RESEARCH   **Open Access**

Land cover classification combining Sentinel-1 and Landsat 8 imagery driven by Markov random field with amendment reliability factors

Xiaofei Shi[1*], Zhiyu Deng[1], Xing Ding[1] and Li Li[2]

* Correspondence: shixiaofei@dlmu.edu.cn
[1]Information Science and Technology College, Dalian Maritime University, Dalian 86-116026, Liaoning, China
Full list of author information is available at the end of the article

Abstract

Reliability factors in Markov random field (MRF) could be used to improve classification performance for synthetic aperture radar (SAR) and optical images; however, insufficient utilization of reliability factors based on characteristics of different sources leaves more room for classification improvement. To solve this problem, a Markov random field (MRF) with amendment reliability factors classification algorithm (MRF-ARF) is proposed. The ARF is constructed based on the coarse label field of urban region, and different controlling factors are utilized for different sensor data. Then, ARF is involved into the data energy of MRF, to classify the sand, vegetation, farmland, and urban regions, with the gray level co-occurrence matrix textures of Sentinel-1 imagery and the spectral values of the Landsat 8 imagery. In the experiments, Sentinel-1 and Landsat-8 images are used with overall accuracy and Kappa coefficient to evaluate the proposed algorithm with other algorithms. Results show that the overall accuracy of the proposed algorithm has the superiority of about 20% in overall precision and at least 0.2 in Kappa coefficient than the comparison algorithms. Thus, the problem of insufficient utilization of different sensors data could be solved.

Keywords: Markov random field, Amendment reliability factors, SAR image, Optical image, Classification

1 Introduction

The availability of reliable land cover information is of great importance for many earth scientific applications, such as the transition of land, increasing demand for food and fiber, biodiversity, and climate [1–5]. Earth observation has been proven to be one of the most useful and efficient approach for land cover classification because it can acquire large scale land cover information quickly and repeatedly [6–8]. Mapping and monitoring land cover, optical remote sensing data has been extensively used for decades, because of its ability to cover large areas in high temporal frequency and overcome the problem of inaccessible areas [9]. However, optical remote sensing data

could be affected by unfavorable weather condition, and there are some difficulties related to the so-called spectral confusions that lower the classification accuracy [10]. Synthetic aperture radar (SAR) as an active remote sensing technique can capture information in all-day, and even in unfavorable weather conditions, but could not provide spectral information, resulting in difficulties in image interpretation [11]. So the data information provided by a single sensor is incomplete, inconsistent, or inaccurate. Thus, utilizing SAR and optical data as sources are indeed highly complementary, where SAR data could ensure all-day and all-weather coverage, and optical data could provide abundant spectral information. So the joint usage SAR and optical data have been adopted in many applications [12]. In this regard, combination of two kinds of data could be utilized to improve good performance on land cover classification.

Joint optical-SAR data classification has been addressed for a couple of decades and many methodological approaches have been proposed, including statistical pattern recognition, neural networks, decision fusion, evidence theory, kernel-based learning, and Markov random field (MRF) [13]. Among these methods, MRF is a probabilistic model that is used to integrate spatial information into image classification [14], demonstrating the advantage to improve classification performance. Moser integrated SVMs and Markov random field models in a unique formulation for spatial contextual classification [15]. Hedhli proposed a classification framework based on hierarchical Markov random fields [16]. Tarabalka involved the edge information into spatial energy term to improve the classification performance [17]. Solberg proposed the Markov random fields with reliability factors, with GIS data to complete multisource classification [18].

The motivation of this paper is to propose a MRF algorithm with amendment reliability factors by utilizing Sentinel-1 and Landsat8 images for land cover classification at coastal regions. Inspired by [18, 19], the information of different sources appear not equally reliable. Thanks giving to the usage of GIS data, [18] could provide good classification results. If GIS data is not available, MRF with reliable factors may result in classification performance degeneration due to insufficient utilization of several sensors' data.

In this paper, a classification algorithm based on MRF with amendment reliability factors is proposed. Based on the coarse urban label field, the additional controlling factors are involved in reliability factors to construct the amendment reliability factors. The amendment reliability factors could fully utilize the advantage of Landsat 8 and Sentinel-1 to balance the contribution of weight in the data term of MRF. The performance is compared to some existing algorithms. The MRF with amendment reliability factors performs better than those algorithms.

The rest of this paper is organized as follows. In Section 2, related works are briefly reviewed, and the proposed classification algorithm driven by amendment reliability factors is described. Experimental results and discussions are provided in Section 3 and 4. Finally, conclusion is drawn in Section 5.

2 Methods
2.1 Related works
Assuming that the images are derived from n sensors, the size of the image taken from one of the sensors with number s is $M \times N$; that means each image provided by this

sensor has $M \times N$ pixels. If feature information has been extracted from this image, the feature vector of each pixel in this sensor could be expressed as $X_S(1, 1)$, ..., $X_S(M, N)$, $S = 1, 2, . . . n$. Similarly, if the images provided by the sensor s have D_s bands, we can also use images $X_S(i, j)$ representing a grayscale vector for whole bands of the sensor s at the location of (i, j). That is, $X_S(i, j) = (X_S(i, j, 1), ..., X_S(i, j, D_s))$, where $X_S(i, j, g)$ represents the grayscale value of the band g. In this paper, n is 2, indicating the considered images consisting of a Sentinel-1 SAR image and a Landsat-8 OLI optical image.

