

Data fusion in data scarce areas using a back-propagation artificial neural network model: a case study of the South China Sea

Zheng WANG^{1,2,3}, Zhihua MAO (✉)^{1,2,3}, Junshi XIA⁴, Peijun DU (✉)^{1,2}, Liangliang SHI^{3,5}, Haiqing HUANG³, Tianyu WANG³, Fang GONG³, Qiankun ZHU³

¹ School of Geographic and Oceanographic Sciences, Nanjing University, Nanjing 210023, China

² Collaborative Innovation Center for the South China Sea Studies, Nanjing University, Nanjing 210023, China

³ States Key Laboratory of Satellite Ocean Environment Dynamics, Second Institute of Oceanography, State Oceanic Administration, Hangzhou 310012, China

⁴ Research Center for Advanced Science and Technology, The University of Tokyo, Tokyo 153-8904, Japan

⁵ Ocean College, Zhejiang University, Hangzhou 310058, China

© Higher Education Press and Springer-Verlag Berlin Heidelberg 2017

Abstract The cloud cover for the South China Sea and its coastal area is relatively large throughout the year, which limits the potential application of optical remote sensing. A HJ-charge-coupled device (HJ-CCD) has the advantages of wide field, high temporal resolution, and short repeat cycle. However, this instrument suffers from its use of only four relatively low-quality bands which can't adequately resolve the features of long wavelengths. The Landsat Enhanced Thematic Mapper-plus (ETM+) provides high-quality data, however, the Scan Line Corrector (SLC) stopped working and caused striping of remote sensed images, which dramatically reduced the coverage of the ETM+ data. In order to combine the advantages of the HJ-CCD and Landsat ETM+ data, we adopted a back-propagation artificial neural network (BP-ANN) to fuse these two data types for this study. The results showed that the fused output data not only have the advantage of data intactness for the HJ-CCD, but also have the advantages of the multi-spectral and high radiometric resolution of the ETM+ data. Moreover, the fused data were analyzed qualitatively, quantitatively and from a practical application point of view. Experimental studies indicated that the fused data have a full spatial distribution, multi-spectral bands, high radiometric resolution, a small difference between the observed and fused output data, and a high correlation between the observed and fused data. The excellent performance in its practical application is a

further demonstration that the fused data are of high quality.

Keywords data fusion, South China Sea, BP-ANN model

1 Introduction

About 60%–80% of the South China Sea and its surrounding area are covered by clouds at any given moment in time, which seriously affects the quality of optical remote sensing data. For this reason, making full use of whatever limited remote sensing data exists in this region is important. Landsat ETM+ data have the advantages of high data quality, and high spectral and spatial resolution. However, since the repeat cycle of the ETM+ sensor is 16 days, Landsat has relatively poor temporal coverage. Making matters worse, the Scan Line Corrector (SLC) on the ETM+ instrument failed on May 31, 2003, which led to a gap in the data record that amounted to nearly 22% missing data over the duration of the Landsat mission (Mohammdy et al., 2014). Therefore, due to the long repeat cycles and missing data, it is difficult to obtain valid data for some temporal scales. The new HJ-CCD observational network (consisting of satellites HJ-1A & HJ-1B), with a repeat cycle of 2 days, has the advantages of a wide field of view and high temporal resolution. However, its low spectral resolution (only four bands) significantly limits the potential applicability of the CCD data in the study area. Therefore, to obtain the best possible coverage and make full use of all available remote sensing data, it is necessary to study how to combine these different

Received September 10, 2016; accepted March 5, 2017

E-mails: mao@sio.org.cn (Zhihua MAO); dupjrs@gmail.com (Peijun DU)

data sources to best leverage their advantages and mitigate any disadvantages of the CCD and ETM+ SLC-Off data. Multi-source remote sensing data fusion can combine the advantages of each data source and avoid or suppress the disadvantages of being incomplete and eliminate errors in single images. The aim of multi-source remote sensing data fusion is to create more accurate, more complete, and more reliable results than the results from the single source image information. Therefore, studies of multi-source remote sensing data fusion should be carried out to solve the existing problems inherent in both HJ-CCD and ETM+ SLC-Off data.

Multi-source remote sensing data fusion (MSRSDF) is a technology that uses a special algorithm to process spatial-temporal redundancy and complementary multi-source remote sensing data to create a fused data set containing advantages of each of the input data sources and minimizing their weaknesses. MSRSDF fuses new composite data, which are more complete, more accurate, and more informative than previous data (Khaleghi et al., 2013). In their example, the fused multi-source remote sensing data had more complementary spectral characteristics than single data sources. With the development of remote sensing techniques there has been a rapid growth in multi-source remote sensing data. The integration of multi-source remote sensing data has enormous potential applicability in both academic and commercial domains (Gigli et al., 2007).

The concept of data fusion was first proposed in the 1970s in the United States. Thanks to its high application potential and academic significance, there have been many studies of multi-source data fusion in recent decades. This methodology was first applied to military purposes using both modeling and simulation (Bossé et al., 2000). Multi-source data fusion technology was also applied to multi-sensor technology in dimensional metrology (Weckenmann et al., 2009), fire detection (Zervas et al., 2011), and intelligent transportation systems (Faouzi et al., 2011).

Developing the theoretical basis of multi-source data fusion in remote sensing science has been an important research direction. Daily et al. (1979) has been credited with performing the first simple multi-source remote sensing image fusion when they composited radar and Landsat imagery for geologic interpretation. With this success, multi-source remote sensing data fusion began to attract more attention in the 1980s. Muskat (1983) applied the method to merge Seasat-A radar images and Landsat MSS images for use in geological interpretation. Following this, Welch and Ehlers (1987) merged SPOT panchromatic and Landsat TM data to create multi-sensor, multi-resolution, multi-spectral, and multi-temporal composite image products. Mohan and Mehta (1988) combined radar and Landsat data using data fusion for application in land use and land cover change studies. Sharma et al. (1990) fused synthetic aperture radar (SAR) data, aerial photographs, and MSS Landsat data for geological appraisal.

Ehlers (1991) merged SPOT HRV, Landsat TM, and Shuttle Imaging Radar (SIR-B) images for the purpose of integrative rectification, enhancement of cartographic, feature extraction and improvement of spatial resolution. Toutin (1995) evaluated the impact of geometric and radiometric processing of data fusion on the resulting composite images using SPOT panchromatic and SAR image data. In the 21st century, research on multi-source remote sensing data fusion is progressing rapidly with the launch of a number of satellites and the desire to connect the new measurements with the earlier ones. Kiema (2002) used data fusion of panchromatic imagery taken by the French SPOT satellite fused with Landsat imagery to derive an algorithm for the automatic extraction of topographic objects from satellite data. Turker and San (2003) merged the pre-earthquake and post-earthquake SPOT HRV data to detect earthquake-induced changes. Amici (2004) fused SAR image based decision level data to develop better flood inundated area detection. Chen et al. (2006) studied multi-source remote sensing data fusion of SPOT and TM images using a wavelet transform. Zhang et al. (2007) used histogram matching and Kriging interpolation methods to de-stripe and data-fuse ETM+ SLC-Off with ETM+ data. Maxwell et al. (2007) used a multi-scale segment model to fill the image gaps and data fusion on the Landsat SLC-on image. Nachouki and Quafafou (2008) developed a new data fusion approach for heterogeneous source data. Hilker et al. (2009) proposed a new data fusion model for high spatial- and temporal-resolution data, such as Landsat ETM+ and EOS-MODIS. Wu et al. (2012) presented a new spatial and temporal data fusion model for fusing a new high spatial-temporal resolution data based on Landsat and MODIS surface reflectance. Chen et al. (2011a) recovered the thermal band of Landsat 7 SLC-off ETM+ images using CBERS as auxiliary data. Following this, Chen et al. (2011b) developed a simple but effective method for filling gaps in Landsat ETM+ SLC-off images based on the neighborhood similar pixel interpolator (NSPI) approach. Based on this approach, Zhu et al. (2012) developed a new geo-statistical data fusion approach using a geo-statistical neighborhood similar pixel interpolator (GNSPI). Zeng et al. (2013) proposed an integrated method to recover missing image pixels with the results assessed using land-cover classification and a Normalized Difference Vegetation Index (NDVI). They found that the recovered data were of adequate accuracy for further remote sensing applications. Nguyen et al. (2014) studied spatial-temporal data fusion in very large remote sensing datasets. Novelli et al. (2016) studied a multi-source remote sensing data fusion algorithm based on the Kalman filter. Liu et al. (2017b) applied statistical analysis to the performance assessment in the image fusion research.

In summary, multi-source remote sensing data fusion can be classified as multi-band, multi-temporal, multi-sensor, or multi-resolution image fusion. From the

information representation point-of-view, it can also be classified by the level on which it is applied, ranging from the pixel level, feature level, to the decision level. Pixel level fusion has several advantages; among these are low information loss, high fusion accuracy, and small computational requirements. However, it also has some disadvantages, like poor real-time performance, poor fault tolerance, and poor ability of anti-interference. Decision level fusion benefits from good real-time performance, strong fault tolerance, strong anti-interference ability, and high level of integration; however it suffers from high information loss, low precision, and high computational cost. Finally, while feature level fusion results can reflect most of the information and effectively reduces the computational effort, it suffers from an information loss problem and it doesn't provide the detailed information that the fused image needs to show. An artificial neural network (ANN) is the simulation of a simplified biological neural network. With the development of the ANN theory, a new method is available for multi-source remote sensing data fusion. The benefits of ANN are that it possesses inherent parallelism, self-organization, self-learning, and high fault tolerance to the input data. ANN can perform data fusion through a self-learning process rather than through some particular set of parameter values, which provides it with a good fault-tolerance capability. Further, the nonlinear characteristics of ANNs can better reflect the complex relationship that can exist between multi source remote sensing datasets. Therefore, ANN has recently seen increasing application to remote sensing studies. Benediktsson et al. (1989) applied a neural network approach in the classification of multisource remote sensing and geographic data. A feed-forward neural network, which can be applied to fuzzy pattern classification problems, was proposed by Karayiannis and Purushothaman (1994). Chen et al. (1997) successfully applied the feed-forward neural network to the classification of Landsat TM image. Tedesco et al. (2004) used ANN technology to effectively retrieve snow depth information from the Special Sensor Microwave Imager (SSM/I) data. Farifteh et al. (2007) used ANN to retrieve the reflectance spectrum of salt-affected soil. Maeda et al. (2009) applied ANN to monitor forest fire in the Brazilian Amazon forest using MODIS data. Al-Sbou (2012) applied ANN to de-noise an image. Fan et al. (2014) constructed a remote sensing image classification algorithm using a novel evolutionary neural network to monitor the GaoFen (GF-1) satellite marine oil spill. Mehta et al. (2015) used a multi-layer feed-forward neural network in supervised classification of medical imagery. Liu et al. (2017a) used a deep convolutional neural network model to solve the activity level measurement and fusion rule issue in image fusion.

With its high parallelism, nonlinear approximation capability, good fault tolerance, associative memory function, and very strong self-adaptive, self-learning capability, ANN is now being widely used in remote

sensing. However, ANN has rarely been applied to the fusion of multi-source remote sensing data.

The study area for this paper is the coastal area of the South China Sea, which is a data scarce region due to persistent cloud cover. In order to improve the utility of the limited existing data and to demonstrate the unique advantage of ANNs for data fusion, multi-source remote sensing data fusion was performed using HJ-CCD and Landsat ETM+ data. Our aim of incorporating the ANN model to the data fusion approach was to combine the advantages of the Landsat ETM+ SLC-Off and HJ-CCD data while avoiding their disadvantages. Our ultimate goal is to develop a methodology that can generate new high-quality fused data with high spectral and temporal resolution that maintains data integrity. If successful, this will allow for the fullest use of all existing data sources. This paper is organized as follows: The study area, datasets and method are described in Section 2; Quantitative and qualitative results, actual application and discussion are presented in Section 3; Conclusions are given in Section 4.

