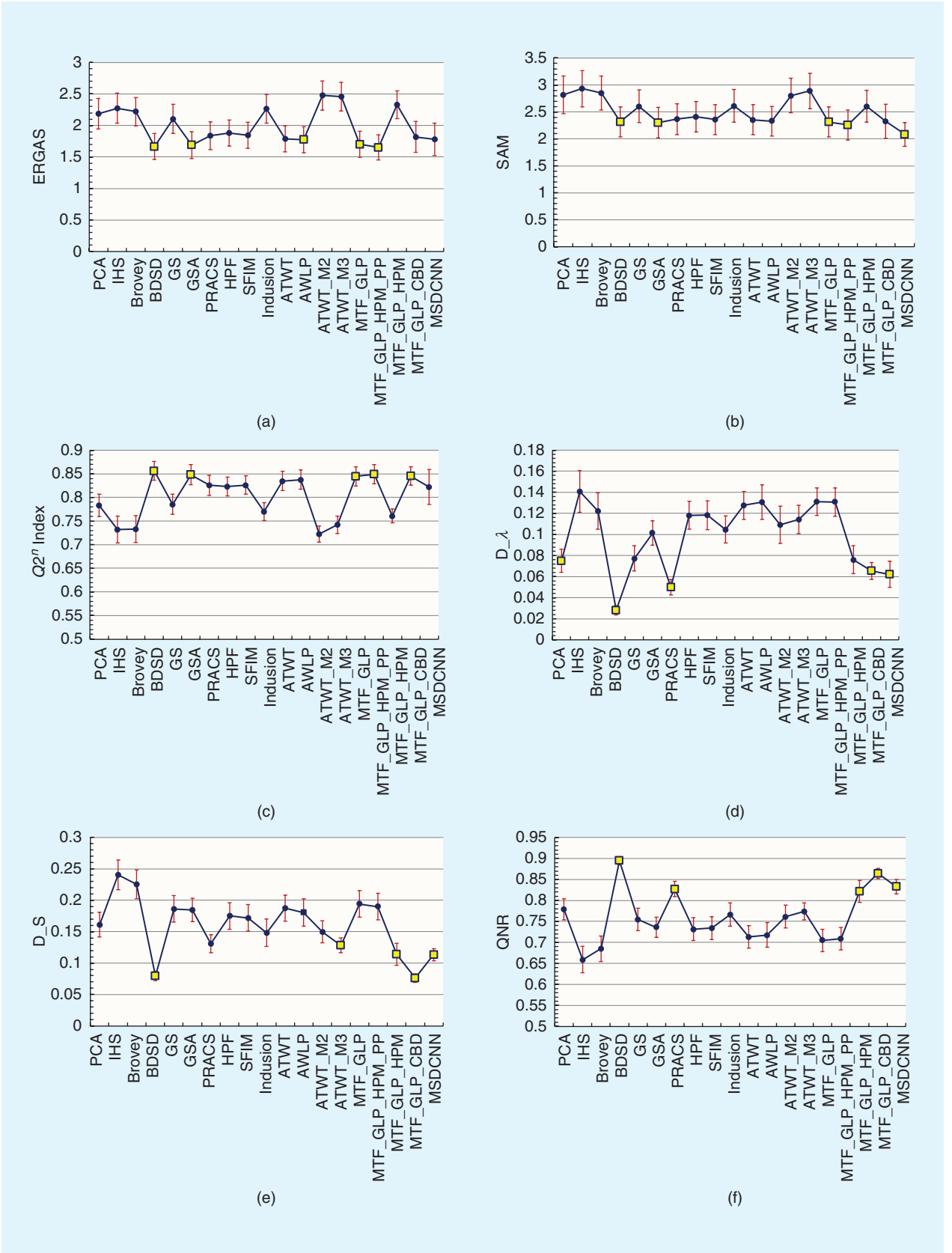

FIGURE 6. The experiment results based on the *Ikonos* data set. (a)–(c) The results with the ERGAS, SAM, and $Q2^n$ index, respectively, in the reduced-resolution experiment. (d)–(f) The results with D_λ, D_s, and the QNR, respectively, in the full-resolution experiment. SFIM: smoothing filter-based intensity modulation.

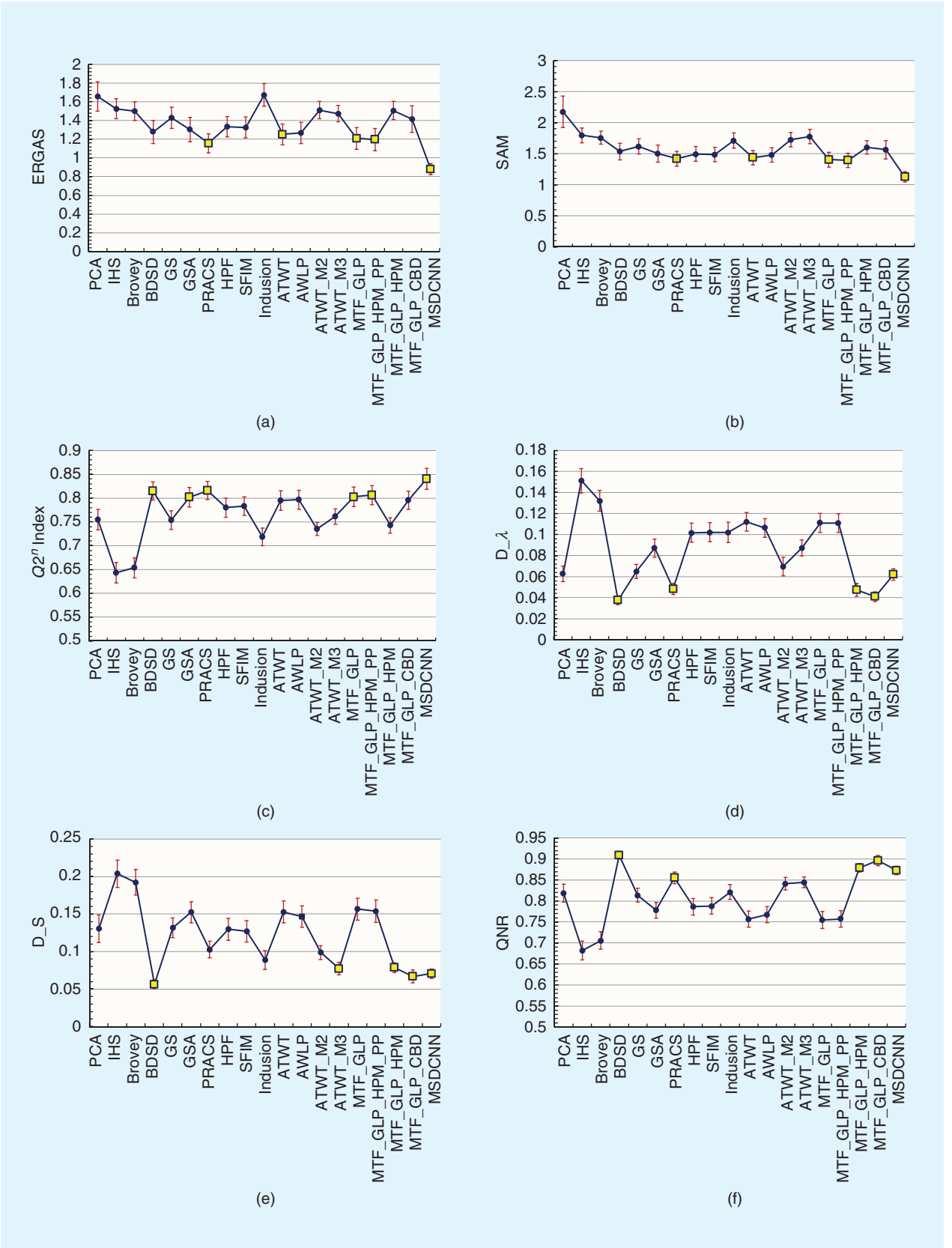

FIGURE 7. The experiment results based on the *QuickBird* data set. (a)–(c) The results with the ERGAS, SAM, and $Q2^n$ index, respectively, in the reduced-resolution experiment. (d)–(f) The results with D_λ, D_s, and the QNR, respectively, in the full resolution experiment.

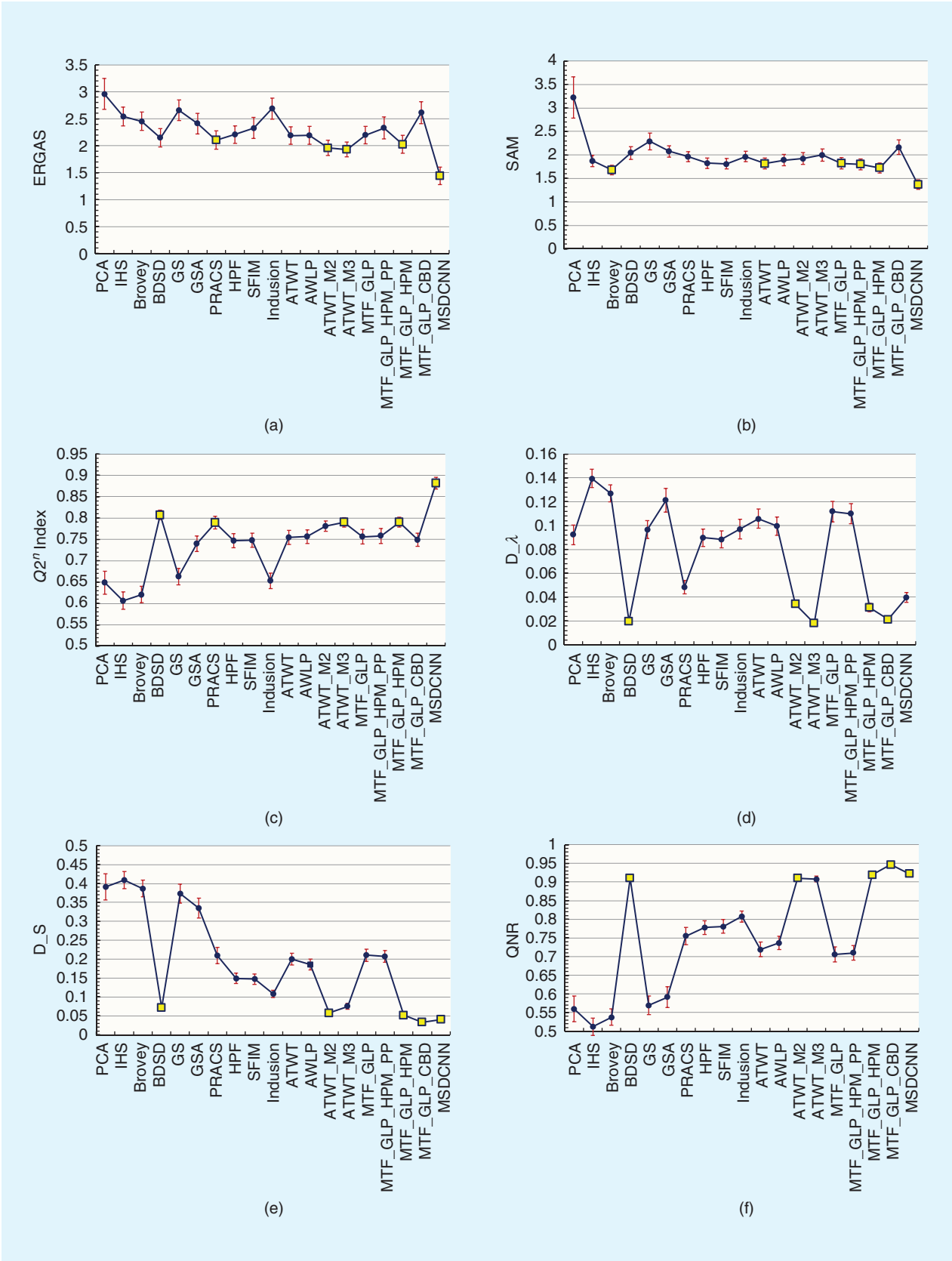

FIGURE 8. The experiment results based on the *GF-1* data set. (a)–(c) The results with the ERGAS, SAM, and $Q2^n$ index, respectively, in the reduced-resolution experiment. (d)–(f) The results with D_λ, D_s, and the QNR, respectively, in the full-resolution experiment.

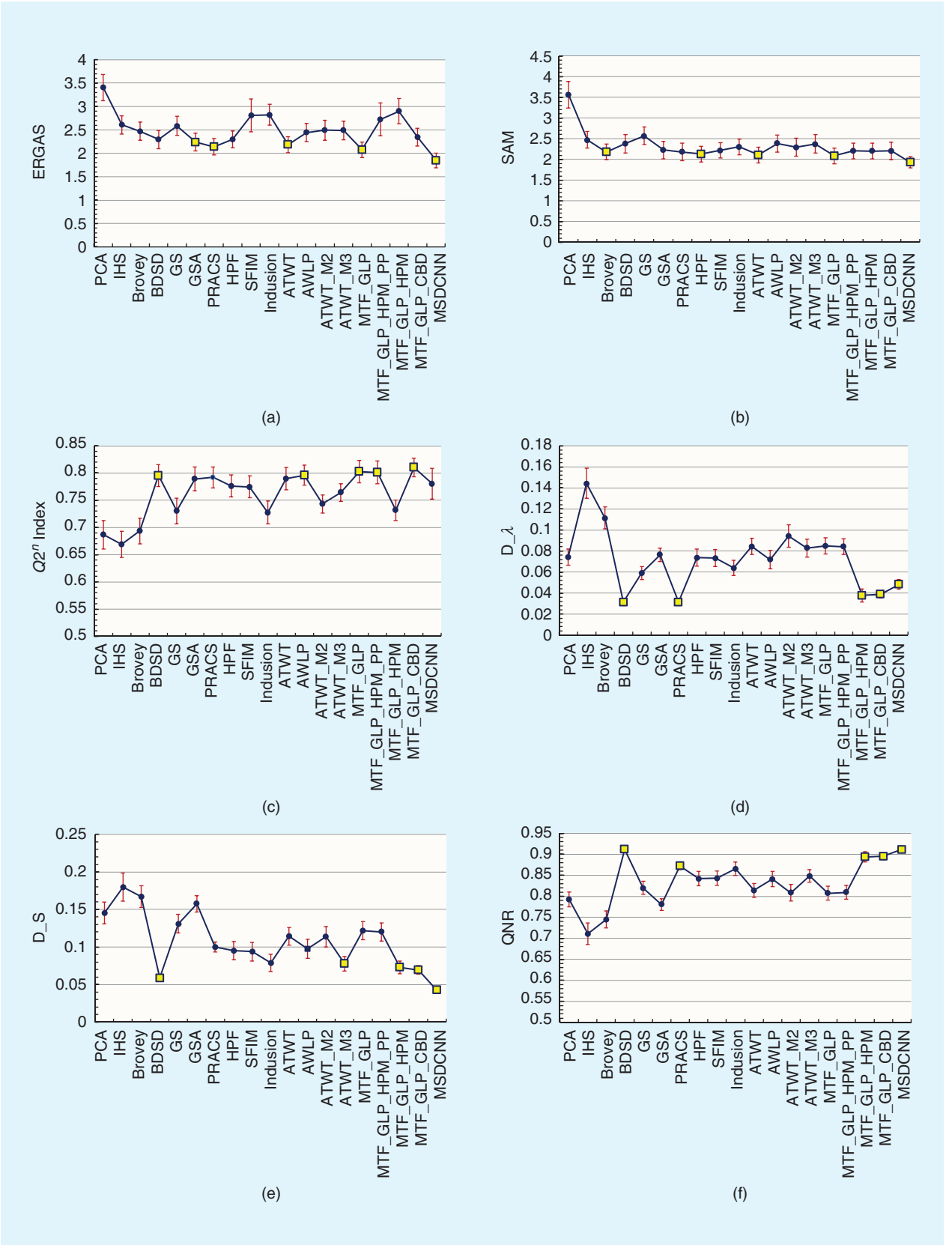

FIGURE 9. The experiment results based on the *WorldView-4* data set. (a)–(c) The results with the ERGAS, SAM, and $Q2^n$ index, respectively, in the reduced-resolution experiment. (d)–(f) The results with D_λ, D_s, and the QNR, respectively, in the full-resolution experiment.