In the images derived from n sensors, it is assumed that there are K objects, namely ω_1, ..., ω_K. For prior probability $P(\omega_1)$, ..., $P(\omega_K)$, $C = \{C(i, j); 1 \leq i \leq M, 1 \leq j \leq N\}$ denotes the label set for the whole scene, where $C(i, j) \in \{\omega_1, \omega_2, ..., \omega_k\}$. All the pixels in the images could be represented by $X_S = \{X_S(i, j); 1 \leq i \leq M, 1 \leq j \leq N\}$.

The task of multisource classification is to maximize the posterior probability $P(C|X_1, ..., X_n)$ of each pixel, depicted as [18]:

$$P(C|X_1, ..., X_n) = \frac{P(X_1, ..., X_n|C)P(C)}{P(X_1, ..., X_n)} \tag{1}$$

where $P(X_1, ..., X_n| C)$ is the conditional probability of feature vector X_1, ..., X_n with the label C. $P(C)$ is a prior probability, and $P(X_1, ..., X_n)$ is the probability of n sensors data.

Assume the images of different sensors are identical independent distributions, we could get the formula as $P(X_1, ..., X_n| C) = P(X_1| C) \cdots P(X_n| C)$. The corresponding weight value is given according to the reliability factors for each sensor data. Thus, the posterior probability could be formulated as [18]:

$$L(C|X_1, ..., X_n) = P(X_1|C)^{\alpha_1} \cdots P(X_n|C)^{\alpha_n} P(C) \tag{2}$$

where α_s denotes a reliability factor with $0 \leq \alpha_s \leq 1$. If the sensor s has low reliability, α_s is zero resulting $P(X_s|C)^{\alpha_s} = 1$. This means that the conditional probability will have no effect on the likelihood function, while, for a sensor with lesser reliability, the closer the reliability factor is close to 0, the more larger it will be to the posterior probability. By using spatial information, prior probability for class labels $P(C)$ could be depicted as [20]

$$P(C(i, j)|C(k, l), \{k, l\} \neq \{i, j\}) = P(C(i, j)|C(k, l), \{k, l\} \in \xi_{ij}) = \frac{1}{Z} e^{-U(C)/T} \tag{3}$$

where U is the potential energy function, C denotes the clique, Z is the normalization constant, T is the temperature constant, ξ_{ij} is the local neighborhood pixel set of the pixel (i, j).

Literature [21] pointed out that the reliability of a sensor could upgrade or downgrade the data classes. The conditional probabilities can be grouped into a matrix R, and denoted by

$$R = \begin{bmatrix} P(\omega_1|X_{S,1}) & P(\omega_2|X_{S,1}) & ... & P(\omega_K|X_{S,1}) \\ P(\omega_1|X_{S,2}) & P(\omega_2|X_{S,2}) & ... & P(\omega_K|X_{S,2}) \\ . & . & ... & . \\ . & . & ... & . \\ . & . & ... & . \\ P(\omega_1|X_{S,Z}) & P(\omega_2|X_{S,Z}) & ... & P(\omega_K|X_{S,Z}) \end{bmatrix} \tag{4}$$

where $X_{S,1}$ represents the feature vector of the first pixel in the sensor image s or the gray vector at the first pixel position (assuming there are multiple bands), i.e., $X_{S,1} = X_S(1, 1)$. $X_{S,Z}$ represents the feature vector of the last pixel in the sensor image or the last pixel position gray vector, i.e., $X_{S,Z} = X_S(M, N)$ with $Z = M \times N$. If the sensor data is reliable, the class label information for each data of the sensor will be unique. Specifically, each row in the matrix has only one value with 1, and all the others would be zero. If the sensor data is extremely unreliable, then the class label information for each data is random. Thus, there is an uncertainty of $\log[1/P(\omega_j|X_{s,i})]$ for a certain observation data $X_{s,i}$ at sensor s with class information ω_j, and the average uncertainty of the data about the class information can be calculated as [21]

$$H(\omega|X_{s,i}) = \sum_j P(\omega_j|X_{s,i}) \log \frac{1}{P(\omega_j|X_{s,i})} \tag{5}$$

If the uncertainty of the sensor data can be measured as $H(\omega|X_s)$, it could be expressed by

$$\begin{aligned} H(\omega|X_s) &= \sum_i P(X_{s,i})H(\omega|X_{s,i}) \\ &= \sum_j \sum_i P(X_{s,i})P(\omega_j|X_{s,i}) \log \frac{1}{P(\omega_j|X_{s,i})} \\ &= \sum_j \sum_i P(X_{s,i}, \omega_j) \log \frac{1}{P(\omega_j|X_{s,i})} \end{aligned} \tag{6}$$

If only $\alpha_s = H(\omega|X_s)$ of the image is taken as the sole basis for the reliability factor of the image, it could have influence on the posterior probability (2). If α_s is small except 0, the sensor data has more effect on (2). Conversely, it plays less role. However, due to the different advantages of different sensors on certain land cover, the reliability factors may result in low classification performance or misclassification. When a reliability factor is not zero, a small reliability factor will make the posterior probability large, and a large reliability factor will make the posterior probability small. Taking two sensors as an example, sensor 1 has an advantage on land cover A over sensor 2, but the reliability factor of sensor 1 may be large and the reliability factor of sensor 2 is small, resulting in a large contribution of sensor 2 to classification. So the classification performance will not be as good as that of sensor 1 for land cover A. Furthermore, the land cover A is classified as another class with small reliability factor, resulting in misclassifications. Therefore, for certain land cover, a sensor with good classification advantage may be chosen. Assuming correct classification, the reliability factor should be small, and the posterior probability is large. The sensor without such an advantage should be with a large reliability factor. Then, the posterior probability is small. This will improve the classification performance to a certain extent and reduce the problem of misclassifications.