2 Data and method

2.1 Study area, sample area, validation area and validation sites

The study area (Fig. 1) is in the northern edge of the South China Sea between 109.10°–111.13°E, 20.36°–22.43°N. The annual mean temperature and precipitation in the area is 23°C, and between 1393–1758 mm, respectively. During the summer season, this area is influenced primarily by the southeast monsoon from the western Pacific Ocean and the southwest monsoon from the Indian Ocean. Because the study area has a tropical monsoonal climate, it is rainy and cloudy most of the time, which significantly limits the potential application of the optical remote sensing. In addition, missing data and the low temporal resolution of the satellite data further erode its utility in this area.

Data from the sample area is used as the source of the ANN model training data. The location and region covered by the sample area is shown in Fig. 1(b). The training samples selected from this area only include positive values. A wide variety of terrestrial objects and their dynamic ranges should be contained within this area. A total of 44,458 pairs of training data points taken from the sample area (excluding the validation area) from the input and output datasets of the data fusion model were used for training the data fusion model.

The validation area was used to verify the data fusion results resulting from application of the model. The validation area should contain all types of ground objects that are typically present in the study area; such as various water body types (rivers, lakes, ponds, marshlands, etc.), mountains, plains, farmlands, man-made objects, forests. With the number of high-resolution satellites being

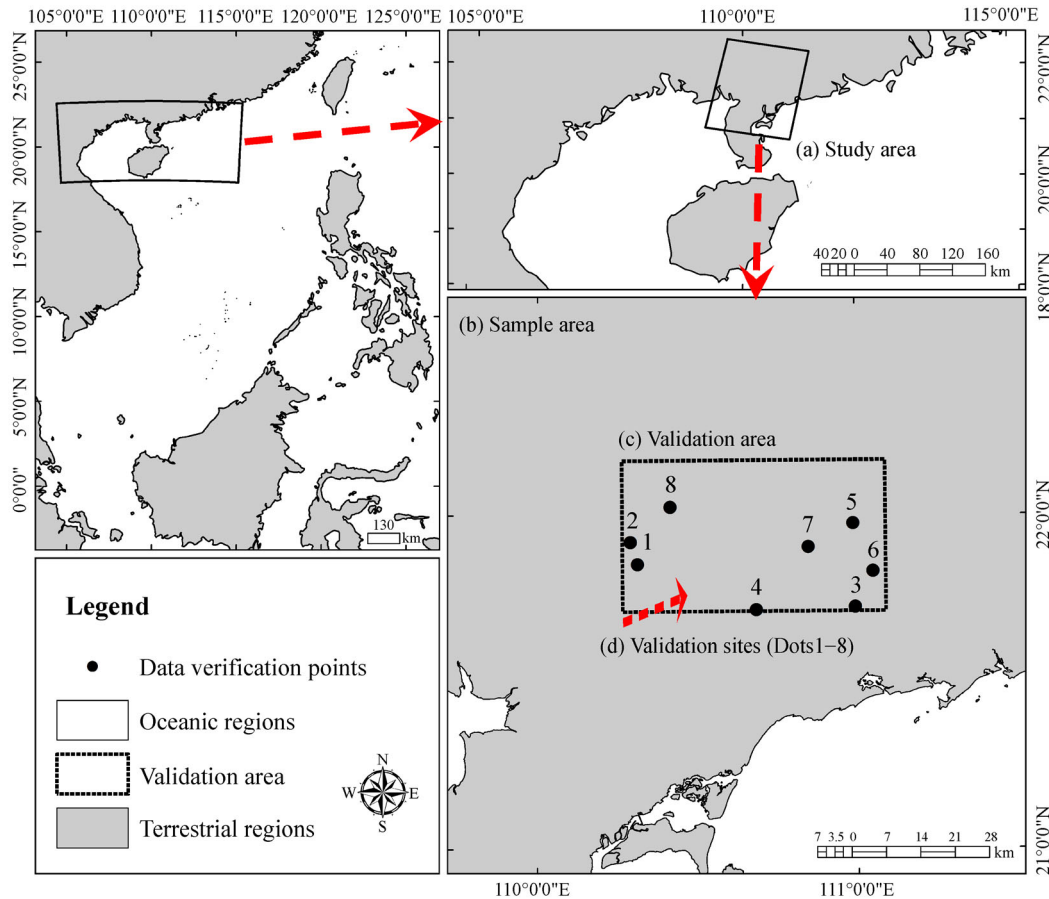


Fig. 1 Study area (a), sample area (b), validation area (c), and validation points (d) used in this study.

launched rapidly increasing, a quickly expanding database of imagery has been cataloged into Google Earth (GE, Hu et al., 2013). GE is an open source database, which provides clear views of a wide variety of ground objects (buildings, roads, etc). Hence, GE can be used for preparing the region of interest or for validating datasets and related applications (Jacobson et al., 2015). As a result, the specific validation area was chosen based on high-resolution images obtained from the GE software (shown in Fig. 1(c)). Figure 2 presents several true-color, high-resolution, composite images acquired from the GE software. It shows that there are a wide variety of ground objects within the validation area. The eight sites shown in Fig. 2 were used as the validation sites.

The eight validation sites were used to quantitatively verify the simulation results of the data fusion model. Specifically, the reflectance values acquired from ETM+ and fused data of eight sites were used to validate the differences in the absolute value of reflectance between the fused data and the observed ETM+ SLC-Off data. The eight validation sites were divided into four groups according to terrain types: water body, forest, farmland, and man-made objects. We chose two sites for each terrain

type and the details of these eight sites are shown in Fig. 2.

2.2 Data and preprocessing

Landsat 7 with its ETM+ imager was successfully launched on April 15, 1999. The main parameters for the Landsat ETM+ imager are shown in Table 1 (Zeng et al., 2013). The two Chinese environmental satellites, HJ-1A/B, were both successfully launched (one rocket) on September 6, 2008. The HJ-A/B satellites each carry four charge-coupled device (CCD) sensors with uniform spatial resolution and spectral range. The main parameters of the HJ-1A/1B CCD satellites are shown in Table 1 (Liu et al., 2011). For the data fusing effort, data from HJ1A-CCD2 and Landsat ETM+ SLC-off from October 2, 2013 were acquired. Imaging times of the two datasets were 3:00 am for the ETM+ and 2:15 am for the CCD.

After radiometric calibration, atmospheric and geometric corrections, Landsat ETM+ SLC-Off and HJ-1A CCD2 data were converted to reflectance. The reflectance data were used in ArcMap, then a new vector point file was created in Arc Catalog. Finally, the vector point file was

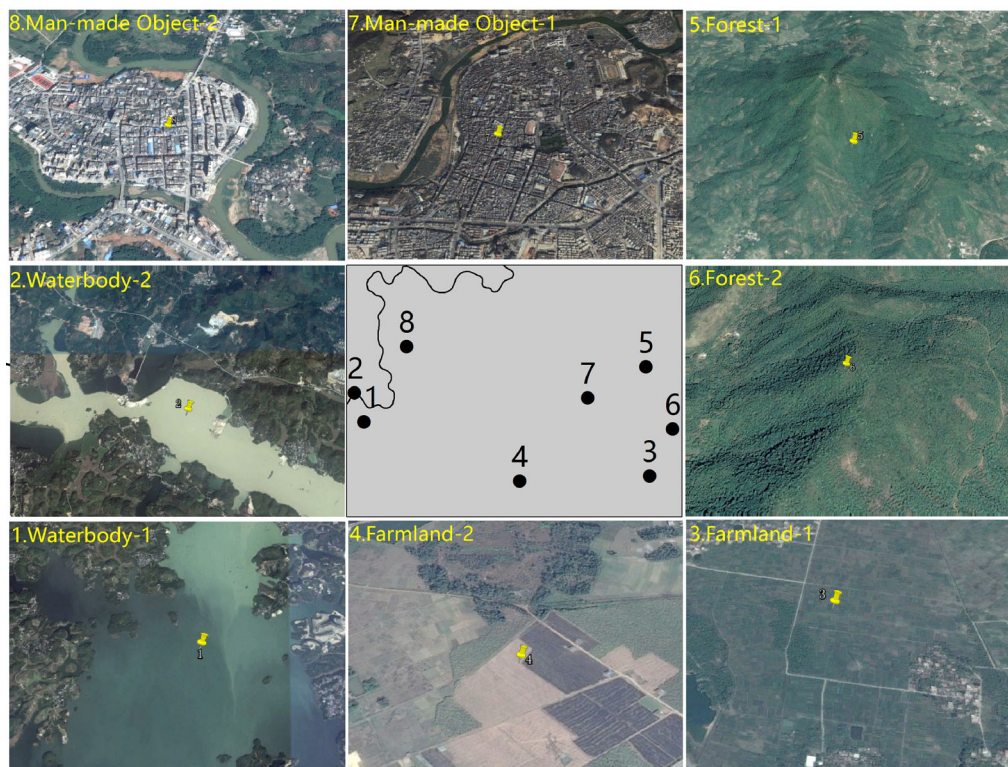


Fig. 2 Details of validation sites. Sites 1 and 2 were different water bodies, sites 3 and 4 were farmland sites on a plain, sites 5 and 6 are forested mountainous areas and sites 7 and 8 were urban areas. All of the 8 validation sites were chosen from Google Earth.

Table 1 Specifications of Landsat ETM+ and HJ-1A CCD satellite missions

Parameters	Landsat ETM+	HJ-1A CCD
Band1/ μm	0.45–0.52	0.43–0.52
Band2/ μm	0.52–0.60	0.52–0.60
Band3/ μm	0.63–0.69	0.63–0.69
Band4/ μm	0.77–0.90	0.76–0.90
Band5/ μm	1.55–1.75	N/A
Band7/ μm	2.09–2.35	N/A
Spatial resolution/m	30	30
Swath width/km	185	700
Imaging time (GTM)	3:00	2:15
Imaging date	2 Oct, 10, 2013	2 Oct, 10, 2013
Cycle/day	16	2

then used to collect samples from the reflectance data beyond the validation area of Landsat ETM+ SLC-Off and HJ-1A CCD2.

2.3 Back-propagation ANN model

A BP-ANN model is a multi-layer, feed-forward, neural network using an algorithm of error reverse transfer

training. This model has one or more hidden layers between input and output layers, and makes the output result approximate the ideally simulated value by adjusting the weight of the neural network (Tedesco et al., 2004; Farifteh et al., 2007). Reference showed that the three-layer ANN model (which has only one hidden layer) can approximate any function. But, adding more hidden layers to the ANN model will lead to a dramatic decrease in computational speed. But, since it makes the neural network more complex it may improve the computational accuracy to some degree. However, if the number of neurons is increased, the accuracy and computing speed will be dramatically improved without the use of additional layers. Consequently, the model can be observed and easily adjusted at the same time (Maeda et al., 2009). Therefore, a three layer BP-ANN model, with only one hidden layer between the input and output layers, is adopted for this study (shown in Fig. 3). The left most column of Fig. 3 (4 nodes enclosed in a dashed rectangle) is the input layer, which corresponds to external stimuli, is the source of stimulation to the neural net. The center column (28 nodes surrounded by the dashed rectangle) is the hidden layer, which connects the neurons between input layer and output layers and allows transfers of information, is equivalent to a human brain. The right most column (6 nodes enclosed in the dashed rectangle) is the output layer, its structure indicates that the neural net reacts to the external

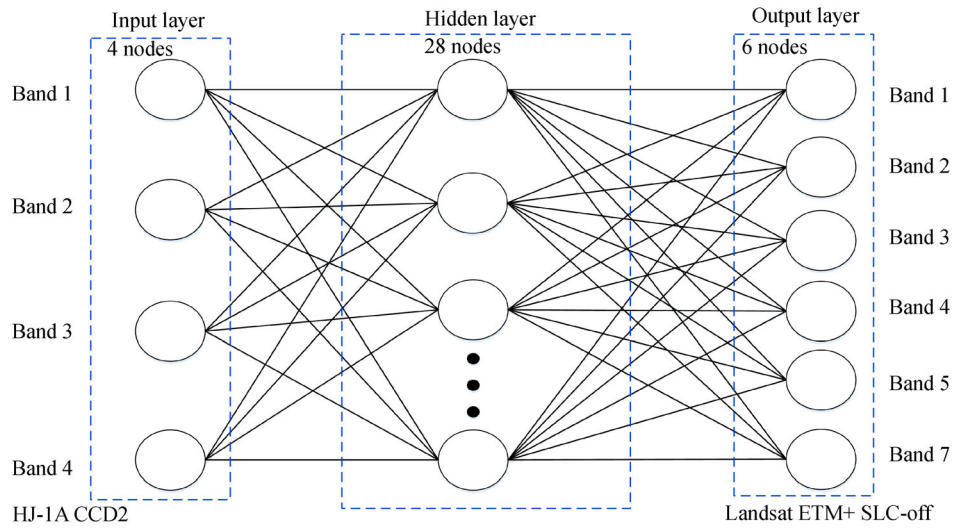


Fig. 3 A basic neural network model structure for data fusion between HJ-CCD and Landsat ETM+ satellite.

environment through multiple levels.