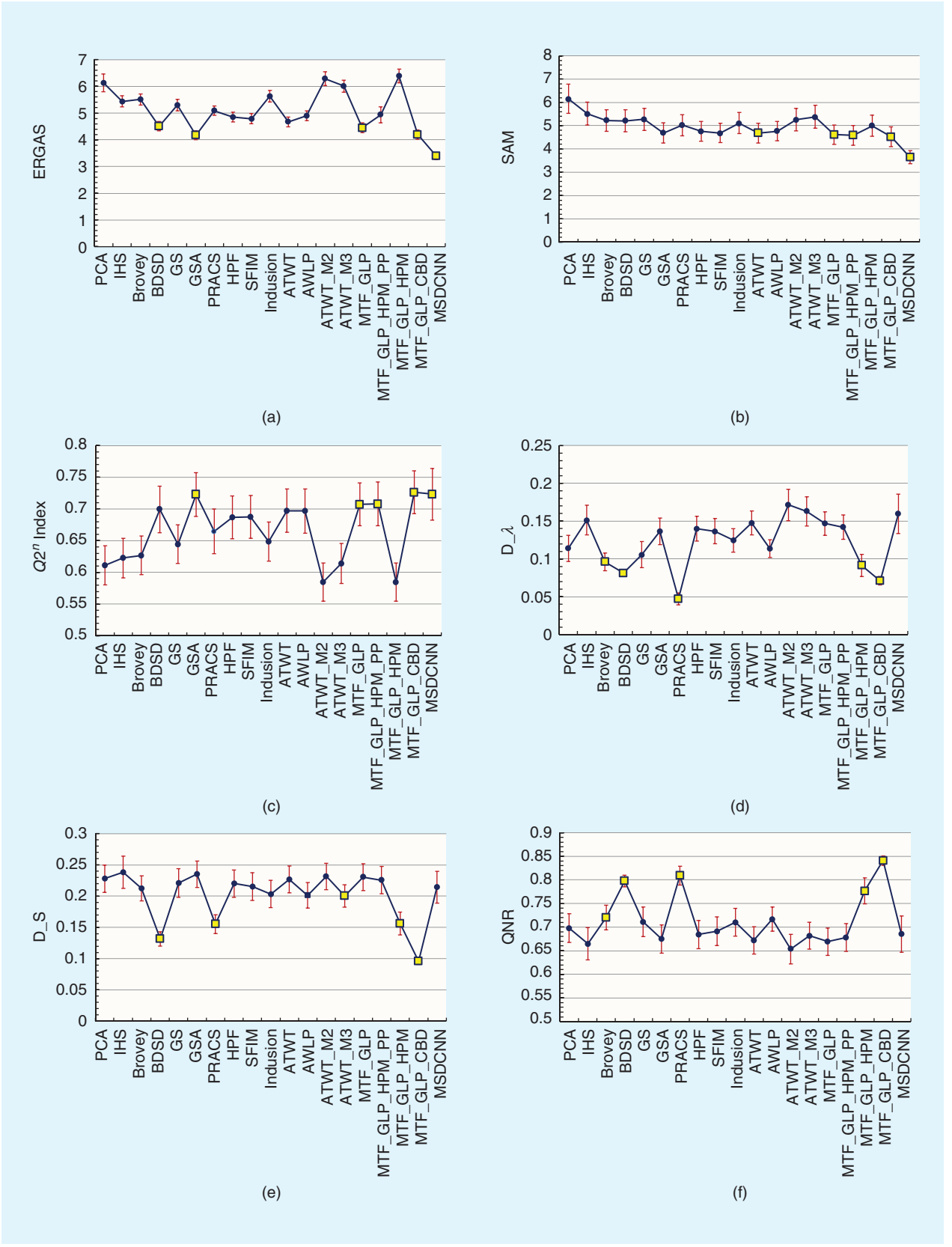

FIGURE 10. The experiment results based on the *WorldView-2* data set. (a)–(c) The results with the ERGAS, SAM, and $Q2^n$ index, respectively, in the reduced-resolution experiment. (d)–(f) The results with D_λ, D_s, and the QNR, respectively, in the full-resolution experiment.

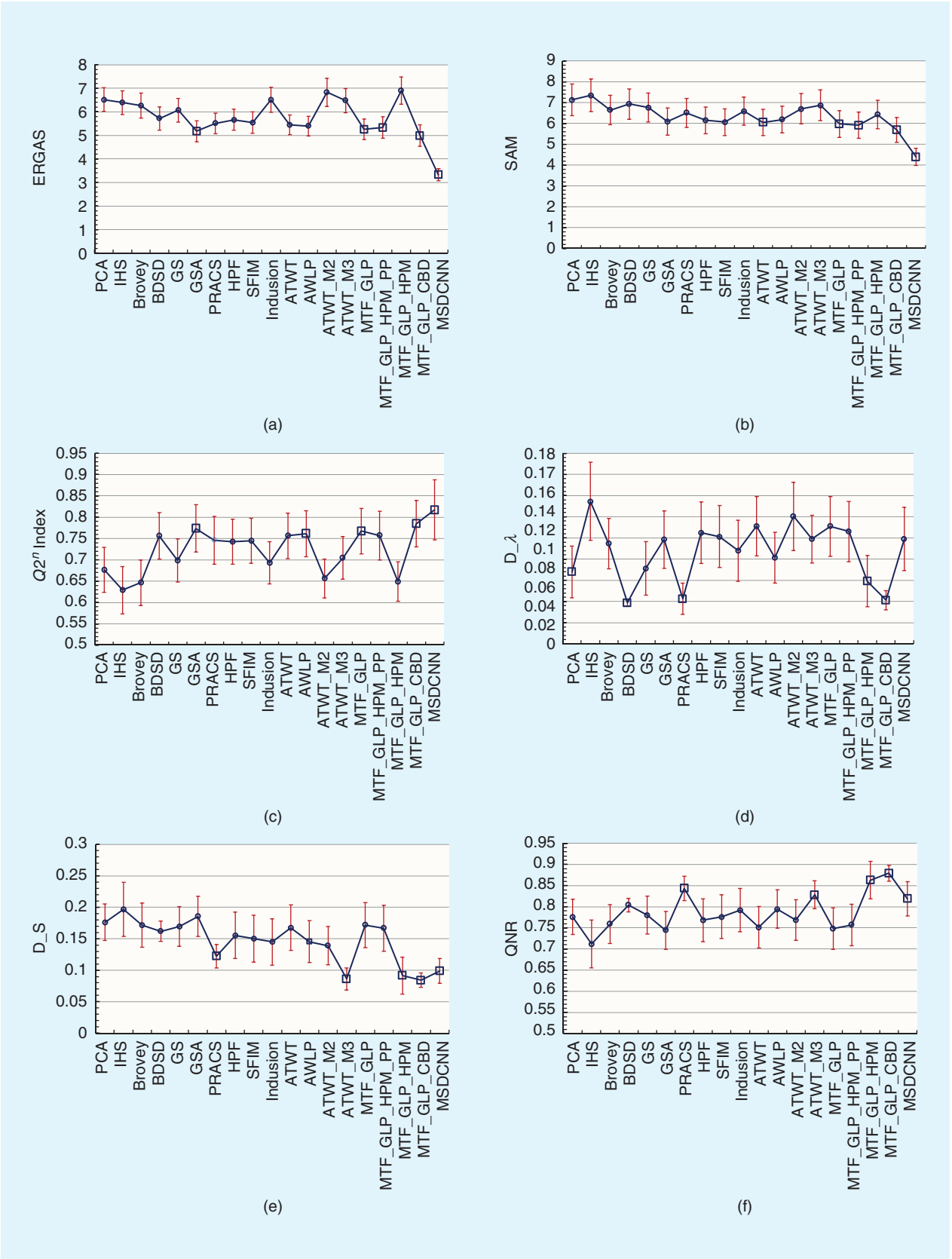

FIGURE 11. The experiment results based on the *WorldView-3* data set. (a)–(c) The results with the ERGAS, SAM, and $Q2^n$ index, respectively, in the reduced-resolution experiment. (d)–(f) The results with D_λ, D_s, and the QNR, respectively, in the full-resolution experiment.

TABLE 4. THE DETAILED STATISTICAL RESULTS, WITH A 95% CONFIDENCE INTERVAL, BASED ON THE *IKONOS* DATA SET.

	REDUCED RESOLUTION			FULL RESOLUTION		
METHOD	**ERGAS**	**SAM**	**$Q2^N$-INDEX**	**D_λ**	**D_S**	**QNR**
PCA	[1.942, 2.429]	[2.471, 3.169]	[0.76, 0.807]	**[0.064, 0.086]**	[0.141, 0.181]	[0.754, 0.804]
HIS	[2.036, 2.512]	[2.597, 3.267]	[0.704, 0.76]	[0.121, 0.161]	[0.216, 0.264]	[0.627, 0.691]
Brovey	[1.993, 2.445]	[2.535, 3.164]	[0.704, 0.761]	[0.105, 0.14]	[0.203, 0.248]	[0.655, 0.715]
BDSD	**[1.456, 1.874]**	**[2.036, 2.595]**	**[0.836, 0.877]**	**[0.024, 0.032]**	**[0.072, 0.087]**	**[0.885, 0.905]**
GS	[1.872, 2.333]	[2.303, 2.907]	[0.765, 0.807]	[0.065, 0.09]	[0.165, 0.208]	[0.728, 0.781]
GSA	**[1.475, 1.904]**	**[2.022, 2.581]**	**[0.827, 0.869]**	[0.09, 0.113]	[0.166, 0.203]	[0.712, 0.76]
PRACS	[1.616, 2.06]	[2.082, 2.649]	[0.804, 0.847]	**[0.043, 0.057]**	[0.117, 0.145]	**[0.809, 0.845]**
HPF	[1.671, 2.089]	[2.127, 2.695]	[0.803, 0.843]	[0.105, 0.131]	[0.154, 0.196]	[0.704, 0.759]
SFIM	[1.639, 2.051]	[2.079, 2.637]	[0.807, 0.846,]	[0.105, 0.132]	[0.151, 0.193]	[0.707, 0.762]
Indusion	[2.038, 2.492]	[2.314, 2.914]	[0.751, 0.789,]	[0.092, 0.117]	[0.127, 0.17]	[0.739, 0.795]
ATWT	[1.58, 1.993]	[2.075, 2.635]	[0.814, 0.855]	[0.114, 0.141]	[0.166, 0.208]	[0.686, 0.74]
AWLP	**[1.568, 1.978]**	[2.056, 2.6]	[0.817, 0.858]	[0.114, 0.147]	[0.159, 0.202]	[0.689, 0.747]
ATWT-M2	[2.246, 2.707]	[2.483, 3.126]	[0.705, 0.74]	[0.091, 0.127]	[0.132, 0.167]	[0.734, 0.789]
ATWT-M3	[2.231, 2.684]	[2.564, 3.219]	[0.724, 0.761]	[0.101, 0.128]	**[0.116, 0.14]**	[0.754, 0.794]
MTF-GLP	**[1.495, 1.907]**	**[2.033, 2.593]**	**[0.824, 0.865]**	[0.118, 0.144]	[0.173, 0.215]	[0.678, 0.731]
MTF-GLP-HPM-PP	**[1.449, 1.853]**	**[1.982, 2.533]**	**[0.829, 0.869]**	[0.117, 0.144]	[0.169, 0.211]	[0.682, 0.735]
MTF-GLP-HPM	[2.107, 2.549]	[2.308, 2.902]	[0.747, 0.775]	[0.063, 0.089]	**[0.096, 0.132]**	**[0.795, 0.848]**
MTF-GLP-CBD	[1.572, 2.062]	[2.009, 2.644]	**[0.826, 0.865]**	**[0.058, 0.073]**	**[0.069, 0.083]**	**[0.852, 0.876]**
MSDCNN	[1.524, 2.035]	**[1.861, 2.305]**	[0.785, 0.859]	**[0.05, 0.075]**	**[0.103, 0.123]**	**[0.815, 0.85]**

For Tables 4–13, the top five results for each quantitative evaluation index are marked in bold.

TABLE 5. THE DETAILED STATISTICAL RESULTS, WITH A 95% CONFIDENCE INTERVAL, BASED ON THE *QuickBird* DATA SET.

	REDUCED RESOLUTION			FULL RESOLUTION		
METHOD	**ERGAS**	**SAM**	**$Q2^N$-INDEX**	**D_λ**	**D_S**	**QNR**
PCA	[1.498, 1.812]	[1.92, 2.428]	[0.733, 0.777]	[0.055, 0.07]	[0.112, 0.149]	[0.798, 0.84]
HIS	[1.42, 1.632]	[1.676, 1.915]	[0.622, 0.665]	[0.14, 0.163]	[0.185, 0.222]	[0.66, 0.704]
Brovey	[1.4, 1.599]	[1.652, 1.866]	[0.633, 0.675]	[0.122, 0.142]	[0.175, 0.209]	[0.685, 0.726]
BDSD	**[1.154, 1.4]**	[1.401, 1.672]	**[0.795, 0.834]**	**[0.033, 0.042]**	**[0.05, 0.063]**	**[0.9, 0.918]**
GS	[1.315, 1.54]	[1.496, 1.743]	[0.734, 0.773]	[0.058, 0.072]	[0.118, 0.145]	[0.797, 0.830]
GSA	[1.172, 1.433]	[1.362, 1.636]	**[0.782, 0.823]**	[0.079, 0.096]	[0.138, 0.166]	[0.759, 0.796]
PRACS	**[1.055, 1.259]**	**[1.301, 1.539]**	**[0.797, 0.835]**	**[0.043, 0.054]**	[0.092, 0.114]	**[0.842, 0.869]**
HPF	[1.224, 1.445]	[1.381, 1.614]	[0.76, 0.8]	[0.093, 0.111]	[0.115, 0.144]	[0.767, 0.806]
SFIM	[1.215, 1.439]	[1.371, 1.603]	[0.763, 0.803]	[0.093, 0.111]	[0.113, 0.141]	[0.769, 0.808]
Indusion	[1.553, 1.794]	[1.59, 1.838]	[0.7, 0.737]	[0.093, 0.112]	[0.076, 0.101]	[0.804, 0.839]
ATWT	[1.14, 1.363]	**[1.32, 1.554]**	[0.774, 0.816]	[0.103, 0.121]	[0.138, 0.167]	[0.737, 0.776]
AWLP	[1.153, 1.381]	[1.362, 1.594]	[0.776, 0.817]	[0.098, 0.115]	[0.132, 0.161]	[0.748, 0.786]
ATWT-M2	[1.418, 1.603]	[1.612, 1.843]	[0.721, 0.749]	[0.061, 0.079]	[0.089, 0.108]	[0.826, 0.856]
ATWT-M3	[1.384, 1.562]	[1.659, 1.893]	[0.745, 0.778]	[0.08, 0.095]	**[0.069, 0.086]**	[0.831, 0.857]
MTF-GLP	**[1.093, 1.326]**	**[1.286, 1.523]**	**[0.782, 0.823]**	[0.102, 0.12]	[0.142, 0.171]	[0.735, 0.775]
MTF-GLP-HPM-PP	**[1.078, 1.316]**	**[1.276, 1.512]**	**[0.786, 0.826]**	[0.102, 0.12]	[0.139, 0.168]	[0.737, 0.777]
MTF-GLP-HPM	[1.409, 1.603]	[1.495, 1.714]	[0.727, 0.759]	**[0.041, 0.053]**	**[0.072, 0.085]**	**[0.868, 0.889]**
MTF-GLP-CBD	[1.273, 1.559]	[1.412, 1.715]	[0.776, 0.815]	**[0.036, 0.046]**	**[0.058, 0.076]**	**[0.884, 0.908]**
MSDCNN	**[0.822, 0.937]**	**[1.048, 1.209]**	**[0.819, 0.862]**	**[0.056, 0.068]**	**[0.064, 0.077]**	**[0.862, 0.883]**

TABLE 6. THE DETAILED STATISTICAL RESULTS, WITH A 95% CONFIDENCE INTERVAL, BASED ON THE *GF-1* DATA SET.