In order to solve this problem, we use different reliability factors for different land covers. Different sensors in the same region have different reliability factors, so that the data with better classification advantage could be more important, that is, the reliability factor with better classification ability is lesser. Different sensors use different reliability factors in different land covers. Different sensor data have different classification ability for different land covers, such as SAR images having

better classification for urban areas, but for those areas without much detail, such areas could not obtain good results. The optical image has rich spectral information, so the discrimination ability of areas without detailed information is larger. Therefore, we divide the image to be classified into the urban area and the non-urban area and give different reliability factors to different sensor data for different areas.

2.2 Proposed method

Since the reliability factors of the two sensors are fixed in (5), reliability factors in different object regions could not be given full play to their classification advantages. In order to solve this problem, we assume that different object areas should adopt different reliability factors, so that data with good classification advantage could play a greater role in different land covers, that is, the reliability factor with good classification ability is lesser and vice versa. Starting from the reliability factors, we divide the reliability factor of Eq. (5) into two quantities as

$$
\begin{cases}
\lambda'_{\text{SAR},i} = H\big(\omega|X_{\text{SAR},i}\big) \\
\qquad = \sum_j P\big(\omega_j|X_{\text{SAR},i}\big) \log \dfrac{1}{P\big(\omega_j|X_{\text{SAR},i}\big)} \\
\\
\lambda'_{\text{Optical},i} = H\big(\omega|X_{\text{Optical},i}\big) \\
\qquad = \sum_j P\big(\omega_j|X_{\text{Optical},i}\big) \log \dfrac{1}{P\big(\omega_j|X_{\text{Optical},i}\big)}
\end{cases}
\tag{7}
$$

where $\lambda'_{\text{SAR},i}$ and $\lambda'_{\text{Optical},i}$ are the reliability factors of the i pixel in the SAR image and the optical image, respectively. $X_{\text{SAR},\,i}$ and $X_{\text{Optical},\,i}$ represent the i pixel in the SAR image and the optical image, respectively. Normalizing the SAR image and the optical image reliability factors, the reliability factors at the same position in the SAR image and the optical image both could reach 1. However, when the two reliability factors are close to each other, the advantage of one image cannot be highlighted, which leads to the inability to improve the classification accuracy. To solve this problem, according to the literature [22], we introduce the idea of stretching. We stretch (7) and get

$$
\begin{cases}
\lambda_{\text{SAR},i} = \dfrac{1/\Big(1+\ \exp\Big(-16 \times \lambda'_{\text{SAR},i} + 4\Big)\Big)}{1/\Big(1+\ \exp\Big(-16 \times \lambda'_{\text{SAR},i} + 4\Big)\Big) + 1/\Big(1+\ \exp\Big(-16 \times \lambda'_{\text{Optical},i} + 4\Big)\Big)} \\
\\
\lambda_{\text{Optical},i} = \dfrac{1/\Big(1+\ \exp\Big(-16 \times \lambda'_{\text{Optical},i} + 4\Big)\Big)}{1/\Big(1+\ \exp\Big(-16 \times \lambda'_{\text{SAR},i} + 4\Big)\Big) + 1/\Big(1+\ \exp\Big(-16 \times \lambda'_{\text{Optical},i} + 4\Big)\Big)} \\
\\
\lambda_{\text{SAR},i} + \lambda_{\text{Optical},i} = 1
\end{cases}
\tag{8}
$$