Specifically, in this model the stimulus will be transmitted from the input layer through the hidden layer to the output layer, using the strength of the relationship between neurons (weights) and transfer rules (activation functions). The model's learning process is finished by feeding the errors back into the neural network and modifying the weights of the connection between the neurons (as shown in Fig. 3). The mathematical expression of its input and output is:

$$x_j = y_1 * w_{j1} + \dots + y_i * w_{ji} + \dots + y_n * w_{jn} = \sum_{i=1}^n y_i * w_{ji}, \quad (1)$$

where x_j is the cumulative amount of stimulation and the corresponding weights transferred from other neurons, y_i is the amount of stimulus transmitted from neuron i and w_{ji} is the linking weight of the stimulus j to neuron i .

When the stimulus accumulation of x_j is finished, the stimulus is transmitted to the surrounding neurons by the neurons that finished the accumulation, y_i . The equation is

$$y_j = f(x_j). \quad (2)$$

The stimulus, y_j , is transmitted to the external neurons by processing according to the accumulation results of x_j . This process is represented by the function f in Eq. (2). A sigmoid or S type function is chosen as the active function for the neurons, because this type of function is continuous and smooth:

$$y_j = f(x_j) = \frac{1}{1 + e^{-x_j}}. \quad (3)$$

An iterative procedure in the BP-ANN model is performed using Eqs. (1) and (3) where y_j is the stimulus for each neuron and x_j is generated after the weight accumulation using weights w_{ji} . The stimulus y_j is then generated by the active function, Eq. (3), and transmitted to the connected neurons of the next layers sequentially until the final results are generated.

In this study, the input nodes are the four bands of HJ-1A CCD2 reflectance data and the output nodes are the six bands of Landsat ETM+ reflectance data. Pairs of data (a total of 44,458 pairs) were taken from the Landsat ETM+ SLC-Off and HJ-1A CCD2 reflectance data as the training samples for the neural network. The training samples should be selected from a cloud-free area without gaps or zero values and a variety of object types covering a wide dynamic range should be included. The S type Hyperbolic Tangent function is used for the transmission function in the ANN model with a Gradient Descent Training Algorithm employed for the learning function. Due to its fast convergence and time-saving capability, a Scaled Conjugate Gradient Method was used as the training function to train the ANN model. We selected 60%, 20% and 20% of the samples as the training, validation and prediction data, respectively. The overall fusion framework is summarized in flowchart (Fig. 4).

Determining the node number of the hidden layer is a vital part of neural network design. It is hard to meet both the requirements of sample dataset learning and reflect the regularities of the sample dataset if the node number of the hidden layer is too small. By contrast, too many nodes in the hidden layer will lead to a longer training time and irregular noise will be retained by the neural network, which will cause neural network over fitting. However, there is no universal analytic formula to determine the

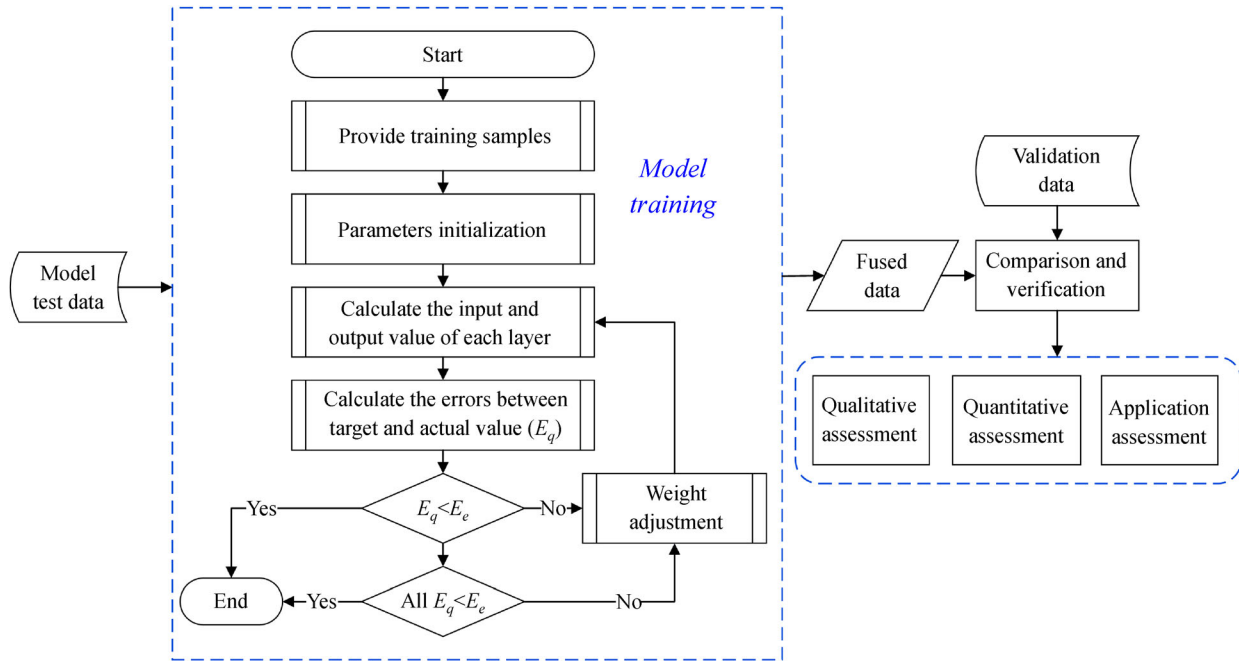


Fig. 4 Flow chart of the overall fusion framework. E_e is the expected training errors. Parameters initialization include: maximum training times, learning accuracy, the number of hidden nodes, initial weights, a threshold value, initial learning rate, etc.

number of nodes to use in the hidden layer. Commonly used methods to estimate the number of nodes needed are based on existing empirical models. A multiple trial method was used to determine the number of nodes to use in the hidden layer in this paper. The empirical formula used to determine the number of initial nodes is $l = \sqrt{m + n} + a$, where m and n are the number of nodes in the input and output layers, respectively (we use $m = 4$ and $n = 6$ in this paper); a is constant between 1 and 10. The initial number of nodes in the hidden layer is 9, and the error is 0.035. The error decreased significantly as the number of nodes was increased. When there were 26 nodes in the hidden layer, the output error of the network was reduced to about 0.001, which meets the error requirements for an acceptable neural network. To include a safety margin, the number of nodes in the hidden layer was chosen to be 28. This number is acceptable because there is a large quantity of remote sensing data. Through continuous experiment and adjustment of the ANN model, the results show that the neural network has an excellent performance with 28 nodes. The input nodes are the four input bands of the HJ-1A CCD2 data, as shown in Fig. 3. The values from the input layer, which performs the summation and activation functions, were received by the 28 nodes of the hidden layer. The output values and information about the hidden layer gathered by computation were then used to compare with the target values of the network's output layer. The weights of the connection were modified if the error between the output and target values was larger than expected. The weight of the

connection was continually modified until the error between the output and target value was equal to or smaller than the required value. The result from the output layer contains the information we are concerned about, referred to the bio-optical properties as shown in Eq. (4):

$$Output = f \left(\sum_{i=1}^n w_{ji} y_i + b_j \right), \quad (4)$$

where w_{ji} is the connection weight of the neurons; y_i is the amount of input stimulus transmitted from neuron i , b_j is the bias or threshold value of the output layer; j is the number of nodes of the hidden layer; $f(*)$ is the activation function.

3 Results and discussion

3.1 Spatial distribution of experimental results

3.1.1 Comparison of simulated ETM+ and HJ-CCD data

Figure 5 shows the observed HJ-CCD and the synthetic ETM+ data simulated by the ANN model. Generally speaking, the false color composite image contains abundant information. However, the false color composite image is a mixture of three channels and some useful information may be lost, disappear, or be masked by other bands in the color composite display. Hence, single gray scale original and simulated bands for these data are

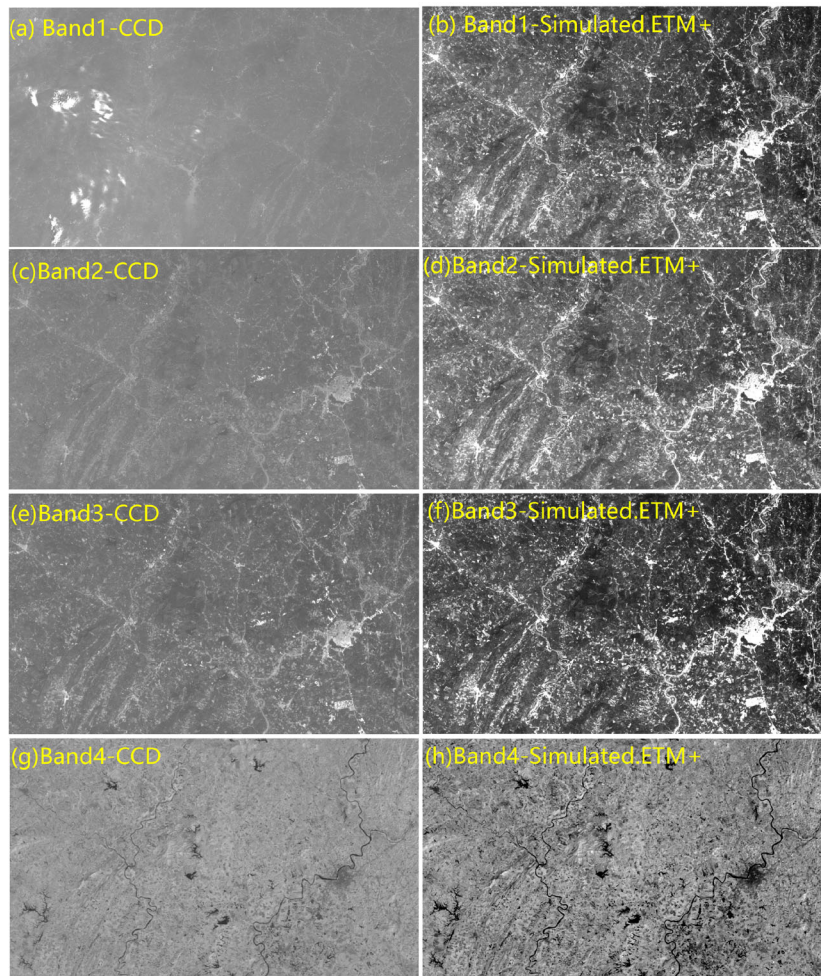


Fig. 5 Comparison of bands 1-4 in the observed HJ-CCD data (left column, panels (a), (c), (e) and (g)) and the synthetic ETM+ data simulated by the ANN model (right column, panels (b), (d), (f), (h)).

displayed in Fig. 5 and used for a comparative study between observed CCD and simulated ETM+ data. We can see that the simulated ETM+ data have more definition in the water body images and are more prominent in urban areas; they also have a larger dynamic range over vegetation, and show more clarity and sharpness overall than the observed CCD data.