	REDUCED RESOLUTION			FULL RESOLUTION		
METHOD	**ERGAS**	**SAM**	**$Q2^N$-INDEX**	**D_λ**	**D_S**	**QNR**
PCA	[2.677, 3.244]	[2.785, 3.666]	[0.621, 0.676]	[0.084, 0.1	[0.356, 0.425]	[0.525, 0.594]
HIS	[2.369, 2.719]	[1.752, 1.994]	[0.586, 0.627]	[0.132, 0.147]	[0.386, 0.432]	[0.489, 0.535]
Brovey	[2.281, 2.627]	**[1.576, 1.781]**	[0.602, 0.64]	[0.12, 0.134]	[0.365, 0.408]	[0.516, 0.559]
BDSD	[1.98, 2.314]	[1.905, 2.179]	**[0.796, 0.818]**	**[0.018, 0.022]**	**[0.063, 0.08]**	**[0.901, 0.919]**
GS	[2.471, 2.85]	[2.107, 2.47]	[0.644, 0.682]	[0.089, 0.104]	[0.348, 0.398]	[0.545, 0.595]
GSA	[2.217, 2.6]	[1.957, 2.193]	[0.722, 0.758]	[0.111, 0.131]	[0.308, 0.361]	[0.563, 0.619]
PRACS	**[1.934, 2.272]**	[1.853, 2.07]	**[0.774, 0.804]**	[0.043, 0.054]	[0.188, 0.232]	[0.732, 0.778]
HPF	[2.042, 2.37]	[1.708, 1.938]	[0.731, 0.763]	[0.082, 0.097]	[0.136, 0.163]	[0.76, 0.796]
SFIM	[2.131, 2.523]	[1.698, 1.925]	[0.732, 0.765]	[0.081, 0.096]	[0.134, 0.161]	[0.763, 0.799]
Indusion	[2.494, 2.884]	[1.86, 2.075]	[0.635, 0.671]	[0.089, 0.105]	[0.098, 0.117]	[0.793, 0.823]
ATWT	[2.025, 2.347]	**[1.699, 1.936]**	[0.739, 0.772]	[0.098, 0.114]	[0.185, 0.216]	[0.7, 0.739]
AWLP	[2.030, 2.355]	[1.767, 2.011]	[0.741, 0.772]	[0.092, 0.107]	[0.171, 0.2]	[0.718, 0.755]
ATWT-M2	**[1.818, 2.099]**	[1.8, 2.054]	[0.769, 0.793]	**[0.032, 0.037]**	**[0.053, 0.063]**	**[0.905, 0.916]**
ATWT-M3	**[1.79, 2.071]**	[1.867, 2.129]	**[0.779, 0.801]**	**[0.016, 0.02]**	[0.068, 0.083]	[0.9, 0.916]
MTF-GLP	[2.034, 2.359]	**[1.701, 1.941]**	[0.739, 0.774]	[0.103, 0.12]	[0.195, 0.226]	[0.686, 0.726]
MTF-GLP-HPM-PP	[2.124, 2.532]	**[1.685, 1.92]**	[0.741, 0.776]	[0.102, 0.118]	[0.192, 0.223]	[0.69, 0.729]
MTF-GLP-HPM	**[1.86, 2.191]**	**[1.619, 1.832]**	**[0.779, 0.802]**	**[0.028, 0.035]**	**[0.047, 0.057]**	**[0.911, 0.926]**
MTF-GLP-CBD	[2.41, 2.815]	[2.01, 2.317]	[0.734, 0.765]	**[0.019, 0.024]**	**[0.03, 0.037]**	**[0.941, 0.952]**
MSDCNN	**[1.276, 1.601]**	**[1.266, 1.47]**	**[0.868, 0.896]**	[0.036, 0.044]	**[0.036, 0.045]**	**[0.915, 0.929]**

TABLE 7. THE DETAILED STATISTICAL RESULTS, WITH A 95% CONFIDENCE INTERVAL, BASED ON THE *WorldView-4* DATA SET.

	REDUCED RESOLUTION			FULL RESOLUTION		
METHOD	**ERGAS**	**SAM**	**$Q2^N$-INDEX**	**D_λ**	**D_S**	**QNR**
PCA	[3.126, 3.685]	[3.246, 3.887]	[0.66, 0.713]	[0.067, 0.082]	[0.131, 0.16]	[0.775, 0.81]
HIS	[2.408, 2.804]	[2.269, 2.678]	[0.645, 0.693]	[0.13, 0.159]	[0.161, 0.198]	[0.685, 0.736]
Brovey	[2.278, 2.665]	**[1.99, 2.368]**	[0.67, 0.717]	[0.101, 0.122]	[0.153, 0.182]	[0.725, 0.765]
BDSD	[2.102, 2.489]	[2.157, 2.605]	**[0.775, 0.816]**	**[0.028, 0.034]**	**[0.055, 0.062]**	**[0.907, 0.916]**
GS	[2.381, 2.793]	[2.357, 2.782]	[0.707, 0.754]	[0.053, 0.065]	[0.119, 0.143]	[0.805, 0.835]
GSA	**[2.048, 2.426]**	[2.016, 2.438]	[0.767, 0.811]	[0.07, 0.083]	[0.147, 0.168]	[0.766, 0.794]
PRACS	**[1.961, 2.315]**	[1.966, 2.395]	[0.773, 0.811]	**[0.028, 0.035]**	[0.094, 0.107]	**[0.864, 0.881]**
HPF	[2.113, 2.476]	**[1.94, 2.317]**	[0.756, 0.796]	[0.066, 0.082]	[0.083, 0.107]	[0.825, 0.859]
SFIM	[2.461, 3.159]	[2.034, 2.408]	[0.754, 0.795]	[0.065, 0.081]	[0.082, 0.106]	[0.826, 0.861]
Indusion	[2.601, 3.047]	[2.116, 2.494]	[0.707, 0.748]	[0.056, 0.071]	[0.067, 0.09]	[0.849, 0.881]
ATWT	**[2.014, 2.355]**	**[1.918, 2.295]**	[0.769, 0.81]	[0.077, 0.092]	[0.102, 0.126]	[0.798, 0.831]
AWLP	[2.254, 2.634]	[2.18, 2.588]	**[0.777, 0.814]**	[0.063, 0.081]	[0.085, 0.11]	[0.823, 0.860]
ATWT-M2	[2.282, 2.707]	[2.074, 2.515]	[0.726, 0.76]	[0.083, 0.105]	[0.1, 0.127]	[0.788, 0.828]
ATWT-M3	[2.289, 2.69]	[2.152, 2.597]	[0.748, 0.781]	[0.074, 0.091]	**[0.068, 0.087]**	[0.833, 0.864]
MTF-GLP	**[1.908, 2.241]**	**[1.894, 2.275]**	**[0.782, 0.823]**	[0.077, 0.092]	[0.11, 0.134]	[0.791, 0.824]
MTF-GLP-HPM-PP	[2.37, 3.078]	[2.014, 2.395]	**[0.78, 0.822]**	[0.077, 0.092]	[0.108, 0.132]	[0.793, 0.826]
MTF-GLP-HPM	[2.63, 3.167]	[2.01, 2.39]	[0.713, 0.75]	**[0.031, 0.044]**	**[0.065, 0.081]**	**[0.881, 0.907]**
MTF-GLP-CBD	[2.154, 2.538]	[1.996, 2.415]	**[0.793, 0.827]**	**[0.036, 0.042]**	**[0.063, 0.075]**	**[0.887, 0.904]**
MSDCNN	**[1.694, 2.002]**	**[1.791, 2.068]**	[0.752, 0.808]	**[0.044, 0.053]**	**[0.039, 0.047]**	**[0.905, 0.918]**

TABLE 8. THE DETAILED STATISTICAL RESULTS, WITH A 95% CONFIDENCE INTERVAL, BASED ON THE *WorldView-2* DATA SET.

	REDUCED RESOLUTION			FULL RESOLUTION		
METHOD	**ERGAS**	**SAM**	**$Q2^N$-INDEX**	**D_λ**	**D_S**	**QNR**
PCA	[5.791, 6.461]	[5.523, 6.79]	[0.581, 0.642]	[0.097, 0.131]	[0.206, 0.25]	[0.667, 0.729]
HIS	[5.232, 5.642]	[5.027, 6.012]	[0.592, 0.653]	[0.132, 0.171]	[0.213, 0.264]	[0.63, 0.699]
Brovey	[5.306, 5.721]	[4.757, 5.692]	[0.596, 0.657]	**[0.084, 0.108]**	[0.193, 0.233]	**[0.694, 0.746]**
BDSD	**[4.342, 4.678]**	[4.73, 5.68]	[0.662, 0.736]	**[0.077, 0.086]**	**[0.12, 0.143]**	**[0.785, 0.81]**
GS	[5.088, 5.521]	[4.797, 5.742]	[0.614, 0.675]	[0.089, 0.123]	[0.198, 0.244]	[0.68, 0.743]
GSA	**[4.011, 4.329]**	[4.247, 5.121]	**[0.688, 0.757]**	[0.119, 0.154]	[0.214, 0.256]	[0.645, 0.705]
PRACS	[4.911, 5.27]	[4.562, 5.479]	[0.629, 0.7]	**[0.039, 0.055]**	**[0.14, 0.17]**	**[0.789, 0.829]**
HPF	[4.665, 5.038]	[4.326, 5.176]	[0.653, 0.72]	[0.123, 0.157]	[0.198, 0.242]	[0.654, 0.714]
SFIM	[4.6, 4.98]	[4.266, 5.113]	[0.653, 0.721]	[0.12, 0.153]	[0.193, 0.238]	[0.661, 0.722]
Indusion	[5.412, 5.844]	[4.656, 5.562]	[0.618, 0.680]	[0.109, 0.14]	[0.181, 0.225]	[0.68, 0.74]
ATWT	[4.493, 4.851]	**[4.262, 5.099]**	[0.663, 0.731]	[0.132, 0.163]	[0.205, 0.248]	[0.643, 0.701]
AWLP	[4.712, 5.086]	[4.342, 5.19]	[0.662, 0.731]	[0.102, 0.125]	[0.181, 0.221]	[0.691, 0.743]
ATWT-M2	[6.03, 6.541]	[4.781, 5.743]	[0.554, 0.615]	[0.15, 0.192]	[0.21, 0.252]	[0.622, 0.685]
ATWT-M3	[5.789, 6.224]	[4.89, 5.874]	[0.582, 0.646]	[0.144, 0.182]	**[0.182, 0.218]**	[0.654, 0.711]
MTF-GLP	**[4.267, 4.634]**	**[4.194, 5.033]**	**[0.673, 0.741]**	[0.131, 0.162]	[0.209, 0.252]	[0.64, 0.698]
MTF-GLP-HPM-PP	[4.64, 5.234]	**[4.167, 5.006]**	**[0.674, 0.742]**	[0.126, 0.158]	[0.204, 0.248]	[0.648, 0.707]
MTF-GLP-HPM	[6.134, 6.641]	[4.547, 5.445]	[0.554, 0.615]	**[0.077, 0.106]**	**[0.138, 0.175]**	**[0.749, 0.804]**
MTF-GLP-CBD	**[4.021, 4.372]**	**[4.096, 4.952]**	**[0.692, 0.76]**	**[0.065, 0.077]**	**[0.091, 0.101]**	**[0.831, 0.85]**
MSDCNN	**[3.26, 3.519]**	**[3.369, 3.919]**	**[0.682, 0.764]**	[0.134, 0.186]	[0.189, 0.24]	[0.647, 0.724]

TABLE 9. THE DETAILED STATISTICAL RESULTS, WITH A 95% CONFIDENCE INTERVAL, BASED ON THE *WorldView-3* DATA SET.

	REDUCED RESOLUTION			FULL RESOLUTION		
METHOD	**ERGAS**	**SAM**	**$Q2^N$-INDEX**	**D_λ**	**D_S**	**QNR**
PCA	[6.01, 7.033]	[6.377, 7.9]	[0.624, 0.729]	**[0.043, 0.093]**	[0.147, 0.205]	[0.734, 0.818]
HIS	[5.893, 6.899]	[6.57, 8.13]	[0.574, 0.685]	[0.098, 0.171]	[0.154, 0.24]	[0.656, 0.769]
Brovey	[5.738, 6.791]	[5.939, 7.347]	[0.593, 0.7]	[0.071, 0.118]	[0.136, 0.207]	[0.713, 0.805]
BDSD	[5.215, 6.215]	[6.199, 7.666]	[0.701, 0.811]	**[0.036, 0.041]**	[0.146, 0.178]	[0.788, 0.82]
GS	[5.557, 6.575]	[6.069, 7.468]	[0.648, 0.749]	[0.046, 0.097]	[0.138, 0.201]	[0.736, 0.825]
GSA	**[4.731, 5.617]**	[5.433, 6.746]	**[0.719, 0.829]**	[0.071, 0.126]	[0.154, 0.217]	[0.699, 0.789]
PRACS	[5.075, 5.949]	[5.806, 7.21]	[0.69, 0.802]	**[0.028, 0.057]**	**[0.104, 0.141]**	**[0.815, 0.872]**
HPF	[5.213, 6.109]	[5.503, 6.789]	[0.69, 0.795]	[0.076, 0.134]	[0.119, 0.192]	[0.718, 0.819]
SFIM	[5.093, 5.998]	[5.419, 6.698]	[0.692, 0.798]	[0.072, 0.131]	[0.113, 0.188]	[0.725, 0.828]
Indusion	[5.985, 7.044]	[5.918, 7.272]	[0.644, 0.743]	[0.059, 0.117]	[0.108, 0.182]	[0.741, 0.844]
ATWT	[5.026, 5.87]	**[5.419, 6.684]**	[0.703, 0.81]	[0.083, 0.139]	[0.132, 0.204]	[0.703, 0.801]
AWLP	[4.969, 5.817]	[5.544, 6.83]	**[0.708, 0.815]**	[0.057, 0.105]	[0.112, 0.179]	[0.749, 0.84]
ATWT-M2	[6.233, 7.432]	[5.985, 7.43]	[0.611, 0.702]	[0.088, 0.153]	[0.109, 0.169]	[0.721, 0.816]
ATWT-M3	[5.955, 6.994]	[6.135, 7.619]	[0.654, 0.755]	[0.076, 0.121]	**[0.068, 0.104]**	**[0.795, 0.861]**
MTF-GLP	**[4.823, 5.694]**	**[5.34, 6.604]**	**[0.714, 0.821]**	[0.083, 0.139]	[0.136, 0.208]	[0.699, 0.797]
MTF-GLP-HPM-PP	**[4.873, 5.789]**	**[5.284, 6.545]**	[0.7, 0.814]	[0.078, 0.135]	[0.13, 0.203]	[0.707, 0.806]
MTF-GLP-HPM	[6.319, 7.488]	[5.748, 7.105]	[0.603, 0.695]	**[0.035, 0.084]**	**[0.062, 0.121]**	**[0.818, 0.907]**
MTF-GLP-CBD	**[4.528, 5.449]**	**[5.094, 6.286]**	**[0.731, 0.839]**	**[0.032, 0.05]**	**[0.072, 0.096]**	**[0.861, 0.898]**
MSDCNN	**[3.068, 3.587]**	**[3.974, 4.814]**	**[0.747, 0.888]**	[0.069, 0.129]	**[0.079, 0.119]**	**[0.778, 0.859]**

On the whole, the experiment results show that the pansharpening methods have slightly different performances on various remote sensing satellites, especially between the four- and eight-band images. Most of the pansharpening methods have a relatively poorer performance for the eight-band images than for the four-band ones. The BDSD and MSDCNN methods generally show robust and excellent performance in the reduced- and full-resolution experiments. The MTF-GLP and MTF-GLP-HPM -PP methods generally show satisfactory performance in terms of the quantitative-evaluation indices in the reduced-resolution experiment and poorer performance in the full-resolution experiment. Conversely, the MTF-GLP-HPM generally shows satisfactory performance in terms of D_λ, D_s, and the QNR in the full-resolution experiment and relatively poorer performance in the reduced-resolution experiment.