where $\lambda_{\text{SAR},\,i}$ and $\lambda_{\text{Optical},\,i}$ are the normalized reliability factors. The aim of (8) is to make highly reliable sensor data having more contribution in the classification process.
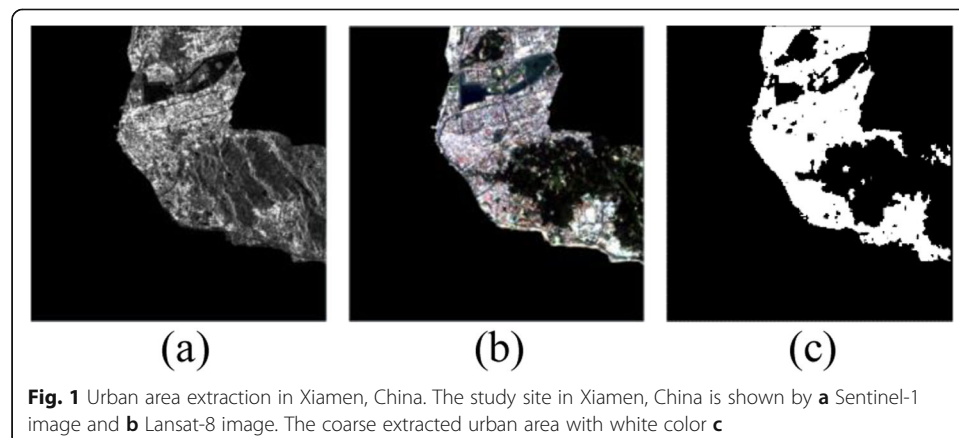
At medium resolution, SAR images have better classification accuracy for texture areas such as urban areas than optical images [23], which means that SAR images can
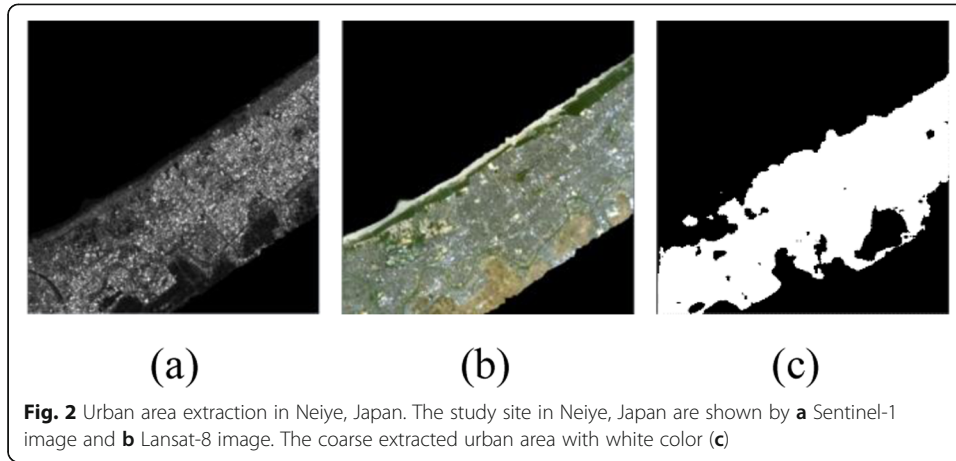
have better recognition ability for urban areas. In SAR images, there are often many bright spots inside the building area that are reflected by objects such as oblique roofs and sharp corners. The middle of the bright spot is mixed with shadows, black roads, and light gray blocks caused by vegetation, and the arrangement of the buildings is usually relatively neat, so it is easy to form a texture with a regular light and dark interval [24]. Therefore, we perform a measure of uniformity on SAR images, to extract urban areas and obtain image classification method with different reliability factors for different sensor images.

At present, the extraction of urban areas based on SAR images has been reported. In the literatures [24, 25], both of them using the gray level co-occurrence matrix texture [26] as the main means for building extraction in SAR images. In [25], the extracted urban area is used as a marking field, and the SAR image is divided into urban area and non-urban area. Then, a joint utilization rule from the SAR image and multi-spectral image based on these two different areas is given.

Inspired by this idea, in order to improve the measurement of reliability factors, we introduce urban area as the label field into the classification of SAR and optical images. In the extraction of urban areas, we use the entropy information of the gray level co-occurrence matrix used in [25]. The difference is that our method does not use the block-based extraction urban area strategy in [25], because the edge fit of [25] is poor according to the experiments. Here, we use a pixel-based approach to extract urban areas. The specific description is as follows: First, the gray level co-occurrence matrix is calculated for the SAR image, and the entropy information is calculated from the gray level co-occurrence matrix [25].

In order to improve the accuracy of urban area extraction, we need to make full use of the information of SAR and optical images. The extracted urban area could be obtained by an entropy threshold for Sentinel-1 image with 0.6 (parameter sensitivity analysis could be seen in experimental section), providing a coarse label field. Figure 1 a–c are the Sentinel-1 image, the Landsat-8 image, and the coarsely extracted urban area image in Xiamen, China. Figure 2a–c are the Sentinel-1 image, the Landsat-8 image, and the coarsely extracted urban area image in Neiye, Japan.

Fig. 1 Urban area extraction in Xiamen, China. The study site in Xiamen, China is shown by **a** Sentinel-1 image and **b** Lansat-8 image. The coarse extracted urban area with white color **c**

Fig. 2 Urban area extraction in Neiye, Japan. The study site in Neiye, Japan are shown by **a** Sentinel-1 image and **b** Lansat-8 image. The coarse extracted urban area with white color (**c**)