3.1.2 Comparison of ANN simulated ETM+ and observed ETM+ data

The left column panels of Fig. 6 are the original single gray scale bands (band 1-5, and band 7) of the observed ETM+ data. The right column panels of Fig. 6 are the single gray scale band images of the synthetic ETM+ results from the ANN model simulation that correspond to the ETM+ data on the left. Figure 6 covers the blue band (band 1), the green band (band 2), the red band (band 3), the near infrared (NIR, band 4), the first shortwave infrared

(SWIR1, band 5), and the second shortwave infrared (SWIR 2, band 7) of the ETM+ data. The NIR, red, and green bands are the bands that contain the absorption and reflection peaks which will highlight the vegetation information in the image. The first shortwave infrared (SWIR1), the NIR, and the red bands typically highlight the differences of various terrain objects in the image. The second SWIR 2 NIR and the blue band, which have good atmospheric transmission capabilities and rarely affected by aerosol or water vapor, contain geological and terrain surface information. With these considerations, compared to the left column panels in Fig. 6, the images in the right column panels show that the image fusion using the ANN model was very successful in that it has totally removed the striped artifact, it fully expresses the spatial distribution information, it contains more land surface information, and it has clarified and sharpened the image. These image data fusion results demonstrate that the application of the ANN model has somewhat improved the data potential.

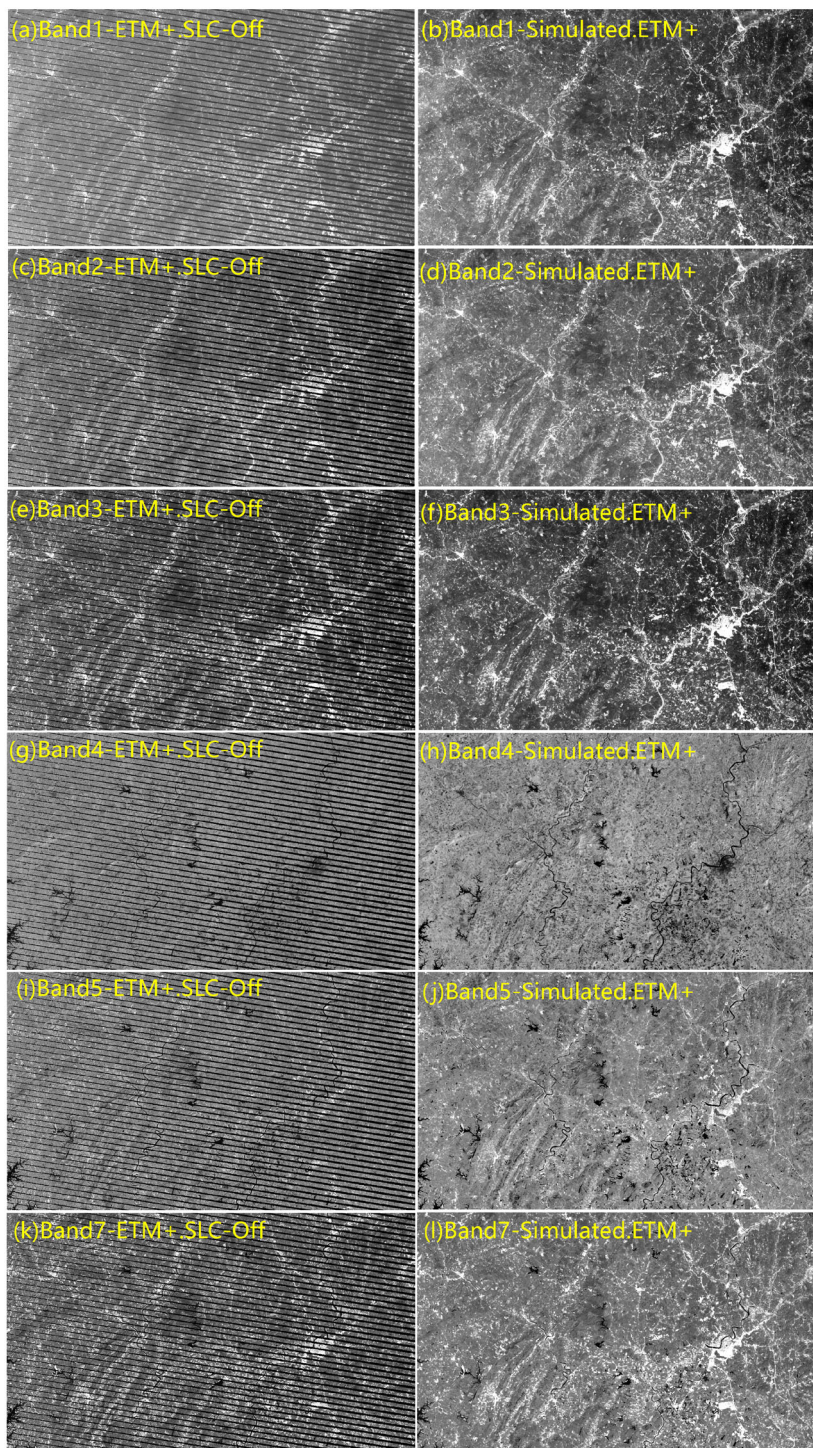


Fig. 6 Comparison between the ANN simulated ETM+ (left column, panels (a), (c), (e), (g), (i) and (k)) and the observed ETM+ data (right column, panels (b), (d), (f), (h), (j) and (l)).

3.2 Absolute accuracy of experimental results

3.2.1 Absolute accuracy of different ground objects

In Fig. 7, we present a comparison of the observed and the ANN simulated reflectance results for the 6 bands at the

eight sample sites in the validation area (Fig. 1 and Fig. 2). As mentioned above, these eight sites were chosen from the high-resolution GE data to represent regions of different terrain object type. The red curve is the reflectance for the ETM+ SLC-Off and the black curve is reflectance for the fused BP-ANN output data. The top row

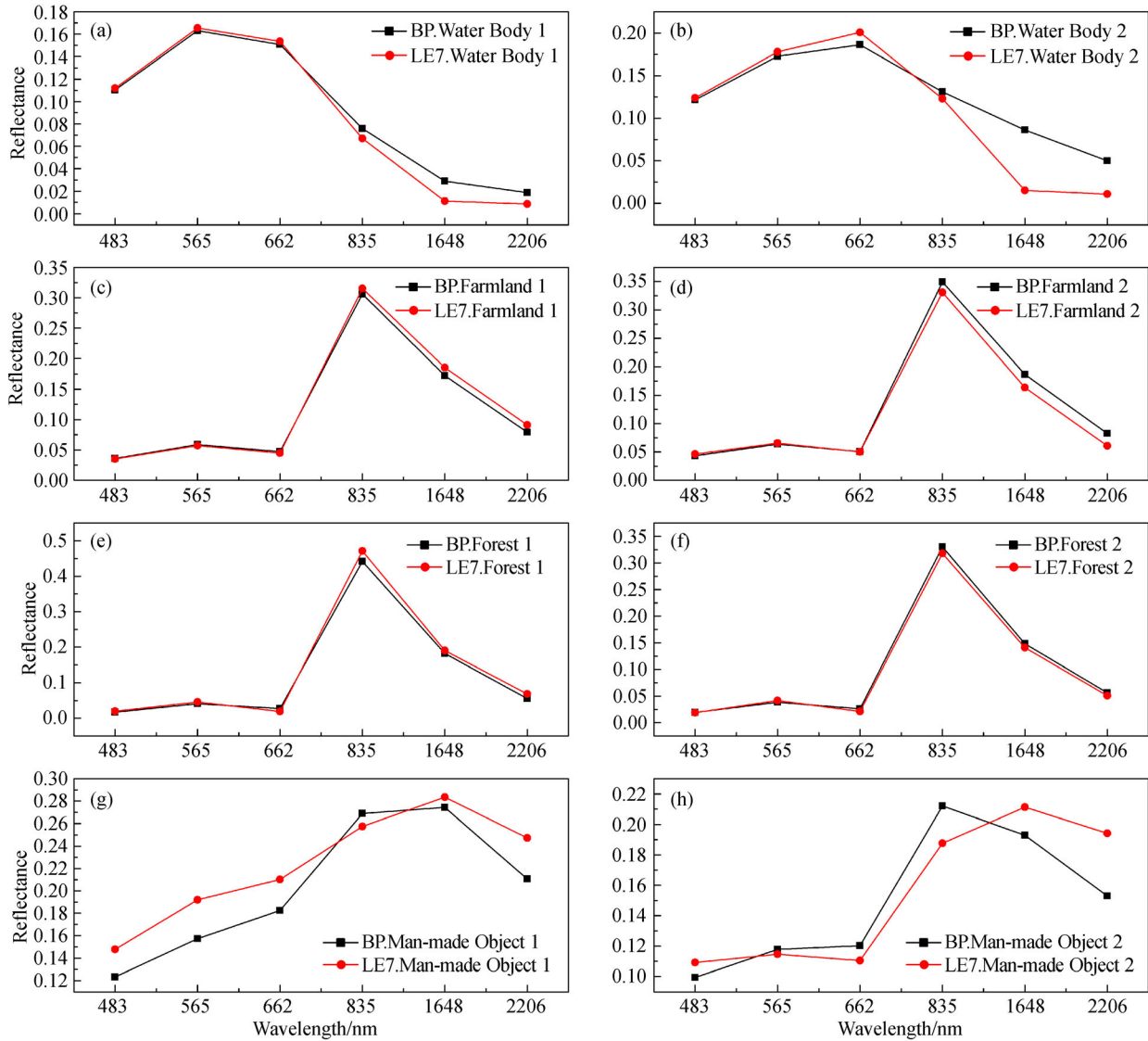


Fig. 7 The difference in absolute value between ANN simulated ETM+ and observed ETM+ data for different terrain objects. Panels (a) and (b) were different water bodies, panels (c) and (d) were farmland sites on a plain, panels (e) and (f) are forested mountainous areas and panels (g) and (h) were urban areas.

of Fig. 7 (panels (a) and (b)) are the water body validation sites, the second row (panels (c) and (d)) are the farmland validation sites, the third row (panels (e) and (f)) are the forest validation sites, and the bottom row (panels (g) and (h)) are the urban sites. The water bodies (Figs. 7(a) and 7 (b)) show very good agreement between the fused and observed reflectance data in the short wavelength bands (483 nm, 565 nm, 662 nm, and 835 nm) with somewhat larger differences in the two longer wavelength bands (1648 nm and 2206 nm). At the farmland sites (Figs. 7(c) and 7(d)), the fused and observed reflectance values are nearly identical in the short wavelength bands (483 nm, 565 nm, 662 nm, and 835 nm), while small differences exist at longer wavelengths (1648 nm and 2206 nm bands) between the fused and observed data. At the forest sites

(Figs. 7(e) and 7(f)), the reflectance values for the fused and observed ETM+ data are nearly the same in all six bands. While at the water body sites, there are small differences in the visible spectral (VIS: bands 1, 2, and 3) and NIR bands between the fused and observed reflectance data in inland waters, a larger difference occurs in the SWIR band BP-ANN. Larger differences between the fused and observed ETM+ reflectance values can be seen for the urban sites (Figs. 7(g) and 7(h)) in all bands but they are particularly prominent in the VIS, NIR, and SWIR bands. Small differences can be seen in the SWIR band from farmland sites with even smaller differences present in the VIS and NIR bands. Forest validation sites show the highest consistency in the VIS, NIR, and SWIR bands between the reflectance of fused and observed ETM+.