EXPERIMENTAL RESULTS BASED ON DATA SETS FOR DIFFERENT THEMATIC SURFACE FEATURES

The experiment results based on different thematic data sets are shown in Figures 12–15 and Tables 10–13, respectively. For each figure, the horizontal axis represents different pansharpening methods, and the vertical axis denotes the statistical results. In addition, the top five results for each evaluation index are marked by dark colors. In the experiments, to avoid the influence of the different spectral range coverages between the HR Pan and LR MS images, only thematic images from *Ikonos*, *QuickBird*, *GF-1*, and *WorldView-4* with four-band MS were employed. Therefore, 305 pairs of HR Pan/LR MS images in the urban data set, 203 pairs of HR Pan/LR MS images in the green vegetation data set, 153 pairs of HR Pan/LR MS images in the water-scenario data set, and 949 pairs of HR Pan/LR MS images in the unlabeled data set were utilized.

As shown in Figure 12 and Table 10, the BDSD and MSDCNN methods show excellent performance based on the urban data set in the reduced- and full-resolution experiments. The ATWT, MTF-GLP, and MTF-GLP-HPM-PP methods show satisfactory results in the reduced-resolution experiment and relatively poorer performance in the full-resolution experiment. Conversely, the MTF-GLP-HPM and MTF-GLP-CBD methods have satisfactory performance in the full-resolution experiment and poorer performance in the reduced-resolution experiment.

Figure 13 and Table 11 provide the experiment results based on the green vegetation data set. It is shown that the BDSD and MSDCNN methods have excellent performance in the reduced- and full-resolution experiments; the MSDCNN method has the best results. Similar to the results based on the urban data set, the ATWT and MTF-GLP methods have better results in the reduced-resolution experiment, and the MTF-GLP-HPM and MTF-GLP-CBD methods achieve better performance in the full-resolution experiment.

The experiment results based on the water-scenario data set are shown in Figure 14 and Table 12. The performance of almost all the pansharpening methods is poorer than for the urban and green vegetation data sets, which can be observed in the $Q2^n$ index. This is because the edges and detailed structures of the images in the water-scenario data set are generally missing, and the advantages of the HR Pan cannot be comprehensively utilized. In addition, the results show that the SFIM, MTF-GLP-HPM-PP, and MTF-GLP-HPM methods have poor robustness with a wider confidence interval in terms of the ERGAS.

For the unlabeled data set with mixed surface features, the BDSD and MSDCNN methods show excellent performance, and the MSDCNN has a competitive advantage over most of the other pansharpening techniques in the reduced- and full-resolution experiments. In addition, similar to the results based on the water-scenario data set, the SFIM, MTF-GLP-HPM-PP, and MTF-GLP-HPM methods display poorer robustness in terms of the ERGAS index.

On the whole, the experiment results based on the proposed large-scale thematic data set show that different pansharpening methods have a relatively larger difference for various thematic images. In addition, the SFIM, MTF-GLP-HPM-PP, and MTF-GLP-HPM methods with the same HPM high-frequency injection model generally show poorer robustness.

HETEROGENEITY ANALYSIS OF PANSHARPENING METHODS FOR THEMATIC DATA SET

The objective of this article is to more comprehensively study the characteristics of pansharpening methods for different thematic remote sensing images. The homogeneity and heterogeneity among pansharpening methods for the four thematic remote sensing data sets were analyzed and discussed. They were measured by the variance among the pansharpening methods in terms of different evaluation indices. The bigger the variance, the larger the heterogeneity of the pansharpening methods for a thematic remote sensing data set; the smaller the variance, the more the pansharpening methods are similar. It should be noted that, to more comprehensively study this characteristic, all the four-band thematic images were pansharpened and analyzed for each data set; this does not include the previous 30% data set volume, which was limited by the DL-based MSDCNN method. Therefore, the 18 pansharpening methods without the MSDCNN were used. The results for the reduced- and full-resolution data sets are shown in Figures 16 and 17, respectively.

The distribution of the variance among pansharpening methods for different thematic data sets based on the reduced-resolution images is shown in Figure 16. The horizontal axis denotes the sequence number of the images, and the vertical axis gives the variance values. Figure 16(a) shows the distribution of the variance for the four thematic data sets (i.e., the urban, green vegetation, water-scenario, and unlabeled data sets) in terms of the ERGAS index. Figure 16(b) and (c) provides the distribution of the variance in terms of the SAM and $Q2^n$ index,

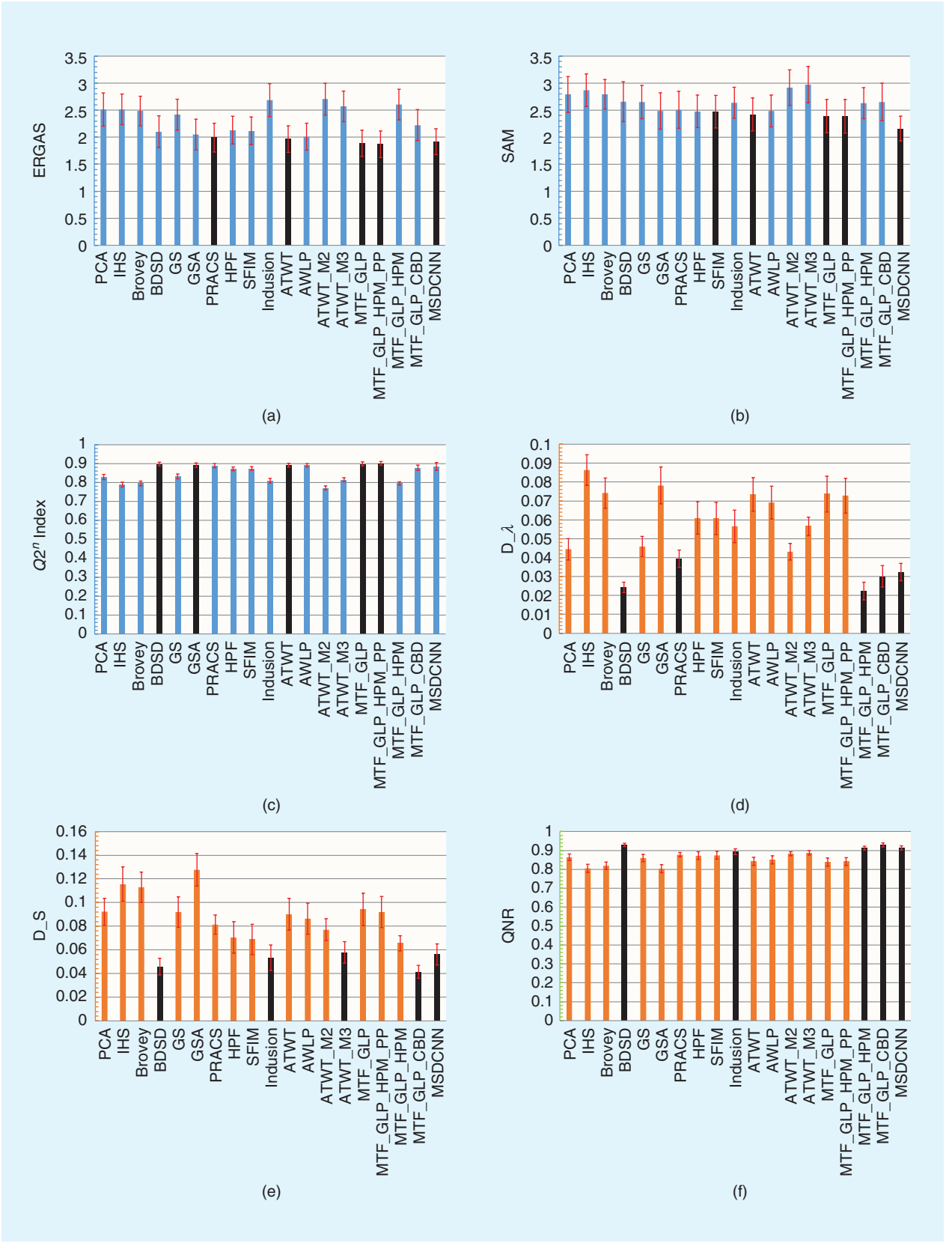

FIGURE 12. The experiment results based on the urban data set. (a)–(c) The results with the ERGAS, SAM, and $Q2^n$ index, respectively, in the reduced-resolution experiment. (d)–(f) The results with D_λ, D_s, and the QNR, respectively, in the full-resolution experiment.

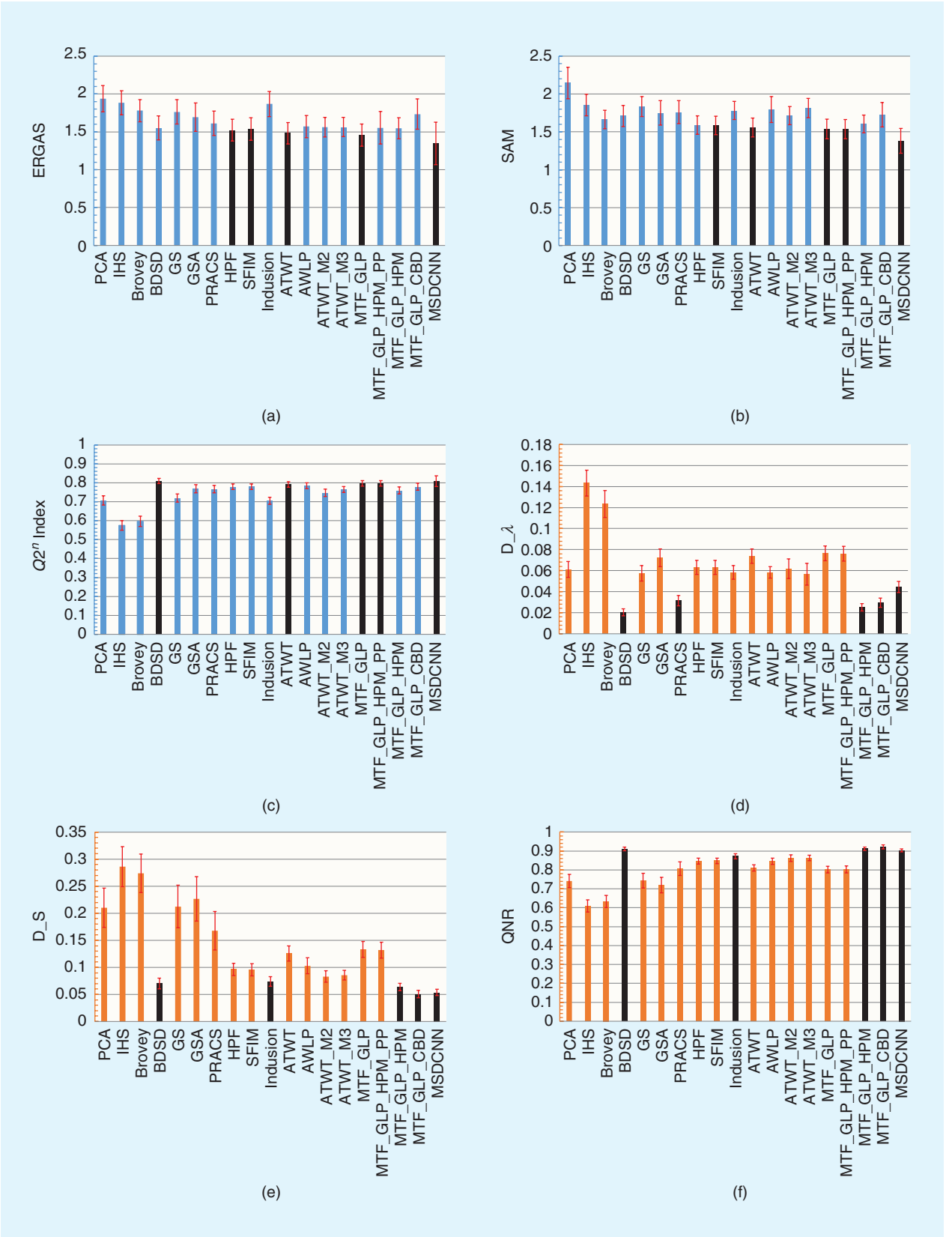

FIGURE 13. The experiment results based on the green vegetation data set. (a)–(c) The results with the ERGAS, SAM, and $Q2^n$ index, respectively, in the reduced-resolution experiment. (d)–(f) The results with D_λ, D_s, and the QNR, respectively, in the full-resolution experiment.

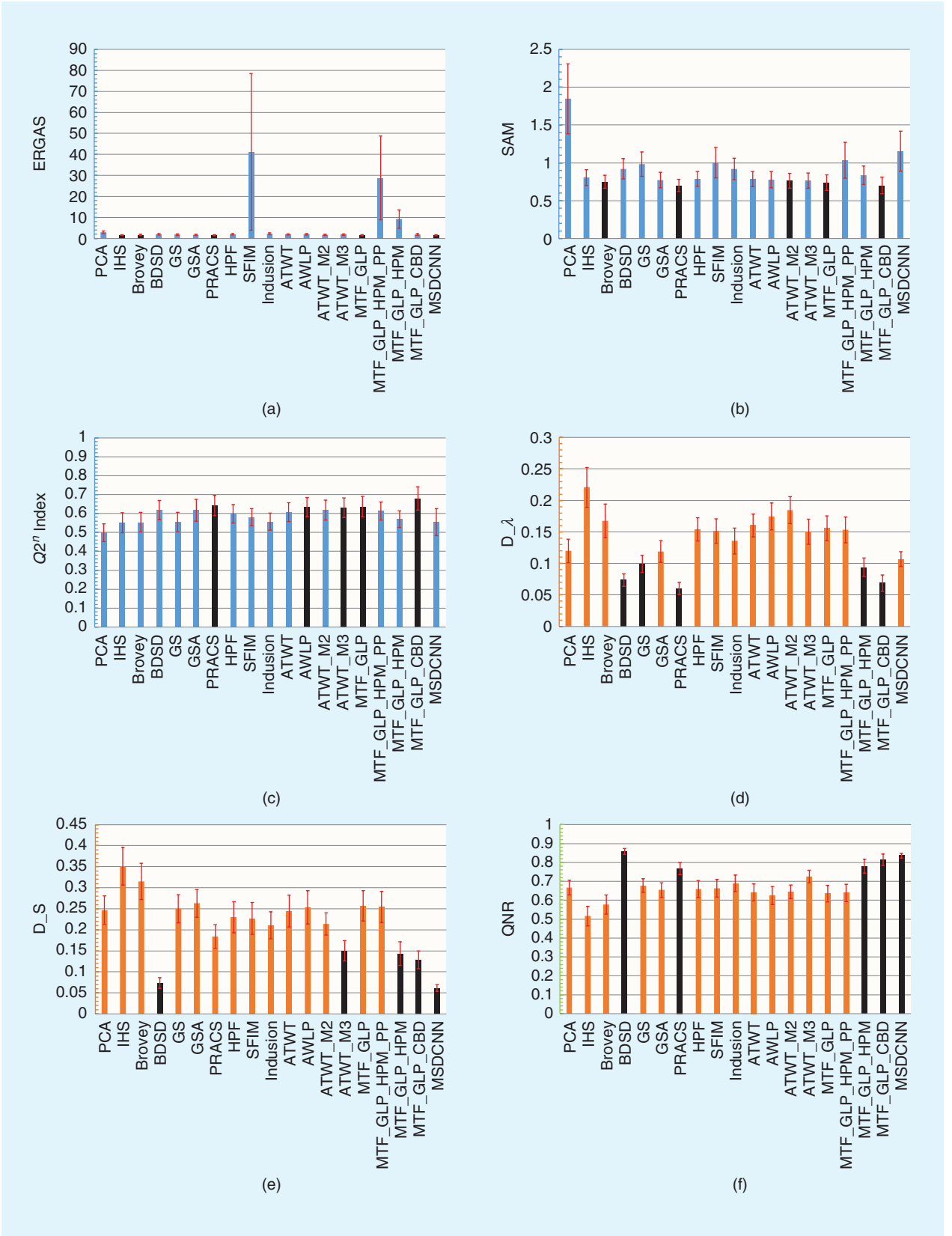

FIGURE 14. The experiment results based on the water-scenario data set. (a)–(c) The results with the ERGAS, SAM, and $Q2^n$ index, respectively, in the reduced-resolution experiment. (d)–(f) The results with D_λ, D_s, and the QNR, respectively, in the full-resolution experiment.