After obtaining the urban area, we propose the strategy to obtain amendment reliability factor for the urban area and the non-urban area. In order to make the contribution of the highly reliable sensor data in classification of these two sources, the strategy with amendment reliability factors could be depicted as: It is assumed that ω_B denotes the urban area label, and $\omega_{B'}$ indicates the label of the non-urban area. We define that the amendment factor could be expressed by the reliability factors in (7) added with controlling factors such as $\lambda_e = 1$, $\lambda'_e = 0$. If the current pixel is in the urban area, the conditional probability $P(X_i|\omega_B)$ belonging to the urban category depicted as $P(X_i|\omega_B) = P(X_{\text{SAR},i}|\omega_B)^{\lambda_{\text{SAR},i}+\lambda'_e} \times P(X_{\text{Optical},i}|\omega_B)^{\lambda_{\text{Optical},i}+\lambda'_e}$ should be increased, where the amendment reliability factor is denoted as $\alpha_{s,i} = \lambda_{s,i} + \lambda'_e$, while the conditional probability of none-urban $P(X_i|\omega_{B'})$ belonging to the urban category could be denoted as $P(X_i|\omega_{B'}) = P(X_{\text{SAR},i}|\omega_{B'})^{\lambda_{\text{SAR},i}+\lambda_e} \times P(X_{\text{Optical},i}|\omega_{B'})^{\lambda_{\text{Optical},i}+\lambda_e}$ and the amendment factor could be denoted with $\alpha_{s,\ i} = \lambda_{s,\ i} + \lambda_e$. The reason is that the conditional probability $P(X_{s,\ i}|\omega_j)$ of the i pixel has the value between 0 and 1. Therefore, the lesser the value of the conditional probability, the lesser the sensor will contribute to the classification. If the current pixel is in the urban area, the value of the amendment reliability factor of a sensor data under the correct label should be as small as possible, and the reliability of the sensor for non-building with incorrect label should be larger, so the former case λ'_e is added, while the latter case λ_e is added. Thus, the probability that the current point is judged to be a non-urban area becomes small, and the probability of being judged as an urban area becomes large. Conversely, if the current pixel is in a non-urban area, the probability that the current point is judged to be an urban should be as small as possible, and the current point is a non-urban should be as large as possible. It means that $P(X_{\text{Fused},i}|\omega_B) = P(X_{\text{SAR},i}|\omega_B)^{\lambda_{\text{SAR},i}+\frac{1}{\text{ep}}} \times P(X_{\text{Optical},i}|\omega_B)^{\lambda_{\text{Optical},i}+\frac{1}{\text{ep}}}$ should be small for the pixel in the non-urban area with the amendment reliability factor $\alpha_{s,i} = \lambda_{s,i} + \frac{1}{\text{ep}}$, and vice versa, where ep is a very small positive real number. The amendment reliability factors in non-urban belong to the non-urban in the SAR image and the optical image could be denoted as $\alpha_{\text{SAR},\ i} = \lambda_{\text{SAR},\ i} + \lambda_e$, and $\alpha_{\text{Optical},i} = \lambda_{\text{Optical},i} + \lambda'_e$ respectively. The reason is that when

the current pixel is in a non-urban area, the value of the amendment reliability factor ω_B under a given label with urban should be as large as possible, and the design $\alpha_{s,i} = \lambda_{s,i} + \frac{1}{ep}$ ensures that the value of the amendment reliability factor is large enough. The smaller the value of the amendment reliability factor $\omega_{B'}$ of the sensor under the given label with non-building should be smaller. In order to highlight the important effect between SAR image and optical images in the non-urban area, λ_e' and λ_e are introduced into $\alpha_{\text{Optical},\ i}$ and $\alpha_{\text{SAR},\ i}$, thus ensuring $\alpha_{\text{SAR},\ i} > \alpha_{\text{Optical},\ i}$, demonstrating the nature that the optical image contributes more to the image classification in the non-urban area than the SAR image. From the above discussion, we get

$$
\begin{cases}
\textit{if } \text{Mask}_i = 1, \lambda_e = 1, \lambda_e' = 0: \\
P(X_i|\omega_j) \\
\quad = \begin{cases} P(X_{\text{SAR},i}|\omega_B)^{\lambda_{\text{SAR},i}+\lambda_e'} \times P(X_{\text{Optical},i}|\omega_B)^{\lambda_{\text{Optical},i}+\lambda_e'} \\ P(X_{\text{SAR},i}|\omega_{B'})^{\lambda_{\text{SAR},i}+\lambda_e} \times P(X_{\text{Optical},i}|\omega_{B'})^{\lambda_{\text{Optical},i}+\lambda_e} \end{cases} \\
\\
\textit{if } \text{Mask}_i \neq 1, \lambda_e = 1, \lambda_e' = 0, ep = 0.00001: \\
P(X_i|\omega_j) \\
\quad = \begin{cases} P(X_{\text{SAR},i}|\omega_B)^{\lambda_{\text{SAR},i}+\frac{1}{ep}} \times P(X_{\text{Optical},i}|\omega_B)^{\lambda_{\text{Optical},i}+\frac{1}{ep}} \\ P(X_{\text{SAR},i}|\omega_{B'})^{\lambda_{\text{SAR},i}+\lambda_e} \times P(X_{\text{Optical},i}|\omega_{B'})^{\lambda_{\text{Optical},i}+\lambda_e'} \end{cases}
\end{cases} \tag{9}
$$

where $\text{Mask}_i = 1$ indicates that the current pixel is in the urban area and $\text{Mask}_i \neq 1$ indicates that the current pixel is in a non-urban area.