Therefore, we can conclude that overall the differences in reflectance between the fused and observed ETM+ data are relatively small, except for urban sites. However, the large differences for urban sites are not of too much concern because these sites represent only a small part of the study area. Overall, with such good performance of the ANN model, fused ETM+ data can be reliably used for further study.

3.2.2 Correlation between the ANN model simulated results and observed ETM+ data

To further assess the accuracy of the fused model outputs, a small area containing water body, forest, farmland, and

urban objects was chosen randomly from within the validation area of Fig. 1, which covers about 25,4284 (=421*604) pixels. We use the correlation coefficient between the observed and simulated ETM+ pixels to assess the accuracy of the data fusion. In this analysis, any zero or abnormal values have been removed from the data.

Figure 8 shows the correlation coefficients between the observed and simulated ETM+ data for the six observed wavelength bands. The correlation coefficients for bands 1–5 and band 7 are 0.813, 0.8, 0.81, 0.9657, 0.92, and 0.85, respectively. The majority of the pixel reflectance values in each band fall between: 0.025–0.100 for band 1; 0.050–0.150 for bands 2 and 3; 0.05 and 0.40 for band 4; 0.025 and 0.300 for band 5 and 0.025–0.250 for band 7.

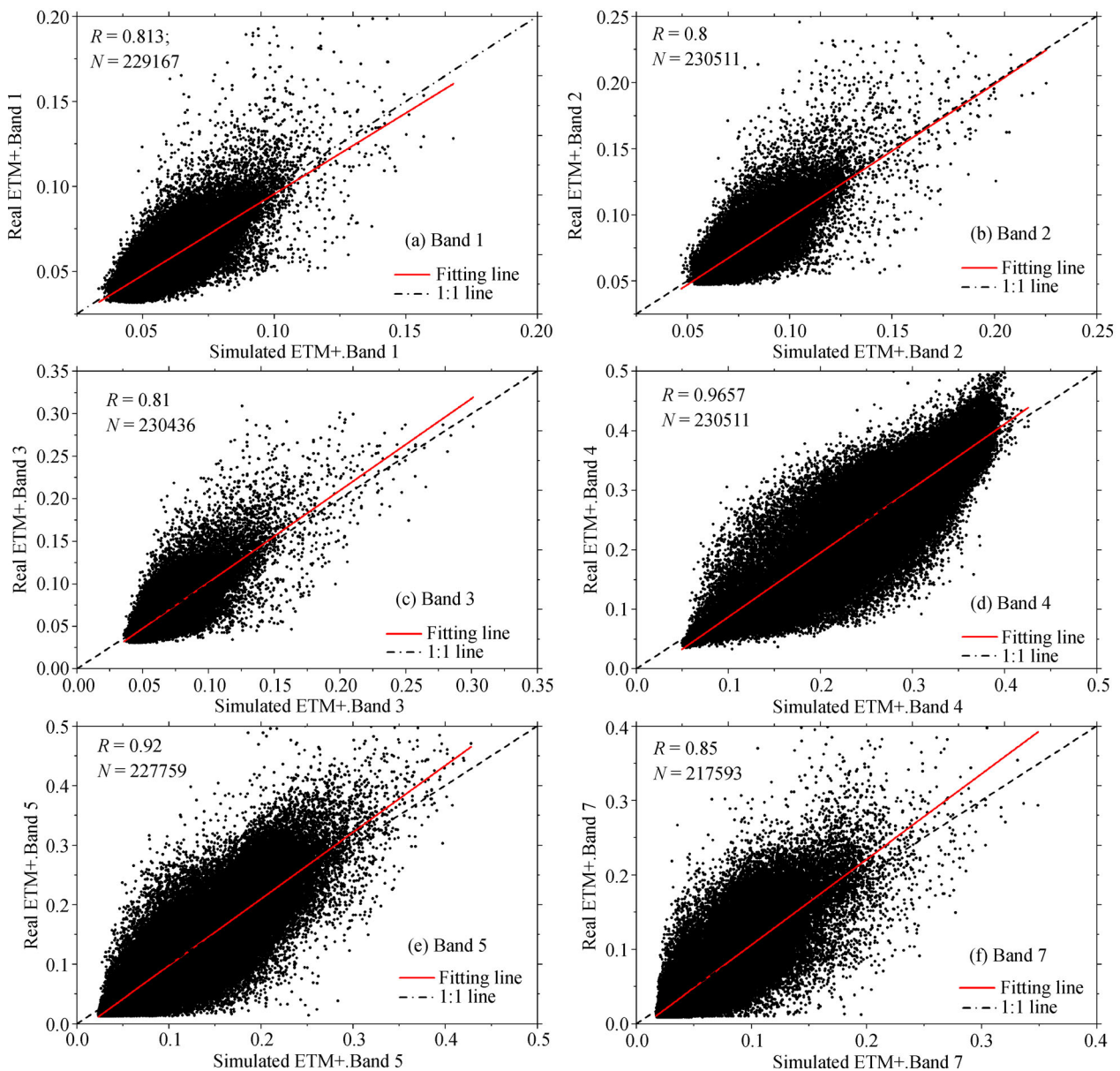


Fig. 8 Scatter plots of observed and simulated ETM+ data. Correlation coefficients in each band are listed at the top left along with the sample size.

There are differences in all bands between the individual observed ETM+ and simulated reflectance data, especially in band 4 and band 5. There are not many discrete data values in band 4 and 5 (Figs. 8(d) and 8(e)), whose reflectance values are higher than those of bands 1–3. These randomly chosen sites mainly contain water body, vegetation and urban areas. As shown in Figs. 7 and 8, the differences in reflectance in all bands and over nearly all terrain object types are small except for those over urban areas. The infrared bands (Figs. 8(d)–8(f)) have higher reflectance and correlation coefficients than do the visible bands (Figs. 8(a)–8(c)). Because the infrared bands have longer wavelengths than the visible bands they are much less affected by the atmosphere. Consequently, the infrared bands (Figs. 8(d)–8(f)) have higher reflectance, less discrete data, and higher correlation coefficients.

3.3 Comparison of fusion results between the BP-ANN model and other similar methods

3.3.1 Qualitative comparison of the BP-ANN model results with other similar methods

A false color composite image of the SWIR, NIR, and red bands from Landsat ETM+ SLC-Off and the simulated BP-ANN model data is shown in Fig. 9. These kinds of composite images are often used to identify water bodies and residential lands. Figure 9(b) contains the data fusion

results using a nearest neighbor interpolation (NNI) method applied to a single image. The data fusion result for NNI is not good enough to represent some key terrestrial features, such as rivers and roads, which are indistinct and intermittent (as shown in the dashed elliptical region). A local adaptive regression model with multiple images based on global linear histogram matching was used to obtain Fig. 9(c), while multiple image fixed window regression analysis model based on local linear histogram matching was used to obtain Fig. 9(d) (Suliman, 2016).

As shown in Fig. 9(c), the features of linear objects such as roads and rivers are presented continuously and with high fidelity by the Global Linear Histogram Match (GLHM) method. The GLHM data fusion results not only show a great improvement compared to the raw ETM+ SLC-Off data, but also show a significant improvement over the data fusion results based on the nearest interpolation method. However, as shown for the urban area enclosed by the dashed elliptical area, the difference between the striped region of ETM+ SLC-Off and its surrounding pixels is relatively large, showing that the GLHM fusion results need to be improved. The fusion results using Local Linear Histogram Match (LLHM, Fig. 9(d)) does provide further improvement relative to those from the GLHM in Fig. 9(c), as the characteristics of rivers and roads are fully represented in the LLHM fusion results (Fig. 9(d)). In addition, the difference between the

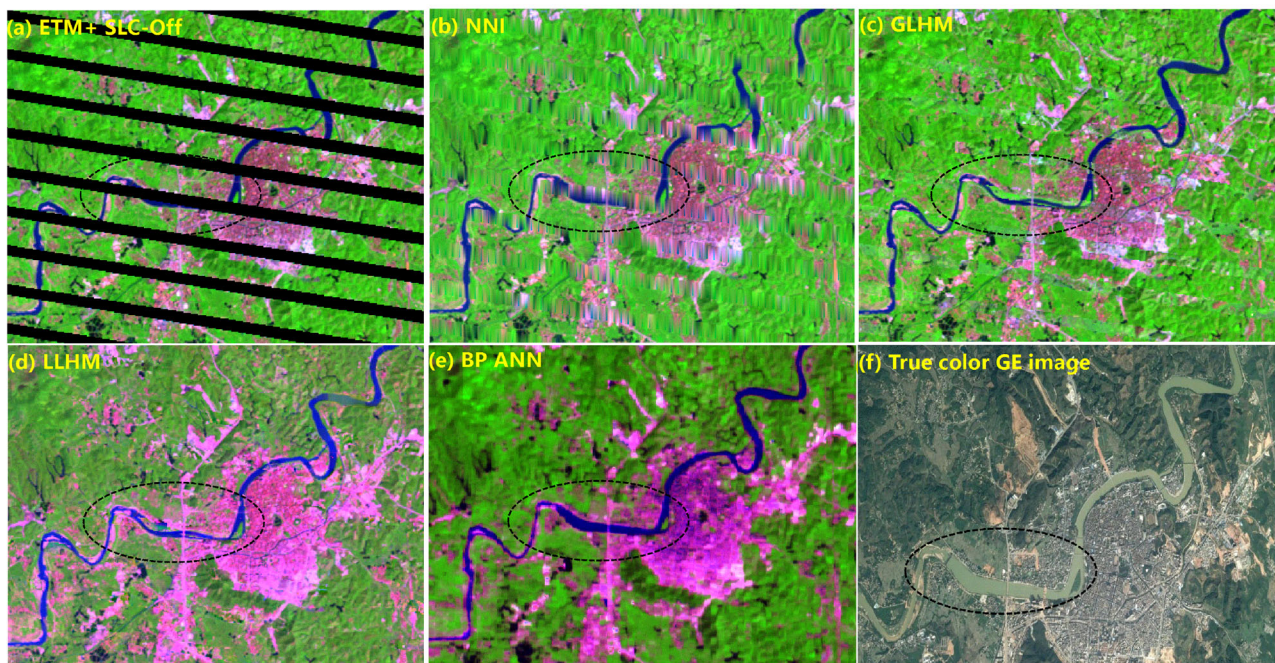


Fig. 9 Qualitative comparison between results of BP-ANN model and other similar methods. (a) ETM+ SLC-Off image; (b) Nearest neighbor interpolation (NNI); (c) Global linear histogram match (GLHM); (d) Local linear histogram match (LLHM); (e) Back-propagation artificial neural network (BP-ANN); and (f) The high-resolution remote sensing image acquired from the Google Earth software of the same region. The acquisition day of the Google Earth image was on January 15, 2014, which was reasonably close to the ETM+ data acquisition time.

gap interpolation region and the peripheral pixel has been eliminated. Furthermore, there is no apparent difference between the gap region of the ETM+ SLC-Off data and its surrounding pixels in the low dynamic regions, such as forest, mountain, farmland, and even the urban areas. However, both of the Linear Histogram Match (GLHM and LHHM) methods (Figs. 9(c) and 9(d)) are affected by historical auxiliary images, which should not be ignored. They both show a certain deflection in the river bank and river region (elliptical dashed box) based on a comparison to the high-resolution data acquired from GE (Fig. 9(f)). By comparison, all the deflection and differences present in similar methods (Figs. 9(b)–9(d)) are totally removed by the BP-ANN approach (Fig. 9(e)). As is shown in Fig. 9(e), there are no obvious interference factors, and the characteristics of linear terrestrial features such as rivers and roads are presented in detail in the BP-ANN fusion results. Consequently, the BP-ANN methodology is the data fusion approach used in this paper.