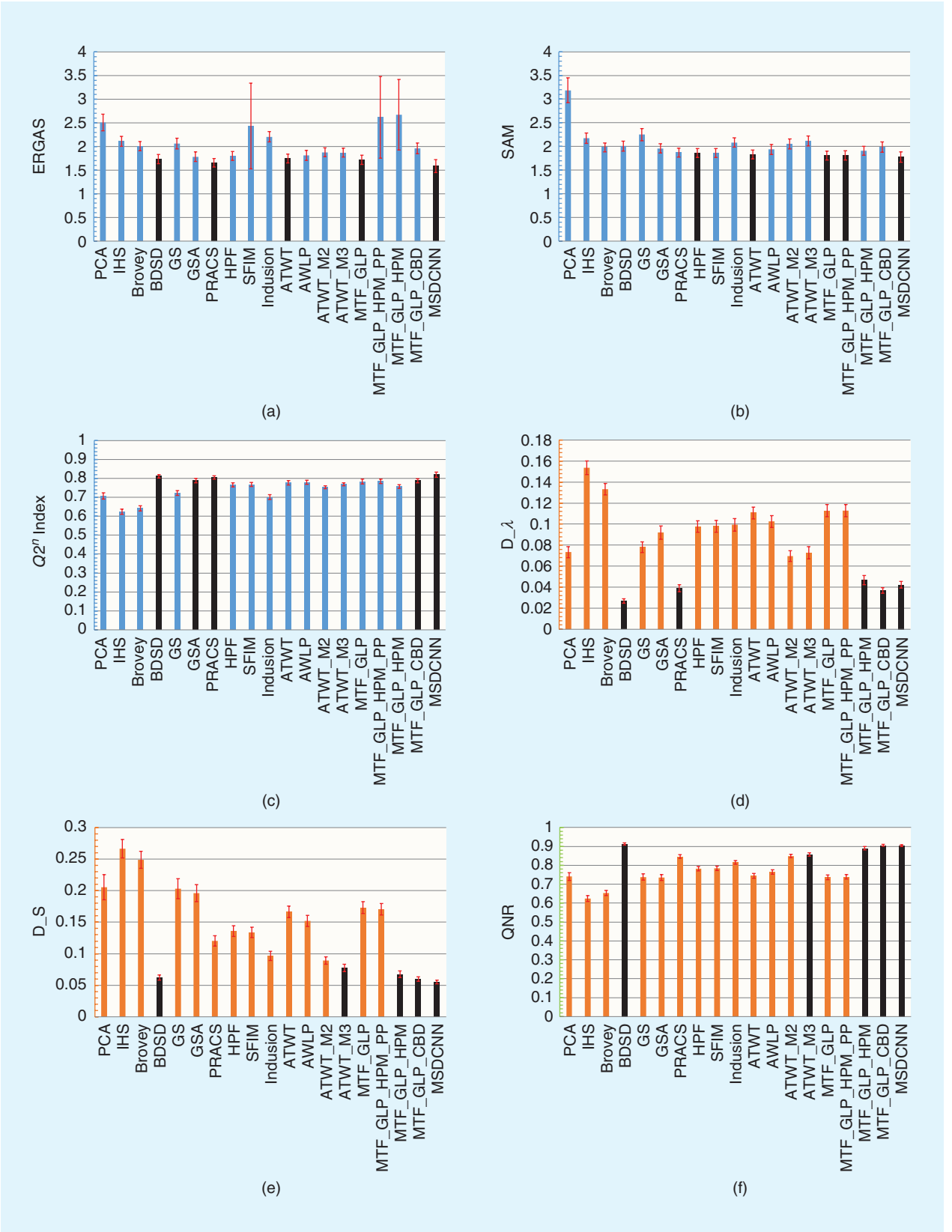

FIGURE 15. The experiment results based on the unlabeled data set with mixed surface features. (a)–(c) The results with the ERGAS, SAM, and $Q2^n$ index, respectively, in the reduced-resolution experiment. (d)–(f) The results with D_λ, D_s, and the QNR, respectively, in the full-resolution experiment.

TABLE 10. DETAILED STATISTICAL RESULTS, WITH A 95% CONFIDENCE INTERVAL, BASED ON THE URBAN DATA SET.

	REDUCED RESOLUTION			FULL RESOLUTION		
METHOD	**ERGAS**	**SAM**	**$Q2^N$-INDEX**	**D_λ**	**D_S**	**QNR**
PCA	[2.206, 2.821]	[2.460, 3.118]	[0.815, 0.842]	[0.039, 0.05]	[0.081, 0.104]	[0.846, 0.882]
HIS	[2.229, 2.798]	[2.564, 3.17]	[0.776, 0.802]	[0.078, 0.094]	[0.101, 0.13]	[0.785, 0.827]
Brovey	[2.212, 2.76]	[2.519, 3.067]	[0.782, 0.808]	[0.066, 0.082]	[0.1, 0.126]	[0.799, 0.838]
BDSD	[1.804, 2.395]	[2.288, 3.024]	**[0.887, 0.907]**	**[0.022, 0.027]**	**[0.039, 0.053]**	**[0.921, 0.938]**
GS	[2.131, 2.705]	[2.34, 2.96]	[0.819, 0.845]	[0.041, 0.051]	[0.079, 0.105]	[0.841, 0.879]
GSA	[1.766, 2.332]	[2.152, 2.823]	**[0.88, 0.903]**	[0.068, 0.088]	[0.114, 0.142]	[0.782, 0.826]
PRACS	**[1.725, 2.261]**	[2.166, 2.845]	[0.878, 0.898]	**[0.035, 0.044]**	[0.073, 0.09]	[0.867, 0.89]
HPF	[1.875, 2.39]	[2.179, 2.779]	[0.862, 0.882]	[0.052, 0.07]	[0.057, 0.084]	[0.852, 0.893]
SFIM	[1.862, 2.377]	**[2.17, 2.771]**	[0.864, 0.883]	[0.052, 0.069]	[0.056, 0.082]	[0.855, 0.895]
Indusion	[2.38, 2.992]	[2.348, 2.922]	[0.799, 0.821]	[0.048, 0.065]	**[0.043, 0.064]**	**[0.879, 0.91]**
ATWT	**[1.721, 2.21]**	**[2.113, 2.721]**	**[0.881, 0.901]**	[0.065, 0.082]	[0.077, 0.104]	[0.822, 0.864]
AWLP	[1.759, 2.261]	[2.193, 2.781]	[0.881, 0.9]	[0.06, 0.078]	[0.073, 0.099]	[0.83, 0.871]
ATWT-M2	[2.411, 2.999]	[2.587, 3.246]	[0.76, 0.782]	[0.039, 0.048]	[0.068, 0.086]	[0.872, 0.894]
ATWT-M3	[2.285, 2.852]	[2.632, 3.303]	[0.803, 0.825]	[0.052, 0.061]	**[0.049, 0.067]**	[0.876, 0.899]
MTF-GLP	**[1.645, 2.131]**	**[2.079, 2.694]**	**[0.888, 0.908]**	[0.064, 0.083]	[0.08, 0.108]	[0.818, 0.861]
MTF-GLP-HPM-PP	**[1.625, 2.114]**	**[2.074, 2.694]**	**[0.891, 0.911]**	[0.064, 0.082]	[0.079, 0.105]	[0.821, 0.863]
MTF-GLP-HPM	[2.319, 2.89]	[2.339, 2.913]	[0.786, 0.804]	**[0.018, 0.027]**	[0.059, 0.072]	**[0.904, 0.923]**
MTF-GLP-CBD	[1.934, 2.515]	[2.307, 2.994]	[0.865, 0.892]	**[0.024, 0.036]**	**[0.036, 0.047]**	**[0.919, 0.94]**
MSDCNN	**[1.677, 2.157]**	**[1.928, 2.385]**	[0.865, 0.905]	**[0.028, 0.037]**	**[0.047, 0.065]**	**[0.902, 0.924]**

TABLE 11. DETAILED STATISTICAL RESULTS, WITH A 95% CONFIDENCE INTERVAL, BASED ON THE GREEN VEGETATION DATA SET.

	REDUCED RESOLUTION			FULL RESOLUTION		
METHOD	**ERGAS**	**SAM**	**$Q2^N$-INDEX**	**D_λ**	**D_S**	**QNR**
PCA	[1.765, 2.11]	[1.938, 2.35]	[0.683, 0.732]	[0.053, 0.069]	[0.174, 0.247]	[0.706, 0.777]
HIS	[1.726, 2.046]	[1.711, 1.997]	[0.55, 0.599]	[0.131, 0.156]	[0.249, 0.323]	[0.577, 0.642]
Brovey	[1.635, 1.928]	[1.541, 1.786]	[0.569, 0.623]	[0.111, 0.136]	[0.238, 0.31]	[0.602, 0.665]
BDSD	[1.392, 1.71]	[1.571, 1.85]	**[0.796, 0.823]**	**[0.017, 0.024]**	**[0.061, 0.08]**	**[0.901, 0.92]**
GS	[1.602, 1.927]	[1.701, 1.966]	[0.697, 0.741]	[0.05, 0.065]	[0.173, 0.252]	[0.704, 0.783]
GSA	[1.505, 1.88]	[1.591, 1.911]	[0.747, 0.79]	[0.064, 0.081]	[0.185, 0.268]	[0.679, 0.762]
PRACS	[1.451, 1.776]	[1.609, 1.91]	[0.746, 0.786]	**[0.027, 0.036]**	[0.132, 0.203]	[0.771, 0.843]
HPF	**[1.379, 1.665]**	[1.467, 1.712]	[0.764, 0.794]	[0.056, 0.07]	[0.085, 0.108]	[0.832, 0.862]
SFIM	**[1.388, 1.689]**	**[1.463, 1.705]**	[0.765, 0.794]	[0.056, 0.07]	[0.084, 0.107]	[0.833, 0.863]
Indusion	[1.704, 2.035]	[1.662, 1.902]	[0.686, 0.724]	[0.052, 0.065]	**[0.065, 0.083]**	**[0.858, 0.885]**
ATWT	**[1.341, 1.624]**	**[1.432, 1.684]**	**[0.777, 0.806]**	[0.067, 0.081]	[0.112, 0.14]	[0.793, 0.827]
AWLP	[1.424, 1.718]	[1.624, 1.968]	[0.769, 0.799]	[0.053, 0.064]	[0.088, 0.118]	[0.829, 0.862]
ATWT-M2	[1.434, 1.691]	[1.596, 1.834]	[0.728, 0.766]	[0.053, 0.071]	[0.072, 0.094]	[0.844, 0.879]
ATWT-M3	[1.438, 1.691]	[1.686, 1.941]	[0.748, 0.781]	[0.046, 0.067]	[0.077, 0.094]	[0.849, 0.877]
MTF-GLP	**[1.313, 1.602]**	**[1.41, 1.667]**	**[0.782, 0.812]**	[0.069, 0.084]	[0.118, 0.148]	[0.784, 0.819]
MTF-GLP-HPM-PP	[1.338, 1.771]	**[1.409, 1.663]**	**[0.782, 0.812]**	[0.069, 0.083]	[0.117, 0.146]	[0.785, 0.82]
MTF-GLP-HPM	[1.408, 1.687]	[1.485, 1.719]	[0.741, 0.778]	**[0.021, 0.029]**	**[0.057, 0.07]**	**[0.904, 0.921]**
MTF-GLP-CBD	[1.534, 1.934]	[1.566, 1.885]	[0.76, 0.797]	**[0.025, 0.034]**	**[0.044, 0.057]**	**[0.912, 0.933]**
MSDCNN	**[1.068, 1.626]**	**[1.216, 1.543]**	**[0.78, 0.836]**	**[0.039, 0.05]**	**[0.048, 0.059]**	**[0.896, 0.911]**

TABLE 12. DETAILED STATISTICAL RESULTS, WITH A 95% CONFIDENCE INTERVAL, BASED ON THE WATER-SCENARIO DATA SET.

	REDUCED RESOLUTION			FULL RESOLUTION		
METHOD	**ERGAS**	**SAM**	**$Q2^N$-INDEX**	**D_λ**	**D_S**	**QNR**
PCA	[2.299, 3.528]	[2.443, 1.847]	[0.452, 0.546]	[0.101, 0.139]	[0.213, 0.28]	[0.629, 0.706]
HIS	**[1.381, 1.871]**	[0.122, 0.805]	[0.499, 0.604]	[0.189, 0.253]	[0.306, 0.396]	[0.465, 0.568]
Brovey	**[1.395, 1.895]**	**[0.087, 0.75]**	[0.502, 0.606]	[0.141, 0.194]	[0.272, 0.359]	[0.528, 0.628]
BDSD	[1.56, 2.232]	[0.21, 0.922]	[0.568, 0.668]	**[0.064, 0.084]**	**[0.061, 0.086]**	**[0.843, 0.874]**
GS	[1.51, 2.155]	[0.298, 0.981]	[0.503, 0.606]	**[0.086, 0.113]**	[0.217, 0.284]	[0.641, 0.714]
GSA	[1.362, 1.973]	[0.118, 0.772]	[0.559, 0.675]	[0.101, 0.136]	[0.23, 0.296]	[0.616, 0.693]
PRACS	**[1.262, 1.709]**	**[0.076, 0.701]**	**[0.588, 0.695]**	**[0.051, 0.07]**	[0.156, 0.212]	**[0.736, 0.8]**
HPF	[1.59, 2.248]	[0.115, 0.788]	[0.55, 0.648]	[0.135, 0.173]	[0.192, 0.267]	[0.614, 0.704]
SFIM	[3.885, 78.517]	[0.464, 1.003]	[0.537, 0.626]	[0.132, 0.171]	[0.189, 0.265]	[0.617, 0.71]
Indusion	[1.839, 2.609]	[0.23, 0.919]	[0.513, 0.603]	[0.115, 0.156]	[0.178, 0.243]	[0.646, 0.733]
ATWT	[1.544, 2.184]	[0.119, 0.785]	[0.557, 0.656]	[0.142, 0.179]	[0.207, 0.282]	[0.597, 0.686]
AWLP	[1.620, 2.331]	[0.134, 0.778]	**[0.583, 0.685]**	[0.153, 0.196]	[0.214, 0.293]	[0.577, 0.673]
ATWT-M2	[1.416, 1.957]	**[0.111, 0.762]**	[0.566, 0.671]	[0.163, 0.206]	[0.188, 0.24]	[0.609, 0.681]
ATWT-M3	[1.501, 2.092]	[0.114, 0.767]	**[0.578, 0.683]**	[0.131, 0.17]	**[0.126, 0.174]**	[0.692, 0.759]
MTF-GLP	**[1.293, 1.869]**	**[0.125, 0.739]**	**[0.582, 0.69]**	[0.136, 0.175]	[0.221, 0.293]	[0.59, 0.679]
MTF-GLP-HPM-PP	[8.758, 48.768]	[0.647, 1.036]	[0.564, 0.662]	[0.133, 0.174]	[0.217, 0.291]	[0.594, 0.685]
MTF-GLP-HPM	[4.747, 13.646]	[0.174, 0.837]	[0.526, 0.614]	**[0.079, 0.108]**	**[0.115, 0.171]**	**[0.744, 0.817]**
MTF-GLP-CBD	[1.487, 2.284]	**[0.137, 0.7]**	**[0.618, 0.741]**	**[0.056, 0.081]**	**[0.107, 0.149]**	**[0.784, 0.845]**
MSDCNN	**[1.184, 1.804]**	[0.798, 1.153]	[0.483, 0.626]	[0.095, 0.118]	**[0.054, 0.07]**	**[0.824, 0.849]**

TABLE 13. DETAILED STATISTICAL RESULTS, WITH A 95% CONFIDENCE INTERVAL, BASED ON THE UNLABELED DATA SET WITH MIXED SURFACE FEATURES.