If $U(C|X_1, ..., X_n) = \log(L(C|X_1, ..., X_n))$, we introduce (8) into the following energy function [21] as

$$
U(C|X_1, ..., X_n) = \sum_{s=1}^{n} \alpha_s U_{\text{data}}(X_S) + U_{sp}(C) \tag{10}
$$

Then, (9) is used in (10), we get the object functions of MRF with amendment reliability factors for classification as follows:

(1) If the current pixel is in the building area, then the energy function that belongs to the building is

$$
\begin{aligned}
& U_{\text{data}}(X_{\text{Fused}}) + U_{sp}(C) \\
& = -\left\{ \left(\lambda_{\text{SAR},i} + \lambda_e' \right) \log\left(P(X_{\text{SAR},i}|\omega_B) \right) \right. \\
& \left. + \left(\lambda_{\text{Optical},i} + \lambda_e' \right) \log\left(P(X_{\text{Optical},i}|\omega_B) \right) \right\} + U_{sp}(C)
\end{aligned} \tag{11}
$$

(2) If the current pixel is in the building area, then the energy function that belongs to the non-building is

$$
\begin{aligned}
&U_{data}(X_{Fused}) + U_{sp}(C) \\
&= -\{(\lambda_{SAR,i} + \lambda_e)\, \log(P(X_{SAR,i}|\omega_{B'})) \\
&\quad + (\lambda_{Optical,i} + \lambda_e)\, \log(P(X_{Optical,i}|\omega_{B'}))\} + U_{sp}(C)
\end{aligned}
\tag{12}
$$

(3) If the current pixel is in a non-building area, then the energy function that belongs to the building is

$$
\begin{aligned}
&U_{\text{data}}(X_{\text{Fused}}) + U_{sp}(C) \\
&= -\left\{\left(\lambda_{\text{SAR},i} + \frac{1}{ep}\right) \log(P(X_{\text{SAR},i}|\omega_B))\right. \\
&\quad \left. + \left(\lambda_{\text{Optical},i} + \frac{1}{ep}\right) \log(P(X_{\text{Optical},i}|\omega_B))\right\} + U_{sp}(C)
\end{aligned}
\tag{13}
$$

(4) If the current pixel is in a non-building area, then the energy function that belongs to the non-building is

$$
\begin{aligned}
&U_{\text{data}}(X_{\text{Fused}}) + U_{sp}(C) \\
&= -\{(\lambda_{\text{SAR},i} + \lambda_e)\, \log(P(X_{\text{SAR},i}|\omega_{B'})) \\
&\quad + (\lambda_{\text{Optical},i} + \lambda'_e)\, \log(P(X_{\text{Optical},i}|\omega_{B'}))\} + U_{sp}(C)
\end{aligned}
\tag{14}
$$

If the current pixel is in the building area, the energy function of each class is calculated according to (11) and (12), then the label that minimizes the energy function is the final label for the current pixel. If the current pixel is in the non-urban region, the energy function of each class is calculated according to (13) and (14), and the class that minimizes the energy function is the final label for the current pixel.

3 Results

3.1 Study sites, data, and evaluation indexes

Based on Windows 7 system and Matlab2015a as the experimental platform, Sentinel-1 image and Landsat-8 image are used in the experiments. The C-band Level-1 products with an imaging mode of IW in Sentinel-1 are used. The single-view spatial resolution is 5 by 20 m. The Landsat8 satellite launched by NASA on February 11, 2013, carries not only the OLI Land Imager, but also the TIRS Thermal Infrared Sensor. The resolution is 30×30 m and the panchromatic resolution is 15×15 m.

Experiments were chosen in two sites of Xiamen, China, and Neiye, Japan. The size of the image in Xiamen is 762×805. The classification categories of this area are: urban area, vegetation and water area. As shown in Fig. 3, Fig. 3a is a preview image of Sentinel-1, and the gray rectangular region in this figure is used in the experiment. Figure 3b shows an image with a resolution of 2.17 m in Google Earth. It is recommended to use the optical image to mark the Ground Truth map of the Xiamen experimental area. The image size of Neiye is 549×504 (the size is the Sentinel-1 image size, the Landsat-8 image needs to be registered
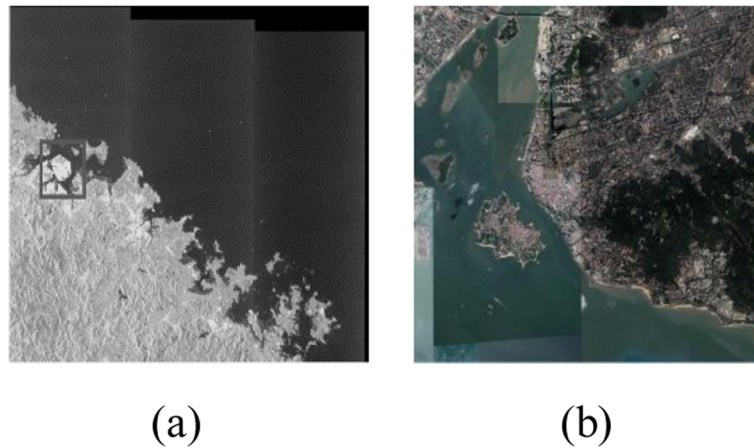
Fig. 3 The site of Xiamen, China. The whole image is provided, containing the site of Xiamen as shown by **a** Sentinel-1 image and **b** Google Earth image

and upsampled to reach this size). The classification categories of the area are urban area, vegetation, farm land, and sand. As shown in Fig. 4, Fig. 4a is a preview image of Sentinel-1, where the gray rectangular region is used in the experiment. Figure 4b is an image with a resolution of 2.17 m in Google Earth. The optical image is also used to mark the Ground Truth map for Neiye. Both study sites have been co-registered with upsampling to 30×30 m.