3.3.2 Quantitative comparison between results of BP-ANN model and other similar methods

For further evaluation and verification of the BP-ANN data fusion method, the model simulated NIR band was compared with the observed NIR band from the ETM+ data. In addition, a Pearson's correlation analysis was carried out to evaluate the relationships between the observed data and the simulated fused data. Furthermore, the results of BP-ANN data fusion model are also compared with those from other similar models.

The results in the NIR band of the BP-ANN data fusion model are compared with the results of other similar data fusion models; such as the data combination model (Busetto et al., 2008), spatial and temporal adaptive reflectance fusion model (STARFM) (Hilker et al., 2009) and spatial and temporal data fusion model (STDFM) (Wu et al., 2012). As shown in Table 2, the NIR band reflectance results from the simulated BP-ANN model are more strongly correlated to the observed NIR reflectance than are the other data fusion models. The HJ-CCD data have the advantage of relatively complete distribution but suffer from low spectral and radiation resolution. In contrast, the ETM+ SLC-Off has the advantages of high spectral and radiation resolution, but the data distribution was incomplete after the scan line

corrector of ETM+ sensor stop working. The ANN model has parallel, self-organizing, self-learning, and fault-tolerant features and in particular its nonlinear properties can better reflect the complex relationships within multi-source remote sensing data. Hence, the ANN model fully combines the advantages of HJ-CCD and Landsat ETM+ data while avoiding their disadvantages. That is the reason the BP-ANN data fusion model gives better fusion results than the other data fusion models.

3.4 Actual application of ANN model fused data

The practical application of the fused data is a necessary and important step to further validate the results of the fused data model. Because urban areas cover such a small part of the validation area, data from those areas can be neglected. Consequently, only forest, farmland, and water body areas are discussed in this section. Because the forest and farmland areas are mostly covered by vegetation, vegetation indices are used for assessing the simulated data. Specifically, enhanced vegetation index (EVI) and modified normalized difference water index (MNDWI) are used to assess the vegetation and water body areas. EVI is predominately captured by the red, infrared and blue bands, while MNDWI is predominately captured by the green and SWIR bands. It is worth noting that these two indices are sensitive to signals in almost every band of the fused data, so they are good proxies for assessing model performance.

3.4.1 Assessment of the BP-ANN model for vegetated regions.

As shown in Sections 3.1 and 3.2, the differences in absolute values of the reflectance between the fused and observed ETM+ data are relatively small. However, for further assessing the quantitative applicability of the fused data for vegetated regions, the EVI of the fused images is compared with those from the observed data.

The absolute differences in the reflectance between the fused and observed ETM+ data for vegetation-covered farmland and forest regions was minimal (Fig. 7). Because most of the validation areas were vegetation-covered, the EVI, calculated from the reflectance in the blue, red, and NIR bands (Eq. (5)), was used for a further assessment of the data fusion accuracy (Sims et al., 2008).

Table 2 Comparison the reflectance with other multi-source remote sensing data fusion methods in the NIR band

Methods	Sensor1	Sensor2	Coefficient of correlation, <i>R</i>
Data combination	MODIS	Landsat-TM	$0.78 \ll r \ll 0.89$
STARFM	MODIS	Landsat-TM	$0.85 \ll r \ll 0.91$
STDFM	MODIS	Landsat-ETM+	$0.90 \ll r \ll 0.94$
BP-ANN	HJ-CCD	Landsat-ETM+	$r = 0.9657$

$$EVI = \frac{G * (\rho_{NIR} - \rho_R)}{\rho_{NIR} + C_1 * \rho_R - C_2 * \rho_B + L}, \quad (5)$$

where ρ_{NIR} is the reflectance in the NIR band, ρ_R is the reflectance in the red band, ρ_B is the reflectance in the blue band, and G is a scaling factor equal to 2.5. The other parameters, C_1 , C_2 , and L were 6.0, 7.5, and 1.0, respectively.

Using data from the validation area in Fig. 1 for this quantitative validation study, we present a scatter plot of the EVI values from the fused ETM+ versus those from the observed ETM+ EVI data in Fig. 10. As expected, they have a high degree of correlation with an $R^2 = 0.87$, excluding zero and abnormal values. This high correlation with a near 1:1 slope shows that the ANN model does an excellent job of reproducing the basic signal within the observed data in these wavelength bands over vegetation-covered areas.

3.4.2 Normalized difference in moisture index inversion using the ANN simulated ETM+ data

Multi-year averages for data from each month precipitation and temperature in the study area are shown in Fig. 11. The precipitation and temperature data were acquired from The National Oceanic and Atmospheric Administration (NOAA) climate data website (<http://www.climate.gov/data/maps-and-data>). As shown in Fig. 11, the study area possesses a predominately monsoon climate, which is easily influenced by the sub-tropic anticyclones in the spring that causes agricultural drought. Because drought

has such a dramatic impact on agricultural production, it is urgent to be able to effectively extract the surface humidity information from the study area to monitor drought conditions. The traditional surface humidity measurement methods rely on observational stations. Remote sensing provides an alternative to surface humidity monitoring. Gao (1996) was the first to use remote sensing to detect the changes in liquid water content of vegetation canopies. Wilson and Sader (2002) used a Normalized Difference Moisture Index (NDMI) to detect moisture content of a forestry harvest in the State of Maine. NDMI is a normalized ratio index based on reflectance of shortwave infrared and NIR bands. However, because the NDMI responds in time to changes in the water content in the vegetation canopy, it can be used to monitor the water content of vegetation canopy, which is important for drought monitoring. The correlation coefficient between the NDMI and vegetation canopy moisture is above 0.95 (Fiorella and Ripple, 1995; Mallick et al., 2009). Hence, NDMI was used in this study. The equation for the NDMI is

$$NDMI = (\rho_{NIR} - \rho_{SWIR}) / (\rho_{NIR} + \rho_{SWIR}), \quad (6)$$

where ρ_{NIR} and ρ_{SWIR} are the reflectance in the NIR and SWIR bands, respectively. The range of the NDMI value is between -1 to 1. The larger (smaller) the NDMI value, the higher (lower) the vegetation canopy moisture level.

The vegetation canopy moisture index (NDMI) can be easily influenced by the presence of different terrain types within the remotely sensed area. For example, water bodies have a different effect on the NDMI than do bare lands and

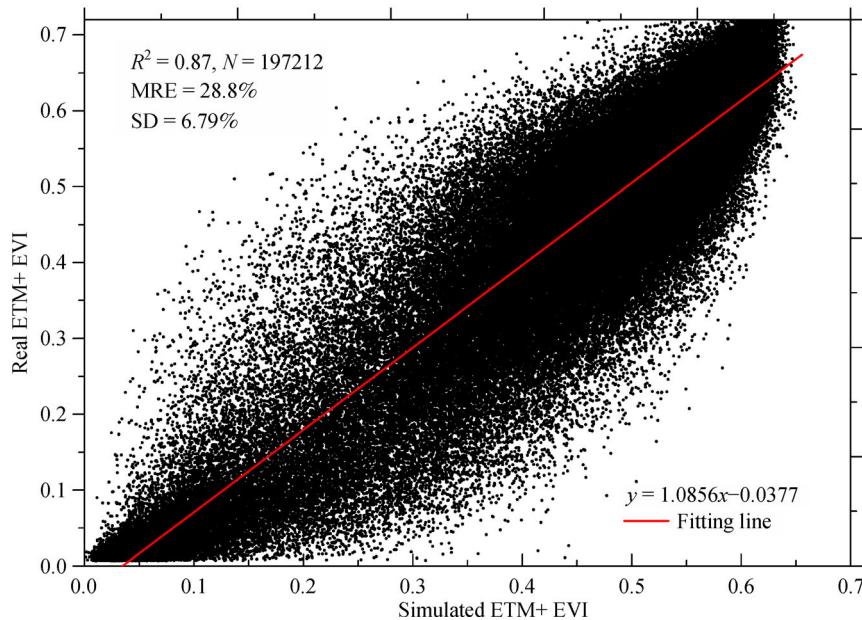


Fig. 10 Comparison of the vegetation index value (EVI) between the ANN model simulated ETM+ and observed ETM+ data. The abscissa is the EVI value computed from the ANN model simulated ETM+ data, and the ordinate is the EVI value computed from the observed ETM+ data.

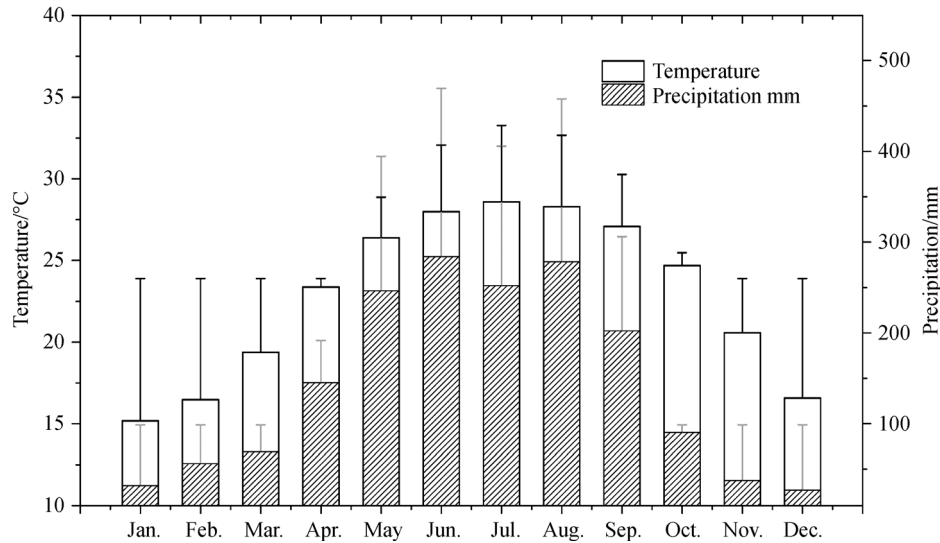


Fig. 11 Multi-year averages for data from each month precipitation and temperature in the study area. The averages for monthly precipitation and temperature datasets is from year 1981 to year 2010.

built-up urban areas. The NDMI value of water bodies is much different from the vegetation index of other terrain types, so a decision-tree classification method was used to extract and eliminate the interference of water bodies on the NDMI. The information on bare land and built-up urban areas is hard to extract due to its complex properties. The normalized difference built-up index (NDBI):

$$NDBI = (\rho_{SWIR} - \rho_{NIR}) / (\rho_{SWIR} + \rho_{NIR}), \quad (7)$$

is a simple but useful metric for distinguishing built-up and bare land areas with larger values of the NDBI from other land terrain types, which have lower NDBI values. This metric has been tested in the classification of Nanjing and found to achieve a classification accuracy of 92.6% (Zha et al., 2003). Since NDBI can be used to identify bare land and built-up areas, the interference coming from those

areas can be isolated.

The result of the NDBI computation (Fig. 12) shows that high vegetation canopy moisture levels exist over the entire study area (Fig. 12), so that there was no obvious sign of agriculture drought in the region. The canopy moisture was highest in the northern part of this region and lowest in the southern part. This may be a consequence of the fact that the southern area is mainly farmland, which has smaller canopy fraction than the more forested northern area. In addition, low canopy moisture values (or NDMI value) can also be found along the perimeter of built-up urban areas, such as roads, man-made objects, etc. Comparing the low canopy moisture value areas in Fig. 12 to the high-resolution image of GE indicates that the lower NDMI areas coincide with areas of lower vegetation canopy density.