	REDUCED RESOLUTION			FULL RESOLUTION		
METHOD	**ERGAS**	**SAM**	**$Q2^N$-INDEX**	**D_λ**	**D_S**	**QNR**
PCA	[2.329, 2.682]	[2.922, 3.446]	[0.69, 0.724]	[0.068, 0.079]	[0.185, 0.225]	[0.721, 0.762]
HIS	[2.003, 2.216]	[2.067, 2.282]	[0.61, 0.638]	[0.147, 0.16]	[0.252, 0.281]	[0.609, 0.639]
Brovey	[1.911, 2.108]	[1.885, 2.075]	[0.628, 0.655]	[0.128, 0.139]	[0.235, 0.262]	[0.639, 0.667]
BDSD	**[1.633, 1.827]**	[1.904, 2.115]	**[0.802, 0.82]**	**[0.025, 0.029]**	**[0.058, 0.067]**	**[0.908, 0.918]**
GS	[1.948, 2.174]	[2.12, 2.382]	[0.709, 0.735]	[0.073, 0.083]	[0.187, 0.219]	[0.721, 0.755]
GSA	[1.681, 1.882]	[1.86, 2.059]	**[0.777, 0.799]**	[0.085, 0.098]	[0.183, 0.209]	[0.719, 0.75]
PRACS	**[1.564, 1.746]**	[1.774, 1.965]	**[0.795, 0.814]**	**[0.036, 0.042]**	[0.112, 0.128]	[0.837, 0.856]
HPF	[1.702, 1.897]	**[1.769, 1.956]**	[0.754, 0.777]	[0.092, 0.103]	[0.127, 0.144]	[0.771, 0.794]
SFIM	[1.526, 3.34]	[1.769, 1.957]	[0.756, 0.779]	[0.092, 0.104]	[0.125, 0.142]	[0.773, 0.796]
Indusion	[2.094, 2.318]	[1.986, 2.181]	[0.688, 0.713]	[0.093, 0.105]	[0.089, 0.104]	[0.805, 0.826]
ATWT	**[1.649, 1.842]**	**[1.736, 1.923]**	[0.764, 0.788]	[0.105, 0.116]	[0.157, 0.176]	[0.732, 0.757]
AWLP	[1.707, 1.917]	[1.834, 2.041]	[0.768, 0.79]	[0.097, 0.108]	[0.143, 0.161]	[0.752, 0.776]
ATWT-M2	[1.781, 1.972]	[1.952, 2.156]	[0.745, 0.761]	[0.064, 0.075]	[0.083, 0.095]	[0.84, 0.859]
ATWT-M3	[1.774, 1.963]	[2.014, 2.221]	[0.76, 0.775]	[0.067, 0.079]	**[0.072, 0.083]**	**[0.848, 0.867]**
MTF-GLP	**[1.618, 1.812]**	**[1.715, 1.902]**	[0.771, 0.795]	[0.107, 0.119]	[0.164, 0.182]	[0.725, 0.749]
MTF-GLP-HPM-PP	[1.751, 3.483]	**[1.716, 1.907]**	[0.772, 0.797]	[0.107, 0.119]	[0.161, 0.18]	[0.727, 0.751]
MTF-GLP-HPM	[1.925, 3.42]	[1.814, 2.005]	[0.748, 0.766]	**[0.042, 0.051]**	**[0.062, 0.073]**	**[0.882, 0.898]**
MTF-GLP-CBD	[1.848, 2.075]	[1.877, 2.098]	**[0.777, 0.797]**	**[0.034, 0.039]**	**[0.056, 0.064]**	**[0.9, 0.912]**
MSDCNN	**[1.449, 1.721]**	**[1.663, 1.884]**	**[0.806, 0.834]**	**[0.039, 0.045]**	**[0.052, 0.058]**	**[0.9, 0.91]**

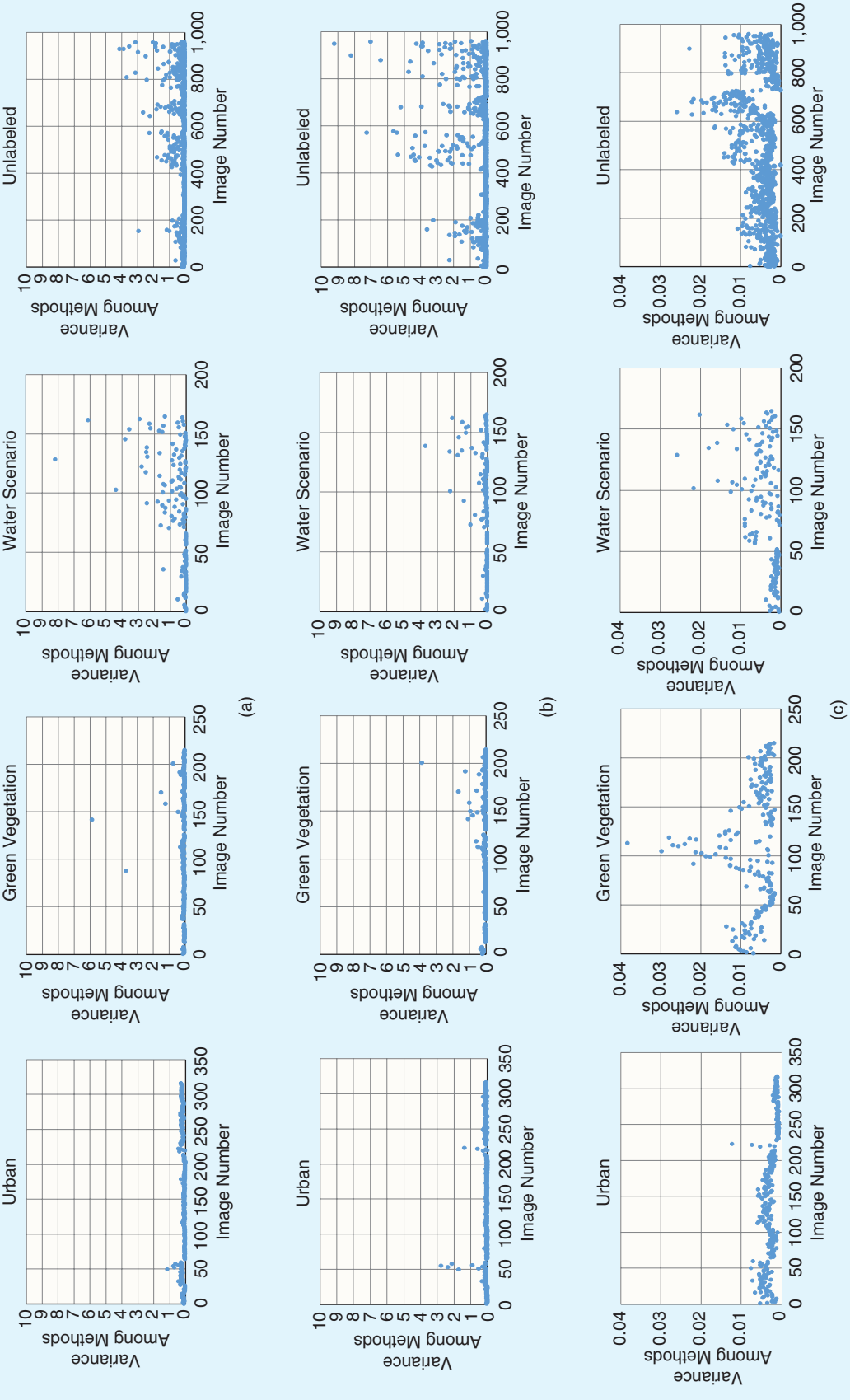

FIGURE 16. A heterogeneity analysis among pansharpening methods for thematic data sets (urban, green vegetation, water scenario, and unlabeled with mixed surface features) based on reduced-resolution images. (a) The variance distribution in terms of the ERGAS. (b) The variance distribution in terms of the SAM. (c) The variance distribution in terms of the $Q2^n$ index.

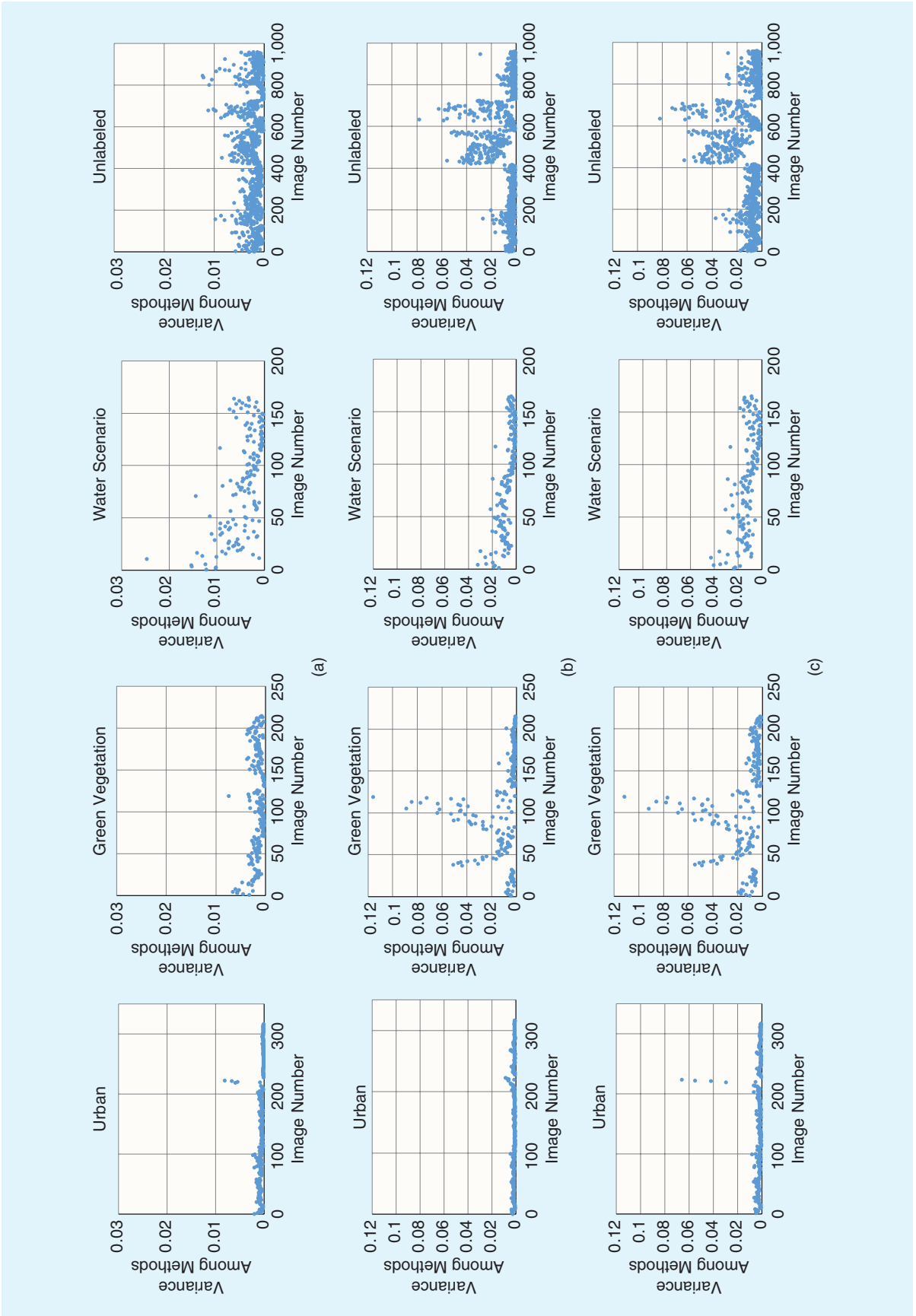

FIGURE 17. The heterogeneity analysis among pansharpening methods for thematic data sets (urban, green vegetation, water scenario, and unlabeled) based on full-resolution images. (a) The variance distribution in terms of D_{λ}. (b) The variance distribution in terms of D_s. (c) The variance distribution in terms of the QNR.

respectively. It is shown that the distribution of the variances in the urban data set is relatively consistent in terms of the ERGAS, SAM, and $Q2^n$ index, and the variance values are small. This indicates that the performance of different pansharpening methods for building-oriented remote sensing images is similar. For the images in the green vegetation and water-scenario data sets, the distribution of the variances is scattered, and the variance values are relatively large. In addition, for the unlabeled data set with mixed surfaces, the heterogeneity among the pansharpening methods is the largest. This indicates that different pansharpening methods have a more obvious difference for the green vegetation, water-oriented, and mixed-surface images.

In addition, heterogeneity analysis based on full-resolution images was implemented; it is shown in Figure 17. The first column denotes the distribution of the variance for different thematic data sets in terms of D_λ. The second and third columns describe the distribution for different thematic data sets in terms of D_s and the QNR, respectively. It can be seen that, for the urban data set, most of the variances are small, and their distribution is consistent. This indicates that different pansharpening methods have a robust performance for building scenes. For the green vegetation and water-scenario data sets, the distribution of the variance of almost all the evaluation indices is relatively scattered. This indicates that, on the one hand, there is a relatively larger performance difference for the various pansharpening methods. On the other hand, it is relatively more sensitive in spectral fidelity and spatial enhancement for the green vegetation and water scenes compared with the building scenes. Overall, this is consistent with the results based on the reduced-resolution data sets.

DISCUSSION

This article proposed a large-scale data set for pansharpening, and the performance of 19 popular pansharpening methods was evaluated based on the proposed data set.