The accuracy evaluation index is dependent on the confusion matrix. Product's accuracy (PA), user accuracy (UA), overall accuracy (OA), and Kappa coefficients are calculated for quantitative evaluation [16].

3.2 Parameters setting

For SAR images, the grayscale co-occurrence matrix window sizes of SAR images in Xiamen and Neiye are 9 and 33 (according to parameter performance analysis), and the spatial smoothing weights of Markov random fields are 0.01 and 0.001, respectively
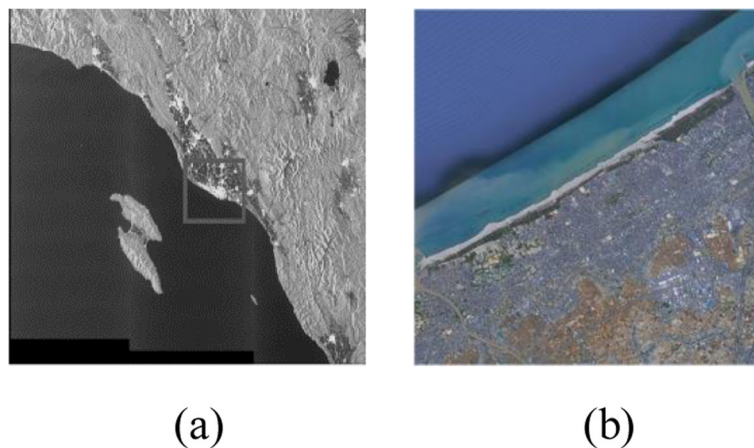

Fig. 4 The site of Neiye, Japan. The whole image is provided, containing the site of Neiye as shown by **a** Sentinel-1 image and **b** Google Earth image

(according to parameter performance analysis). The threshold values of the SAR images were selected to be 0.5 respectively (according to parameter performance analysis). The training sample selection and stopping strategy adopt the method consistent with the comparison algorithms. At each iteration, the samples with top 20% of the classification accuracy are selected as the training data set. In order to reduce the running time of the algorithm and ensure the accuracy of the algorithm, we choose the label updating rate less than 5%; then, the iterations are stopped. For the optical image, bands of 4, 3, and 2 in Landsat-8 are used. The spatial smoothing weights and the training data set selection are as same as those in SAR images. The proposed algorithm is named as MRF-ARF.

The comparison algorithm adopts a pixel-based fusion classification algorithm [22], named as ATWT-EMD, and the parameters are set as follows: the number of layers decomposed by the à trous wavelet transform (ATWT) is 3 and the number of layers decomposed by empirical mode decomposition (EMD) is 3. The spatial smoothing weights of the Markov random field in Xiamen and the inland are 0.01 and 0.001, respectively. At each iteration, the samples with top 20% of the classification accuracy are selected as the training data set. When the label updating rate is less than 5%, then the iteration is stopped.
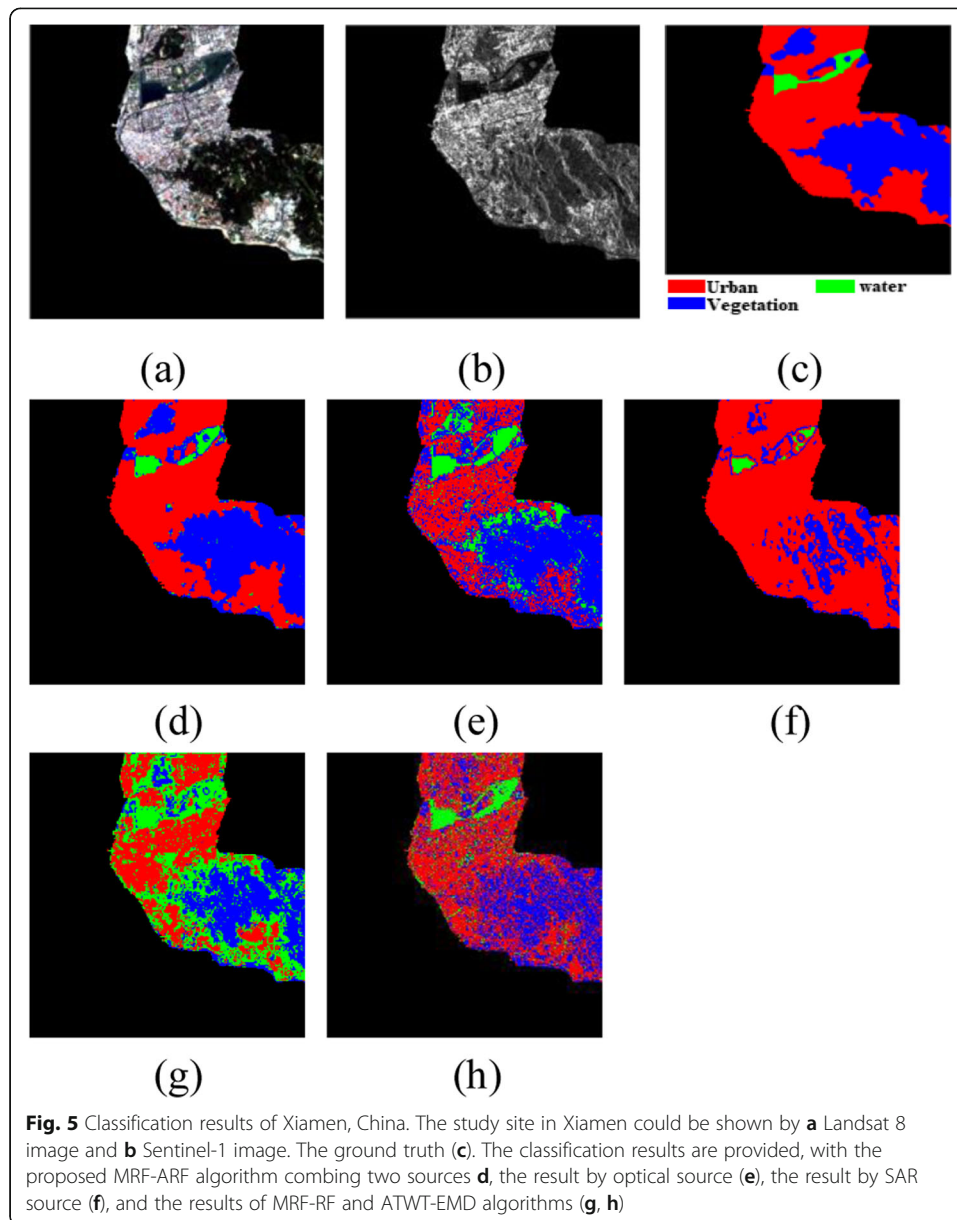
The classification algorithm based on reliability factors without GIS data is adopted [18] named as MRF-RF. The parameters are set as follows: the spatial smoothing weights of the Markov random field in Xiamen and the Neiye are 0.01 and 0.001, respectively. At each iteration, the samples with top 20% of the classification accuracy are selected as the training data set. When the label updating rate is less than 5%, then the iteration is stopped.