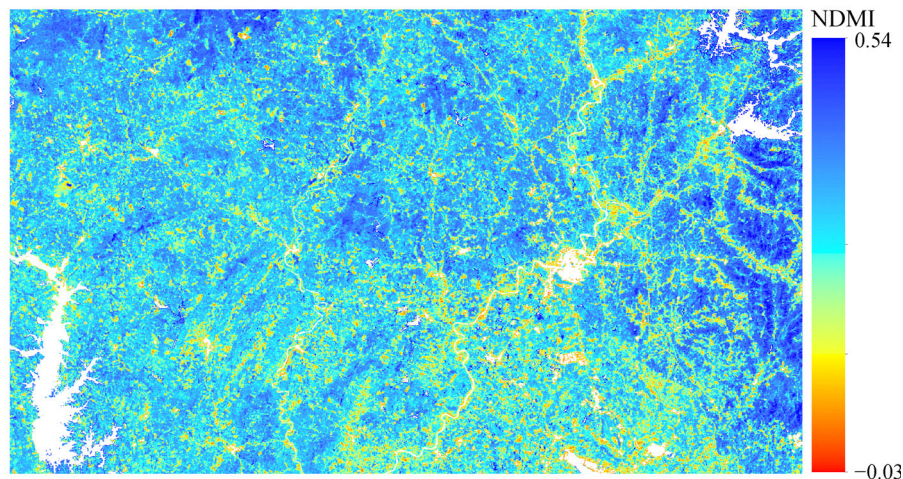


Fig. 12 NDMI in the validation area. Different colors and NDMI values indicate different canopy moisture levels. Blue color with high NDMI value indicates no agriculture drought. By contrast, the red color with low NDMI value indicates agriculture drought.

3.4.3 Water body area extraction in MNDWI using the ANN simulated ETM+ data

It is essential to extract the water body regions efficiently and accurately for water resource investigation, macro-monitoring, wetland protection, shoreline change studies, and flood inundated area assessment, etc. The false color composite image of ETM+ data using SWIR, NIR, and red bands contain the most abundant information on details of terrain feature, which will especially highlight the differences in the various terrain objects in the image. The Normalized Difference Water Index (NDWI) is capable of highlighting water body information by suppressing the information from vegetation and soil as was first proposed by McFeeters (1996). However, because NDWI cannot effectively suppress the information from man-made objects and soil, a new index, MNDWI, that was able to suppress these signals was developed from NDWI (Xu, 2005). Therefore, MNDWI is used to extract water body information for validating the fused data in practical application. The equation for MNDWI is:

$$MNDWI = (\rho_G - \rho_{SWIR}) / (\rho_G + \rho_{SWIR}), \quad (8)$$

where ρ_G is the reflectance in the green band of the fused ETM+ data (corresponds to band 2 of ETM+ data); and ρ_{SWIR} is the reflectance in the SWIR band, which

corresponds to band 5 of the ETM+ data.

We used the SWIR, NIR, and red bands from the fused ETM+ data output from the ANN model to create a false color composite for testing the quality of the ANN fused data for water body extraction. First, the image was converted from reflectance values to MNDWI values through the Basic Tool-Band Math menu of the ENVI software. Then, the water body area was extracted by a decision-tree classification method using a decision condition that the MNDWI value > -0.2 . The white pixels in Figs. 13(a), 13(c), 13(e), and 13(g) are the water body pixels, while the black pixels indicate pixels of any other terrain types.

To further assess the quality of these datasets, the water body type is classified into four typical cases in Fig. 13. Case 1 water bodies are large lakes; Case 2 water bodies are small inter-mountain lakes; Case 3 water bodies are rivers; and Case 4 water bodies are water bodies disturbed by man-made objects. In Figs. 13(a), 13(c), 13(e), and 13(g) correspond to cases 1–4, respectively, extracted by MNDWI, while Figs. 13(b), 13(d), 13(f), and 13(h) are the false color composite images using bands 3–5 fused data for the corresponding water bodies in Figs. 13(a), 13(c), 13(e), and 13(g).

The accuracy of the data fusion results are determined by a comparative analysis between the MNDWI results and

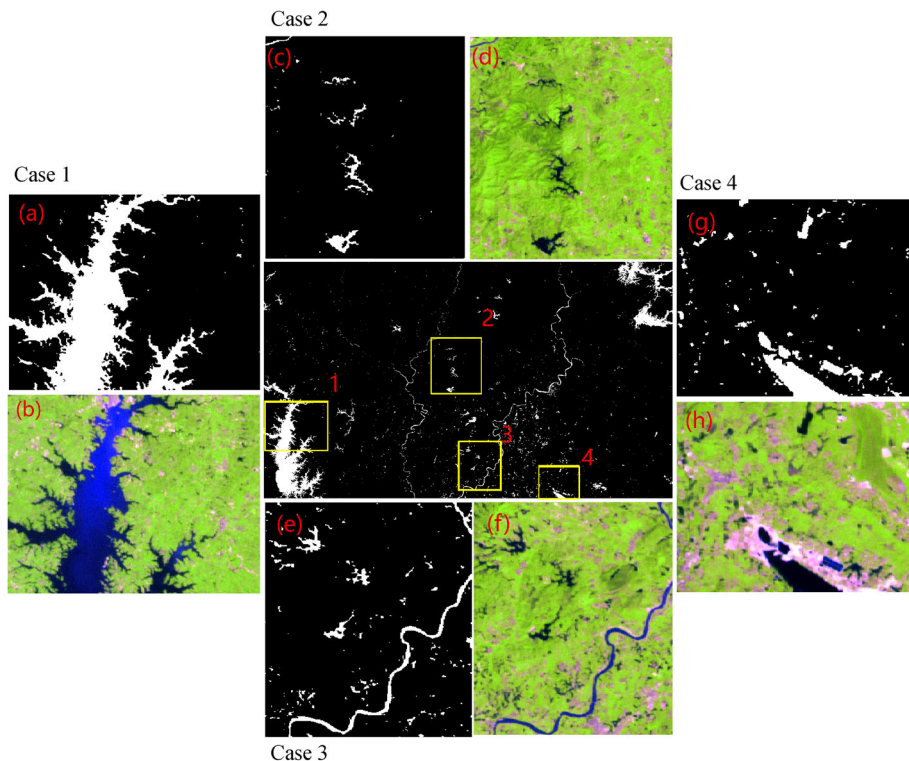


Fig. 13 Detailed comparison between the water body area extracted by MNDWI and high-resolution false color composite image. Panels (a), (c), (e) and (g) correspond to cases 1–4 extracted by MNDWI, while panels (b), (d), (f) and (h) are the false color composite images using bands 3–5 of fused data.

the false color composite images as shown in Fig. 13 (e.g., comparing panels (a) to (b), (c) to (d), (e) to (f), (g) to (h)). For Case 1 lakes (comparing Figs. 13(a) and 13(b)) the information on small islands in the lake, the small branches of large lakes and the edges of land and water was successfully extracted and all the details are presented accurately. For Case 2 lakes, the main body of small inter-mountain lakes that are surrounded by lush vegetation cover shows details of the lakes and their small branches. Case 3 are big rivers with small, connected lakes and pools in the plain areas under the impacts of human activities. All details, even the smallest river pools and larger pools are extracted successfully by the comparison of the MNDWI to the false color image. For case 4, water bodies influenced by man-made objects (buildings, docks, breakwaters, etc.) the comparative images (Figs. 13(g) and 13(h)) clearly delineate the water body information in most areas, except for a small area at the upper right corner of the image. In summary, the ANN fused data can be used to successfully extract the water body information for all water body types. The overall accuracy (OA) of the water body extraction is 97.06% with a Kappa Coefficient of 0.9374. Therefore, the fused data meets the needs of practical applications, and has a high application potential in the future.

3.5 Uncertainty estimates

The signals obtained by satellite sensors are affected by atmospheric water vapor, air molecules, aerosols, and other signals. The information of interest received by the sensor is only a small part of the total signal seen by the satellite. When these data are used in the aforementioned applications, an important issue is to successfully eliminate the effects of aerosols, water vapor, and air molecules that interfere with the radiation propagation process. The Quick Atmospheric Correction (QUAC) model is an easy and fast method that has been shown to perform quite well for atmospheric correction (Bernstein et al., 2005). Therefore, in this study, the widely used QUAC model was chosen as the atmospheric correction model for HJ-CCD and Landsat ETM+ data. The band ratios (such as EVI and MNDWI) can be used to further eliminate atmospheric effects (Lee and Carder, 2000). Therefore, the effects of the atmosphere on the signal propagation can be reduced significantly by atmospheric correction and the application of band ratio analysis.

The HJ-CCD data possess remarkable geometric distortion, while in the Landsat ETM+ data the geometric distortions are accurately eliminated. The most important part of this ANN data fusion model is building a pixel-to-pixel relationship. This relationship can be expressed by a function that does not allow a relative error larger than one pixel. Geometric correction was carried out on the HJ-CCD data by use of the Landsat data with a relative error less than a pixel. In addition, as shown in Fig. 1, the

topographic relief in the validation area is sufficient to complicate the identification of high accuracy ground control sites. Therefore, the ground control sites should be chosen with digital elevation model (DEM) data in an area with significant topographic relief to obtain a higher accuracy geometric correction.

4 Conclusions

In this study, multi-source remote sensing data from the cloudy, coastal area of the South China Sea were used to perform data fusion based on the BP-ANN method using Landsat ETM+ and HJ-CCD data. The data fusion results were assessed qualitatively, quantitatively, and through consideration of the practical application of the fused data results.

Qualitative assessment was accomplished by comparisons between the fused data and Landsat ETM+ data in terms of spectral resolution, spatial resolution and image clarity. The spectral resolution of the HJ-CCD data was improved from 4 bands to 6 bands by the data fusing process. The results of the data fusion showed no gaps in the entire fused image and that the fused data also have 6 bands corresponding to the ETM+ data. Moreover, the clarity of the HJ-CCD data image was greatly improved in the fused data image.

For the quantitative assessment, regions containing water bodies, forests, farmlands, and urban areas were chosen for a further study with the high resolution data. A comparative study was accomplished for the fused and ETM+ data. The results showed the smallest differences in reflectance exist between the fused and ETM+ data for the forest and farmland area, with slightly larger differences occurring in areas containing water bodies. The worst performance in terms of differences between observed and fused reflectance values occurred for areas including man-made objects. In a further assessment of the accuracy of the fused data, the correlation coefficients between observed ETM+ data and ANN model fused ETM+ data were calculated for bands 1–5, and 7. The results show that high correlation coefficients and small differences between the observed and fused data occur in all bands, with the infrared band showing the best performance.

While both the qualitative and quantitative assessments showed that the data fusion process provided results with very good performance, a further practical application of the results was used as a further assessment of the data fusion methodology. As mentioned above, because a large fraction of the validation area is covered by vegetation and water bodies, the practical applications included three parts: 1) Since there were small differences between the fused data and ETM+ data in forest and farmland areas that contain relatively abundant vegetation, the data fusion results were assessed quantitatively by calculating the

correlation coefficient of EVI values between the fused and observed ETM+ data. The results found the correlation coefficient between the EVI of fused and observed images was 0.937 ($R^2=0.87$); 2) A normalized difference moisture index (NDMI) was used to infer the moisture levels in the canopy of study area. The results showed that the NDMI computed for the fused ETM+ data provided reasonable results for canopy moisture; 3) Water body area extraction using MNDWI computed for the fused data was another assessment of the accuracy of data fusion in a practical application. The results showed that by comparative analysis of the water extraction results using MNDWI, with an overall accuracy of 97.06% and a Kappa coefficient of 0.9374, the fused data can totally meet the needs of any practical application.

Relatively high quality data can be fused together with no gaps, multiple-bands, and high clarity using the BP-ANN fusion model. High correlation coefficient values and very small differences in reflectance exist between the observed ETM+ data and the fused ETM+ data generated by the BP-ANN model. The fused ETM+ data can meet the needs of any practical application after proper validation. The advantages of both HJ-CCD data and Landsat ETM+ data can be fully integrated while the disadvantages can be minimized using the BP-ANN data fusion model. In addition, there is no need to have a thorough understanding of the experimental area nor are ground-based sites and huge data operation necessary for using the BP-ANN model. The BP-ANN model is simple, efficient, and of high accuracy, and it can be easily extended as a data fusion method, with potential application in cloudy and rainy coastal areas.