1) Contrary to the existing assessment that depends on a few images, the pansharpening methods were evaluated based on large numbers of remote sensing images from different satellites and with typical thematic surface features.
2) The experiment results show that pansharpening methods have relatively different performance for various remote sensing satellite images and thematic images. On the whole, the BDSD and MSDCNN methods generally show excellent performance in quantitative evaluation. The SFIM, MTF-GLP-HPM, and MTF-GLP-HPM-PP methods are unstable due to the common HPM spatial-detail injection scheme. This is because the HPM may introduce outliers in some cases. Specifically, for the HPM, the HR Pan is first matched to the MS based on moment matching to remove the radiation difference, and the matched Pan image $\tilde{\mathbf{P}}$ is divided by its low-pass version $\tilde{\mathbf{P}}_l$ to calculate the spatial-detail injection weight, represented as $\tilde{\mathbf{P}}/\tilde{\mathbf{P}}_l$. For most cases, the preceding two-step operations contribute to the improvement of the pansharpening methods; however, matched Pan images that have many negative values generally lead to some pixel values closing to zeros in the $\tilde{\mathbf{P}}_l$, and then the outliers are generally introduced. In addition, the experiment results show a superior performance for the DL-based MSDCNN method. Although this cannot completely represent the performance of all the DL-based pansharpening methods, it shows their advantages to some extent.
3) The experiment results show that inconsistent performance exists among some quantitative-evaluation indices. For the MSDCNN in the *Ikonos* experiment, the SAM shows excellent results; however, the ERGAS and $Q2^n$ index show relatively poorer performance. This is because quantitative-evaluation indices have different emphases on the evaluation. On the one hand, some evaluation indices focus on spatial-quality assessment, and others put emphasis on spectral distortion evaluation. On the other hand, quantitative-evaluation indices generally consider different factors. For example, the ERGAS mainly assesses the overall radiation difference between the fused image and the reference image, and the SAM evaluates the distortions of the inner-band relations. Therefore, a comprehensive analysis of several evaluation indices is necessary for the performance assessment of pansharpening methods. Effective and robust comprehensive quality evaluation for pansharpening should be further studied, especially the nonreference quantitative-evaluation methods.
4) To date, pansharpening methods have been developed for nearly 40 years, and a large number of them have been proposed. Moreover, many CS- and MRA-based pansharpening methods have been commercialized in professional remote sensing software. However, there are still some essential and urgent problems to be solved.
   - Application-oriented pansharpening methods should be given attention. Pansharpening is generally a fundamental processing progress in various applications, and different applications may have contrasting requirements for high spectral fidelity and high spatial-detail enhancement. Therefore, application-oriented pansharpening methods should be developed.
   - Robustness-oriented pansharpening methods should be devised. The experiment results based on the large-scale data set in this article show that most of the pansharpening methods may have high accuracy for some images and poor performance in the case of other images, such as those from different thematic scenarios. Therefore, robust pansharpening methods for different thematic scenarios, misregistration, disparate imaging times, noise, cloud contaminations, and so forth should be studied.

- DL-based pansharpening methods have been attracting ever-increasing attention in recent years, and they generally have high precision. However, they are limited by the large-scale data set, and time-consuming network training is also generally performed again for various tasks, such as different remote sensing satellites and thematic scenarios. Therefore, robust, adjustable DL-based methods should be researched.

5) Finally, this article proposed a large-scale benchmark data set for pansharpening, and this can contribute to research in many aspects.
   - The proposed large-scale data set can be utilized to comprehensively assess the performance of pansharpening methods.
   - The data set can be used to develop new pansharpening methods, especially data-driven pansharpening approaches such as DL-based techniques.
   - The data set will assist researchers, especially those who have difficulty obtaining very-high-resolution remote sensing images.

CONCLUSIONS

This article presented a large-scale benchmark data set for performance evaluation of pansharpening methods. The proposed data set has 2,270 pairs of HR Pan/LR MS images. To assist researchers, the proposed data set was grouped in two ways, i.e., different remote sensing satellites and typical thematic surface features. The data set for the different remote sensing satellites consists of 200 pairs of *Ikonos*, 500 pairs of *QuickBird*, 410 pairs of *GF-1*, 500 pairs of *WorldView-4*, 500 pairs of *WorldView-2*, and 160 pairs of *WorldView-3* HR Pan/LR MS images. The thematic data set is composed of the urban data set with 510 pairs of HR Pan/LR MS images, the green vegetation data set with 258 pairs of HR Pan/LR MS images, the water-scenario data set with 318 pairs of HR Pan/LR MS images, and the unlabeled data set with mixed surface features and 1,184 pairs of HR Pan/LR MS images. In addition, this article reviewed pansharpening methods, including the CS-, MRA-, VO-, and DL-based approaches. Finally, 19 pansharpening methods were evaluated and statistically analyzed based on the proposed data set.

ACKNOWLEDGMENTS

We would like to thank the Global Land Cover Facility (https://geog.umd.edu/feature/global-land-cover-facility-(glcf)) and DigitalGlobe (http://www.digitalglobe.com/product-samples) for providing the partial data used in this article. We would also like to thank the anonymous reviewers for their careful reading and valuable comments. This work was supported by the National Natural Science Foundation of China (grant 41801252), by the Natural Science Foundation of Ningbo City (grant 2019A610098), and, in part, by the K.C. Wong Magna Fund at Ningbo University.

AUTHOR INFORMATION

Xiangchao Meng (mengxiangchao@nbu.edu.cn) is with the Faculty of Electrical Engineering and Computer Science, Ningbo University, China, and the College of Electrical and Information Engineering, Hunan University, Changsha, China. He is a Member of the IEEE.

Yiming Xiong (xiongyiming6@163.com) is with the Faculty of Electrical Engineering and Computer Science, Ningbo University, China.

Feng Shao (shaofeng@ nbu.edu.cn) is with the Faculty of Electrical Engineering and Computer Science, Ningbo University, China. He is a Member of the IEEE.

Huanfeng Shen (shenhf@whu.edu.cn) is with the School of Resources and Environmental Sciences, Wuhan University, China. He is a Senior Member of the IEEE.

Weiwei Sun (sunweiwei@nbu.edu.cn) is with the Department of Geography and Spatial Information Techniques, Ningbo University, China. He is a Member of the IEEE.

Gang Yang (yanggang@nbu.edu.cn) is with the Department of Geography and Spatial Information Techniques, Ningbo University, China. He is a Member of the IEEE.

Qiangqiang Yuan (yqiang86@gmail.com) is with the School of Geodesy and Geomatics, Wuhan University, China. He is a Member of the IEEE.

Randi Fu (furandi@nbu.edu.cn) is with the Faculty of Electrical Engineering and Computer Science, Ningbo University, China.

Hongyan Zhang (zhanghongyan@whu.edu.cn) is with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing and the Collaborative Innovation Center for Geospatial Technology, Wuhan University, China. He is a Senior Member of the IEEE.

REFERENCES

[1] X. Meng, H. Shen, H. Li, L. Zhang, and R. Fu, "Review of the pansharpening methods for remote sensing images based on the idea of meta-analysis: Practical discussion and challenges," *Inform. Fusion*, vol. 46, pp. 102–113, Mar. 2019. doi: 10.1016/j.inffus.2018.05.006.

[2] H. Shen, X. Meng, and L. Zhang, "An integrated framework for the spatio-temporal-spectral fusion of remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 12, pp. 7135–7148, 2016. doi: 10.1109/TGRS.2016.2596290.

[3] G. Vivone et al., "A critical comparison among pansharpening algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2565–2586, 2015. doi: 10.1109/TGRS.2014.2361734.

[4] Y. Zhang, "Understanding image fusion," *Photogramm. Eng. Remote Sens.*, vol. 70, no. 6, pp. 657–661, 2004.

[5] G. Cliche, F. Bonn, and P. Teillet, "Integration of the SPOT panchromatic channel into its multispectral mode for image sharpness enhancement," *Photogramm. Eng. Remote Sens.*, vol. 51, no. 3, pp. 311–316, 1985.

[6] R. Haydn, G. W. Dalke, J. Henkel, and J. E. Bare, "Application of the IHS color transform to the processing of multisensor data

and image enhancement," in *Proc. Int. Symp. Remote Sensing Arid and Semi-Arid Lands*, Cairo, 1982, pp. 599–616.

[7] R. A. Schowengerdt, "Reconstruction of multispatial, multispectral image data using spatial frequency content," *Photogramm. Eng. Remote Sens.*, vol. 46, no. 10, pp. 1325–1334, 1980.

[8] R. Welch and M. Ehlers, "Merging multiresolution SPOT HRV and Landsat TM data," *Photogramm. Eng. Remote Sens.*, vol. 53, no. 3, pp. 301–303, 1987.

[9] P. S. Chavez Jr., and J. A. Bowell, "Comparison of the spectral information content of Landsat Thematic Mapper and SPOT for three different sites in the Phoenix, Arizona region," *Photogramm. Eng. Remote Sens.*, vol. 54, no. 12, pp. 1699–1708, 1988.

[10] D. L. Hall and J. Llinas, "An introduction to multisensor data fusion," *Proc. IEEE*, vol. 85, no. 1, pp. 6–23, 1997. doi: 10.1109/5.554205.

[11] C. Pohl and J. L. van Genderen, "Multisensor image fusion in remote sensing: Concepts, methods and applications," *Int. J. Remote Sens.*, vol. 19, no. 5, pp. 823–854, 1998. doi: 10.1080/014311698215748.

[12] L. Wald, "Some terms of reference in data fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1190–1193, 1999. doi: 10.1109/36.763269.

[13] B. Garguet-Duport, J. Girel, J.-M. Chassery, and G. Pautou, "The use of multiresolution analysis and wavelets transform for merging SPOT panchromatic and multispectral image data," *Photogramm. Eng. Remote Sens.*, vol. 62, no. 9, pp. 1057–1066, 1996.

[14] H. Li, B. S. Manjunath, and S. K. Mitra, "Multisensor image fusion using the wavelet transform," *Graph. Models Image Process.*, vol. 57, no. 3, pp. 235–245, 1995. doi: 10.1006/gmip.1995.1022.

[15] J. Nunez, X. Otazu, O. Fors, A. Prades, V. Pala, and R. Arbiol, "Multiresolution-based image fusion with additive wavelet decomposition," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1204–1211, 1999. doi: 10.1109/36.763274.

[16] T. Ranchin and L. Wald, "The wavelet transform for the analysis of remotely sensed images," *Int. J. Remote Sens.*, vol. 14, no. 3, pp. 615–619, 1993. doi: 10.1080/01431169308904362.

[17] P. S. Chavez Jr., S. C. Sides, and J. A. Anderson, "Comparison of three different methods to merge multiresolution and multispectral data: Landsat TM and SPOT panchromatic," *Photogramm. Eng. Remote Sens.*, vol. 57, no. 3, pp. 295–303, 1991.

[18] J. Vrabel, "Multispectral imagery band sharpening study," *Photogramm. Eng. Remote Sens.*, vol. 62, no. 9, pp. 1075–1084, 1996.

[19] T. M. Tu, S. C. Su, H. C. Shyu, and P. S. Huang, "A new look at IHS-like image fusion methods," *Inform. Fusion*, vol. 2, no. 3, pp. 177–186, 2001. doi: 10.1016/S1566-2535(01)00036-7.

[20] C. Ballester, V. Caselles, L. Igual, J. Verdera, and B. Rougé, "A variational model for P+XS image fusion," *Int. J. Comput. Vis.*, vol. 69, no. 1, pp. 43–58, 2006. doi: 10.1007/s11263-006-6852-x.

[21] X. Meng, H. Shen, Q. Yuan, H. Li, L. Zhang, and W. Sun, "Pansharpening for cloud-contaminated very high-resolution remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 5, 2019. doi: 10.1109/TGRS.2018.2878007.

[22] S. Li and B. Yang, "A new pan-sharpening method using a compressed sensing technique," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 2, pp. 738–746, 2011. doi: 10.1109/TGRS.2010.2067219.

[23] X. X. Zhu and R. Bamler, "A sparse image fusion algorithm with application to pan-sharpening," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2827–2836, 2013. doi: 10.1109/TGRS.2012.2213604.

[24] W. Huang, L. Xiao, Z. Wei, H. Liu, and S. Tang, "A new pan-sharpening method with deep neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 5, pp. 1037–1041, 2015. doi: 10.1109/LGRS.2014.2376034.

[25] Y. Liu, X. Chen, Z. Wang, Z. J. Wang, R. K. Ward, and X. Wang, "Deep learning for pixel-level image fusion: Recent advances and future prospects," *Inform. Fusion*, vol. 42, pp. 158–173, July 2018. doi: 10.1016/j.inffus.2017.10.007.

[26] G. Masi, D. Cozzolino, L. Verdoliva, and G. Scarpa, "Pansharpening by convolutional neural networks," *Remote Sens.*, vol. 8, no. 7, pp. 594:1–594:22, July 2016. doi: 10.3390/rs8070594.

[27] Y. Wei, Q. Yuan, H. Shen, and L. Zhang, "Boosting the accuracy of multi-spectral image pan-sharpening by learning a deep residual network," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 10, pp. 1795–1799, 2017. doi: 10.1109/LGRS.2017.2736020.

[28] F. Bovolo, L. Bruzzone, L. Capobianco, A. Garzelli, and S. Marchesi, "Analysis of the effects of pansharpening in change detection on VHR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 1, pp. 53–57, 2010. doi: 10.1109/LGRS.2009.2029248.

[29] L. Bruzzone and F. Bovolo, "A novel framework for the design of change-detection systems for very-high-resolution remote sensing images," *Proc. IEEE*, vol. 101, no. 3, pp. 609–630, 2013. doi: 10.1109/JPROC.2012.2197169.

[30] T. Ranchin and L. Wald, "Benefits of fusion of high spatial and spectral resolutions images for urban mapping," in *Proc. 26th Int. Symp. Remote Sensing Environment and 18th Annu. Symp. Canadian Remote Sensing Society*, Vancouver, BC, 1996, pp. 262–265.

[31] P. Sirguey, R. Mathieu, Y. Arnaud, M. M Khan, and J. Chanussot, "Improving MODIS spatial resolution for snow mapping using wavelet fusion and ARSIS concept," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 1, pp. 78–82, 2008. doi: 10.1109/LGRS.2007.908884.

[32] S. Baronti, B. Aiazzi, M. Selva, A. Garzelli, and L. Alparone , "A theoretical analysis of the effects of aliasing and misregistration on pansharpened imagery," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 3, pp. 446–453, 2011. doi: 10.1109/JSTSP.2011.2104938.

[33] H. Zhang and B. Huang, "A new look at image fusion methods from a Bayesian perspective," *Remote Sens.*, vol. 7, no. 6, pp. 6828–6861, 2015. doi: 10.3390/rs70606828.

[34] A. Garzelli, "A review of image fusion algorithms based on the super-resolution paradigm," *Remote Sens.*, vol. 8, no. 10, pp. 797:1–797:20, 2016. doi: 10.3390/rs8100797.

[35] C. Kwan, J. H. Choi, S. Chan, J. Zhou, and B. Budavari, "Resolution enhancement for hyperspectral images: A super-resolution and fusion approach," in *Proc. IEEE Int. Conf. Acoustics, Speech*

and Signal Processing, New Orleans, 2017, pp. 6180–6184. doi: 10.1109/ICASSP.2017.7953344.

[36] J. Zhou, C. Kwan, and B. Budavari, "Hyperspectral image super-resolution: A hybrid color mapping approach," *J. Appl. Remote Sens.*, vol. 10, no. 3, pp. 035024:1–035024:20, 2016. doi: 10.1117/1.JRS.10.035024.

[37] W. Carper, T. Lillesand, and R. Kiefer, "The use of Intensity-Hue-Saturation transformations for merging spot panchromatic and multispectral image data," *Photogramm. Eng. Remote Sens.*, vol. 56, no. 4, pp. 459–467, 1990.

[38] B. Aiazzi, S. Baronti, and M. Selva, "Improving component substitution pansharpening through multivariate regression of MS +Pan data," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3230–3239, 2007. doi: 10.1109/TGRS.2007.901007.

[39] C. A. Laben and B. V. Brower, "Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening," U.S. Patent 6 011 875 A, 2000.

[40] X. Otazu, M. González-Audícana, O. Fors, and J. Nunez, "Introduction of sensor spectral response into image fusion methods. Application to wavelet-based methods," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 10, pp. 2376–2385, 2005. doi: 10.1109/TGRS.2005.856106.