Acknowledgements The authors would like to thank the reviewers and the editor for their constructive comments. This study was supported by the National Key Research and Development Program of China (2016YFC1400901), the High Resolution Earth Observation Systems of National Science and Technology Major Projects (41-Y20A31-9003-15/17), the National Natural Science Foundation of China (Grant Nos. 41476156 and 41621064), and the Public Science and Technology Research Funds Projects of Ocean (201005030).

References

- Al-Sbou Y A (2012). Artificial neural networks evaluation as an image denoising tool. *World Appl Sci J*, 17(2): 218–227
- Amici G, Dell'Acqua F, Gamba P, Pulina G (2004). A comparison of fuzzy and neuro-fuzzy data fusion for flooded area mapping using SAR images. *Int J Remote Sens*, 25(20): 4425–4430
- Benediktsson J A, Swain P H, Ersoy O K (1989). Neural Network Approaches Versus Statistical Methods in Classification of Multi-source Remote Sensing Data. In: 12th Canadian Symposium on Remote Sensing Geoscience and Remote Sensing Symposium, 489–492
- Bernstein L S, Adler-Golden S M, Sundberg R L, Levine R Y, (2005). A new method for atmospheric correction and aerosol optical property retrieval for VIS-SWIR multi- and hyperspectral imaging sensors: QUAC (QUick atmospheric correction). *Geoscience and Remote Sensing Symposium*, 2005. IGARSS'05. Proceedings. 2005 IEEE International IEEE, 2005: 3549–3552
- Bossé É, Roy J, Paradis S (2000). Modeling and simulation in support of the design of a data fusion system. *Inf Fusio*, 1(2): 77–87
- Busetto L, Meroni M, Colombo R (2008). Combining medium and coarse spatial resolution satellite data to improve the estimation of sub-pixel NDVI time series. *Remote Sens Environ*, 112(1): 118–131
- Chen F, Tang L, Wang C, Qiu Q (2011a). Recovering of the thermal band of Landsat 7 SLC-off ETM+ image using CBERS as auxiliary data. *Adv Space Res*, 48(6): 1086–1093
- Chen J, Zhu X, Vogelmann J E, Gao F, Jin S (2011b). A simple and effective method for filling gaps in Landsat ETM+ SLC-off images. *Remote Sens Environ*, 115(4): 1053–1064
- Chen Y, Deng L, Li J, Li X, Shi P (2006). A new wavelet - based image fusion method for remotely sensed data. *Int J Remote Sens*, 27(7): 1465–1476
- Chen Z Y, Desai M, Zhang X P (1997). Feedforward neural networks with multilevel hidden neurons for remotely sensed image classification. In: *International Conference on Image Processing*, 2: 653–656
- Daily M I, Farr T, Elachi C, Schaber G (1979). Geologic interpretation from composited radar and Landsat imagery. *Photogramm Eng Remote Sensing*, 45(8): 1109–1116
- Ehlers M (1991). Multi sensor image fusion techniques in remote sensing. *ISPRS J Photogramm Remote Sens*, 46(1): 19–30
- Fan J, Zhao D, Wang J (2014). Oil Spill GF-1 Remote Sensing Image Segmentation Using an Evolutionary Feedforward Neural Network. In: *IEEE International Joint Conference on Neural Networks (IJCNN)*, 446–450
- Faouzi N E, Leung H, Kurian A (2011). Data fusion in intelligent transportation systems: progress and challenges – A survey. *Inf Fusion*, 12(1): 4–10
- Farifteh J, Van der Meer F, Atzberger C, Carranza E J M (2007). Quantitative analysis of salt-affected soil reflectance spectra: a comparison of two adaptive methods (PLSR and ANN). *Remote Sens Environ*, 110(1): 59–78
- Fiorella M, Ripple W J (1995). Determining successional stage of temperate coniferous forests with landsat satellite data. *Photogramm Eng Remote Sensing*, 59(2): 239–246
- Gao B C (1996). NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sens Environ*, 58(3): 257–266
- Gigli G, Bossé É, Lampropoulos G A (2007). An optimized architecture for classification combining data fusion and data-mining. *Inf Fusion*, 8(4): 366–378
- Hilker T, Wulder M A, Coops N C, Linke J, McDermid G, Masek J G, Gao F, White J C (2009). A new data fusion model for high spatial- and temporal-resolution mapping of forest disturbance based on Landsat and MODIS. *Remote Sens Environ*, 113(8): 1613–1627
- Hu Q, Wu W, Xia T, Yu Q, Yang P, Li Z, Song Q (2013). Exploring the use of Google Earth imagery and object-based methods in land use/cover mapping. *Remote Sens*, 5(11): 6026–6042
- Jacobson A, Dhanota J, Godfrey J, Jacobson H, Rossman Z, Stanish A, Walker H, Riggio J (2015). A novel approach to mapping land conversion using Google Earth with an application to East Africa.

- Environ Model Softw, 72: 1–9
- Karayiannis N B, Purushothaman G (1994). Fuzzy pattern classification using feedforward neural networks with multilevel hidden neurons. Paper presented at the IEEE International Conference on Neural Networks, 1994. IEEE World Congress on Computational Intelligence
- Khaleghi B, Khamis A, Karray F O, Razavi S N (2013). Multisensor data fusion: a review of the state-of-the-art. *Inf Fusion*, 14(1): 28–44
- Kiema J B K (2002). Texture analysis and data fusion in the extraction of topographic objects from satellite imagery. *Int J Remote Sens*, 23(4): 767–776
- Lee Z, Carder K L (2000). Band-ratio or spectral-curvature algorithms for satellite remote sensing. *Appl Opt*, 39(24): 4377–4380
- Liu R, Sun J, Wang J, Liao X (2011). Data quality evaluation of Chinese HJ CCD sensor. *Advances in Earth Science*, 26(9): 971–979
- Liu Y, Chen X, Peng H, Wang Z (2017a). Multi-focus image fusion with a deep convolutional neural network. *Inf Fusion*, 36: 191–207
- Liu Z, Blasch E, John V (2017b). Statistical comparison of image fusion algorithms: recommendations. *Inf Fusion*, 36: 251–260
- Maeda E E, Formaggio A R, Shimabukuro Y E, Arcoverde G F B, Hansen M C (2009). Predicting forest fire in the Brazilian Amazon using MODIS imagery and artificial neural networks. *Int J Appl Earth Obs Geoinf*, 11(4): 265–272
- Mallick K, Bhattacharya B K, Patel N K (2009). Estimating volumetric surface moisture content for cropped soils using a soil wetness index based on surface temperature and NDVI. *Agric Meteorol*, 149(8): 1327–1342
- Maxwell S K, Schmidt G L, Storey J C (2007). A multi-scale segmentation approach to filling gaps in Landsat ETM+ SLC-off images. *Int J Remote Sens*, 28(23): 5339–5356
- McFeeters S K (1996). The use of the normalized difference water index (NDWI) in the delineation of open water features. *Int J Remote Sens*, 17(7): 1425–1432
- Mehta A, Parihar A S, Mehta N (2015). Supervised Classification of Dermoscopic Images using Optimized Fuzzy Clustering based Multi-Layer Feed-Forward Neural Network. 2015 International Conference on Computer, Communication and Control (IC4)
- Mohamdy M, Moradi H R, Zeinivand H, Temme A J A M, Pourghasemi H R, Alizadeh H (2014). Validating gap-filling of Landsat ETM+ satellite images in the Golestan Province, Iran. *Arab J Geosci*, 7(9): 3633–3638
- Mohan S, Mehta R L (1988). Combined Radar and Landsat data analysis for land use/cover studies over parts of the Punjab plains. *J Indian Soc Remote Sens*, 16(4): 33–36
- Muskat J (1983). Geologic interpretations of Seasat-A radar images and Landsat MSS images of a portion of the southern Appalachian Plateau, Virginia, Kentucky, West Virginia. California State University Northridge
- Nachouki G, Quafafou M (2008). Multi-data source fusion. *Inf Fusion*, 9(4): 523–537
- Nguyen H, Katzfuss M, Cressie N, Braverman A (2014). Spatio-temporal data fusion for very large remote sensing datasets. *Technometrics*, 56(2): 174–185
- Novelli A, Tarantino E, Fratino U, Iacobellis V, Romano G, Gentile F (2016). A data fusion algorithm based on the Kalman filter to estimate leaf area index evolution in durum wheat by using field measurements and MODIS surface reflectance data. *Remote Sens Lett*, 7(5): 476–484
- Sharma S C, Rajendran N, Grover A K, Srivastava G S (1990). Interpretation of Synthetic Aperture Radar (SAR) imagery for geological appraisal: a comparative study in Anantapur district of Andhra Pradesh. *J Indian Soc Remote Sens*, 18(4): 45–64
- Sims D A, Rahman A F, Cordova V D, Elmasri B, Baldocchi D, Bolstad P, Flanagan L, Goldstein A, Hollinger D, Misson L (2008). A new model of gross primary productivity for North American ecosystems based solely on the enhanced vegetation index and land surface temperature from MODIS. *Remote Sens Environ*, 112(4): 1633–1646
- Suliman S I (2016). Locally linear manifold model for gap-filling algorithms of hyperspectral imagery: proposed algorithms and a comparative study. Dissertation for Master Degree. Michigan State University, 1–73
- Tedesco M, Pulliainen J, Takala M, Hallikainen M, Pampaloni P (2004). Artificial neural network-based techniques for the retrieval of SWE and snow depth from SSM/I data. *Remote Sens Environ*, 90(1): 76–85
- Toutin T (1995). Intégration de données multisources: comparaison de méthodes géométriques et radiométriques. *Int J Remote Sens*, 16(15): 2795–2811
- Turker M, San B T (2003). SPOT HRV data analysis for detecting earthquake-induced changes in Izmit, Turkey. *Int J Remote Sens*, 24(12): 2439–2450
- Weckenmann A, Jiang X, Sommer K D, Neuschaefer-Rube U, Seewig J, Shaw L, Estler T (2009). Multisensor data fusion in dimensional metrology. *CIRP Annals- Manufacturing Technology*, 58(2): 701–721
- Welch R, Ehlers M (1987). Merging multiresolution SPOT HRV and Landsat TM data. *Photogramm Eng Remote Sensing*, 53: 301–303
- Wilson E H, Sader S A (2002). Detection of forest harvest type using multiple dates of Landsat TM imagery. *Remote Sens Environ*, 80(3): 385–396
- Wu M Q, Wang J, Niu Z, Zhao Y Q, Wang C Y (2012). A model for spatial and temporal data fusion. *J Infrared Millim W*, 31(1): 80–84
- Xu H Q (2005). A study on information extraction of water body with the modified normalized difference water index (MNDWI). *J Remot Sens*, 9(5): 589–595
- Zeng C, Shen H, Zhang L (2013). Recovering missing pixels for Landsat ETM+ SLC-off imagery using multi-temporal regression analysis and a regularization method. *Remote Sens Environ*, 131: 182–194
- Zervas E, Mpimpoudis A, Anagnostopoulos C, Sekkas O, Hadjiefthymiades S (2011). Multisensor data fusion for fire detection. *Inf Fusion*, 12(3): 150–159
- Zha Y, Gao J, Ni S (2003). Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *Int J Remote Sens*, 24(3): 583–594
- Zhang C, Li W, Travis D (2007). Gaps-fill of SLC-off Landsat ETM+ satellite image using a geostatistical approach. *Int J Remote Sens*, 28(22): 5103–5122
- Zhu X, Liu D, Chen J (2012). A new geostatistical approach for filling gaps in Landsat ETM+ SLC-off images. *Remote Sens Environ*, 124: 49–60