[41] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva, "MTF-tailored multiscale fusion of high-resolution MS and pan imagery," *Photogramm. Eng. Remote Sens.*, vol. 72, no. 5, pp. 591–596, 2006. doi: 10.14358/PERS.72.5.591.

[42] L. Alparone, B. Aiazzi, S. Baronti, and A. Garzelli, "Sharpening of very high resolution images with spectral distortion minimization," in Proc. *IEEE Int. Geoscience and Remote Sensing Symp.*, Toulouse, France 2003, pp. 458–460.

[43] N. Mascarenhas, G. Banon, and A. Candeias, "Multispectral image data fusion under a Bayesian approach," *Int. J. Remote Sens.*, vol. 17, no. 8, pp. 1457–1471, 1996. doi: 10.1080/01431169608948717.

[44] L. Zhang, H. Shen, W. Gong, and H. Zhang, "Adjustable model-based fusion method for multispectral and panchromatic images," *IEEE Trans. Syst. Man Cybern. B, Cybern.*, vol. 42, no. 6, pp. 1693–1704, 2012. doi: 10.1109/TSMCB.2012.2198810.

[45] Q. Yuan, Y. Wei, X. Meng, H. Shen, and L. Zhang, "A multiscale and multidepth convolutional neural network for remote sensing imagery pan-sharpening," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 3, pp. 978–989, 2018. doi: 10.1109/JSTARS.2018.2794888.

[46] J. Zhong, B. Yang, G. Huang, F. Zhong, and Z. Chen, "Remote sensing image fusion with convolutional neural network," *Sens. Imaging*, vol. 17, no. 1, pp. 10:1–10:16, 2016. doi: 10.1007/s11220-016-0135-6.

[47] B. Aiazzi, L. Alparone, S. Baronti, R. Garzelli, and M. Selva, "Twenty-five years of pansharpening: A critical review and new developments," in *Signal and Image Processing for Remote Sensing*, 2nd ed. C.-H. Chen, Ed. Boca Raton, FL: CRC, 2012, pp. 533–548.

[48] L. Alparone, B. Aiazzi, S. Baronti, and A. Garzelli, *Remote Sensing Image Fusion*. Boca Raton, FL: CRC, 2015.

[49] L. Alparone, S. Baronti, B. Aiazzi, and A. Garzelli, "Spatial methods for multispectral pansharpening: Multiresolution analysis demystified," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 5, pp. 2563–2576, 2016. doi: 10.1109/TGRS.2015.2503045.

[50] L. Alparone, L. Wald, J. Chanussot, C. Thomas, P. Gamba, and L. M. Bruce, "Comparison of pansharpening algorithms: Outcome of the 2006 GRS-S data-fusion contest," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3012–3021, 2007. doi: 10.1109/TGRS.2007.904923.

[51] I. Amro, J. Mateos, M. Vega, R. Molina, and A. K. Katsaggelos, "A survey of classical methods and new trends in pansharpening of multispectral images," *J. Adv. Signal Process.*, vol. 2011, no. 1, pp. 79:1–79:22, 2011.

[52] M. Dalla Mura, S. Prasad, F. Pacifici, P. Gamba, and J. Chanussot, "Challenges and opportunities of multimodality and data fusion in remote sensing," *Proc. IEEE*, vol. 103, no. 9, pp. 1585–1601, 2015. doi: 10.1109/JPROC.2015.2462751.

[53] M. Ehlers, S. Klonus, P. Johan Åstrand, and P. Rosso, "Multi-sensor image fusion for pansharpening in remote sensing," *Int. J. Image Data Fusion*, vol. 1, no. 1, pp. 25–45, 2010. doi: 10.1080/19479830903561985.

[54] H. Ghassemian, "A review of remote sensing image fusion methods," *Inform. Fusion*, vol. 32, pp. 75–89, Nov. 2016. doi: 10.1016/j.inffus.2016.03.003.

[55] S. Li, X. Kang, L. Fang, J. Hu, and H. Yin, "Pixel-level image fusion: A survey of the state of the art," *Inform. Fusion*, vol. 33, pp. 100–112, Jan. 2017. doi: 10.1016/j.inffus.2016.05.004.

[56] L. Loncan et al., "Hyperspectral pansharpening: A review," *IEEE Geosci. Remote Sens. Mag. (replaces Newsletter)*, vol. 3, no. 3, pp. 27–46, 2015. doi: 10.1109/MGRS.2015.2440094.

[57] M. Schmitt and X. X. Zhu, "Data fusion and remote sensing: An ever-growing relationship," *IEEE Geosci. Remote Sens. Mag. (replaces Newsletter)*, vol. 4, no. 4, pp. 6–23, 2016. doi: 10.1109/MGRS.2016.2561021.

[58] C. Thomas, T. Ranchin, L. Wald, and J. Chanussot, "Synthesis of multispectral images to high spatial resolution: A critical review of fusion methods based on remote sensing physics," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 5, pp. 1301–1312, 2008. doi: 10.1109/TGRS.2007.912448.

[59] Z. Wang, D. Ziou, C. Armenakis, D. Li, and Q. Li, "A comparative analysis of image fusion methods," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 6, pp. 1391–1402, 2005. doi: 10.1109/TGRS.2005.846874.

[60] J. Zhang, "Multi-source remote sensing data fusion: Status and trends," *Int. J. Image Data Fusion*, vol. 1, no. 1, pp. 5–24, 2010. doi: 10.1080/19479830903561035.

[61] A. Garzelli, F. Nencini, and L. Capobianco, "Optimal MMSE pan sharpening of very high resolution multispectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 1, pp. 228–236, 2008. doi: 10.1109/TGRS.2007.907604.

[62] J. Choi, K. Yu, and Y. Kim, "A new adaptive component-substitution-based satellite image fusion by using partial replacement," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 1, pp. 295–309, 2011. doi: 10.1109/TGRS.2010.2051674.

[63] P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, vol. 31, no. 4, pp. 532–540, 1983. doi: 10.1109/TCOM.1983.1095851.

[64] B. Aiazzi, S. Baronti, F. Lotti, and M. Selva, "A comparison between global and context-adaptive pansharpening of mul-

tispectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 2, pp. 302–306, 2009. doi: 10.1109/LGRS.2008.2012003.

[65] T. Ranchin and L. Wald, "Fusion of high spatial and spectral resolution images: The ARSIS concept and its implementation," *Photogramm. Eng. Remote Sens.*, vol. 66, no. 1, pp. 49–61, 2000.

[66] G. Vivone, "Multispectral and hyperspectral pansharpening: A critical examination and new developments," Ph.D. dissertation, Dipartimento di Ingegneria dell'Informazione, Ingegneria Elettrica e Matematica Applicata, Università degli Studi di Salerno, Fisciano, Italy, 2014.

[67] G. Vivone, R. Restaino, M. Dalla Mura, G. Licciardi, and J. Chanussot, "Contrast and error-based fusion schemes for multispectral image pansharpening," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 5, pp. 930–934, 2014. doi: 10.1109/LGRS.2013.2281996.

[68] F. Fang, F. Li, C. Shen, and G. Zheng, "A variational approach for pan-sharpening," *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2822–2834, 2013. doi: 10.1109/TIP.2013.2258355.

[69] D. Fasbender, J. Radoux, and P. Bogaert, "Bayesian data fusion for adaptable image pansharpening," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 6, pp. 1847–1857, 2008. doi: 10.1109/TGRS.2008.917131.

[70] P. Liu, L. Xiao, J. Zhang, and B. Naz, "Spatial-Hessian-feature-guided variational model for pan-sharpening," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 4, pp. 2235–2253, 2016. doi: 10.1109/TGRS.2015.2497966.

[71] F. Palsson, J. R. Sveinsson, and M. O. Ulfarsson, "A new pansharpening algorithm based on total variation," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 1, pp. 318–322, 2014. doi: 10.1109/LGRS.2013.2257669.

[72] M. Guo, H. Zhang, J. Li, L. Zhang, and H. Shen., "An online coupled dictionary learning approach for remote sensing image fusion," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1284–1294, 2014. doi: 10.1109/JSTARS.2014.2310781.

[73] C. Jiang, H. Zhang, H. Shen, and L. Zhang, "A practical compressed sensing-based pan-sharpening method," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 4, pp. 629–633, 2012.

[74] Z. Li and H. Leung, "Fusion of multispectral and panchromatic images using a restoration-based method," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 5, pp. 1482–1491, 2009. doi: 10.1109/TGRS.2008.2005639.

[75] X. Meng, H. Shen, L. Zhang, Q. Yuan, and H. Li, "A unified framework for spatio-temporal-spectral fusion of remote sensing images," in *Proc. IEEE Int. Geoscience and Remote Sensing Symp. (IGARSS)*, Milan, Italy, 2015, pp. 2584–2587. doi: 10.1109/IGARSS.2015.7326340.

[76] Q. Wei, "Bayesian fusion of multi-band images: A powerful tool for super-resolution," Ph.D. dissertation, Institut national polytechnique de Toulouse (INPT), Univ. Toulouse, France, 2015.

[77] C. Chen, Y. Li, W. Liu, and J. Huang., "Image fusion with local spectral consistency and dynamic gradient sparsity," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Columbus, OH, 2014, pp. 2760–2765. doi: 10.1109/CVPR.2014.347.

[78] F. Fang, F. Li, G. Zhang, and C. Shen, "A variational method for multisource remote-sensing image fusion," *Int. J. Remote Sens.*, vol. 34, no. 7, pp. 2470–2486, 2013. doi: 10.1080/01431161.2012.744882.

[79] M. Moeller, T. Wittman, and A. L. Bertozzi, "Variational wavelet pan-sharpening," submitted for publication.

[80] M. Vega, J. Mateos, R. Molina, and A. Katsaggelos, "Super resolution of multispectral images using ℓ1 image models and interband correlations," *J. Signal Process. Syst.*, vol. 65, no. 3, pp. 509–523, 2011. doi: 10.1007/s11265-010-0554-x.

[81] E. J. Candès, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Commun. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, 2006. doi: 10.1002/cpa.20124.

[82] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006. doi: 10.1109/TIT.2006.871582.

[83] X. X. Zhu, C. Grohnfeldt, and R. Bamler, "Exploiting joint sparsity for pansharpening: The J-SparseFI algorithm," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 5, pp. 2664–2681, 2016. doi: 10.1109/TGRS.2015.2504261.

[84] H. Shen, L. Zhang, B. Huang, and P. Li, "A MAP approach for joint motion estimation, segmentation, and super resolution," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 479–490, 2007. doi: 10.1109/TIP.2006.888334.

[85] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015. doi: 10.1038/nature14539.

[86] L. Zhang, L. Zhang, and B. Du, "Deep learning for remote sensing data: A technical tutorial on the state of the art," *IEEE Geosci. Remote Sens. Mag.*, vol. 4, no. 2, pp. 22–40, 2016. doi: 10.1109/MGRS.2016.2540798.

[87] X. X. Zhu et al., "Deep learning in remote sensing: A comprehensive review and list of resources," *IEEE Geosci. Remote Sens. Mag. (replaces Newsletter)*, vol. 5, no. 4, pp. 8–36, 2017. doi: 10.1109/MGRS.2017.2762307.

[88] L. Wald, "Quality of high resolution synthesised images: Is there a simple criterion?" in *Proc. 3rd Conf. "Fusion of Earth data: Merging Point Measurements, Raster Maps and Remotely Sensed Images"*, Sophia Antipolis, France, 2000, pp. 99–103.

[89] C. Dong, C. C. Loy, K. He, and X. Tang., "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, 2016. doi: 10.1109/TPAMI.2015.2439281.

[90] N. Li, N. Huang, and L. Xiao, "PAN-sharpening via residual deep learning," in *Proc. IEEE Int. Geoscience and Remote Sensing Symp. (IGARSS)*, Fort Worth, TX, 2017, pp. 5133–5136. doi: 10.1109/IGARSS.2017.8128158.

[91] X. Liu, Y. Wang, and Q. Liu, "Remote sensing image fusion based on two-stream fusion network," in *Proc. Int. Conf. Multimedia Modeling*, 2018, pp. 428–439.

[92] Y. Rao, L. He, and J. Zhu, "A residual convolutional neural network for pan-shaprening," in *Proc. Int. Workshop on Remote Sensing Intelligent Processing (RSIP)*, Shanghai, China 2017, pp. 1–4. doi: 10.1109/RSIP.2017.7958807.

[93] Z. Shao and J. Cai, "Remote sensing image fusion with deep convolutional neural network," *IEEE J. Sel. Topics Appl. Earth*

Observ. Remote Sens., vol. 11, no. 5, pp. 1656–1669, 2018. doi: 10.1109/JSTARS.2018.2805923.

[94] J. Yang, X. Fu, Y. Hu, Y. Huang, and X. Ding, "PanNet: A deep network architecture for pan-sharpening," in *Proc. IEEE Int. Conf. Computer Vision (ICCV)*, Venice, Italy, 2017, pp. 1753–1761. doi: 10.1109/ICCV.2017.193.

[95] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2016, pp. 770–778. doi: 10.1109/CVPR.2016.90.

[96] G. Scarpa, S. Vitale, and D. Cozzolino, "Target-adaptive CNN-based pansharpening," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 9, pp. 1–15, 2018. doi: 10.1109/TGRS.2018.2817393.

[97] J. B. Bednar and T. L. Watt, "Alpha-trimmed means and their relationship to median filters," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 32, no. 1, pp. 145–153, 1984. doi: 10.1109/TASSP.1984.1164279.

[98] R. H. Yuhas, A. F. H. Goetz, and J. W. Boardman, "Discrimination among semi-arid landscape endmembers using the Spectral Angle Mapper (SAM) algorithm," in *Proc. Summaries 3rd Annu. JPL Airborne Geoscience Workshop*, 1992, pp. 147–149.

[99] A. Garzelli and F. Nencini, "Hypercomplex quality assessment of multi/hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 4, pp. 662–665, 2009. doi: 10.1109/LGRS.2009.2022650.

[100] L. Alparone, B. Aiazzi, S. Baronti, and A. Garzelli, "Multispectral and panchromatic data fusion assessment without reference," *Photogramm. Eng. Remote Sens.*, vol. 74, no. 2, pp. 193–200, 2008. doi: 10.14358/PERS.74.2.193.

[101] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, 2002. doi: 10.1109/97.995823.

[102] A. R. Gillespie, A. B. Kahle, and R. E. Walker, "Color enhancement of highly correlated images: Channel ratio and 'Chromaticity' transformation techniques," *Remote Sens. Environ.*, vol. 22, no. 3, pp. 343–365, 1987. doi: 10.1016/0034-4257(87)90088-5.

[103] J. Liu, "Smoothing filter-based intensity modulation: A spectral preserve image fusion technique for improving spatial details," *Int. J. Remote Sens.*, vol. 21, no. 18, pp. 3461–3472, 2000. doi: 10.1080/014311600750037499.

[104] L. Wald and T. Ranchin, "Liu 'Smoothing filter-based intensity modulation: A spectral preserve image fusion technique for improving spatial details," *Int. J. Remote Sens.*, vol. 23, no. 3, pp. 593–597, 2002. doi: 10.1080/01431160110088772.

[105] M. M. Khan, J. Chanussot, L. Condat, and A. Montanvert, "Indusion: Fusion of multispectral and panchromatic images using the induction scaling technique," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 1, pp. 98–102, 2008. doi: 10.1109/LGRS.2007.909934.

[106] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva, "An MTF-based spectral distortion minimizing model for pan-sharpening of very high resolution multispectral images of urban areas," in *Proc. 2nd GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas*, Berlin, 2003, pp. 90–94. doi: 10.1109/DFUA.2003.1219964.

[107] J. Lee and C. Lee, "Fast and efficient panchromatic sharpening," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 1, pp. 155–163, 2010. doi: 10.1109/TGRS.2009.2028613.

GRS