3.3 Experimental comparison

The site in Fig. 5 is Xiamen, China. Figure 5a–h are the original Landsat-8 image, the original Sentinel-1 image, Ground truth map, the result of MRF-ARF, the result only on optical image classification, the result only on SAR image classification, the result of MRF-RF, and the result of ATWT-EMD. The ground truth map as shown in Fig. 5c indicates red color for urban, green color for water, and blue color for vegetation. As shown in Fig. 5d, the consistency of MRF-ARF with the data label is the highest, and the urban area is basically consistent with the urban area. Therefore, the extraction results seem to be good. In Fig. 5d, there is a fault in the narrower part of the water. This is because Markov random field has the effect of spatial smoothness, that is, the energy of the data term is not high, resulting into such a fault.

As shown in Fig. 5e, the classification result based solely on the optical image is good in non-urban areas, but some point-like error labels appear in the vegetation and urban regions. This is because the optical image-based classification only uses the spectral characteristics of the optical image, resulting into such problem.

When only the SAR image is used, as shown in Fig. 5f, the classification of the urban area can achieve good result, while those regions lack the detailed texture due to the electromagnetic wave reflection information, misclassification could occur in non-urban areas.

Fig. 5 Classification results of Xiamen, China. The study site in Xiamen could be shown by **a** Landsat 8 image and **b** Sentinel-1 image. The ground truth (**c**). The classification results are provided, with the proposed MRF-ARF algorithm combing two sources **d**, the result by optical source (**e**), the result by SAR source (**f**), and the results of MRF-RF and ATWT-EMD algorithms (**g**, **h**)

For the result of MRF-RF as shown in Fig. 5g, the optical image reliability factor is small because it cannot be excluded in the urban area, that is, the reliability factor in optical image contributes more than that in the SAR image. In the non-urban area, it is also possible that the reliability factor in the SAR image is small, providing more contribution to the objective function than that in the optical image. In other words, MRF-RF method is lack of full utilization of reliability factors. The proposed MRF-ARF algorithm has strong guiding ability due to the existence of urban region marking field. That is, in this area, the advantage of the SAR image is exerted (the SAR image has a good recognition ability for the building), and in the non-urban area, the optical image is fully utilized (using the spectral characteristics of the optical image).

For the result of ATWT-EMD, as shown in Fig. 5h, it can be seen that the ATWT-EMD method has significantly classification labels in the vegetation area. That is to say,

this method also has a guiding effect. It means, in the urban area, the characteristics of the SAR image are more prominent, while in the non-urban area, the spectral characteristics of the optical image are more prominent.

The site in Fig. 6 is Neiye, Japan. Figure 6a–h are the original Landsat-8 image, the original Sentinel-1 image, Ground truth map, the result of MRF-ARF, the result only on optical image classification, the result only on SAR image classification, the result of MRF-RF, and the result of ATWT-EMD. The ground truth map as shown in Fig. 6c indicates red color for urban, green color for sand, pink color for farmland, and blue color for vegetation.

From Fig. 6d, compared with the experimental result in Xiamen area, the classification effect of the urban area in Neiye is slightly worse, which is due to the deviation according to extraction of the urban area. The urban areas may not be so dense, and
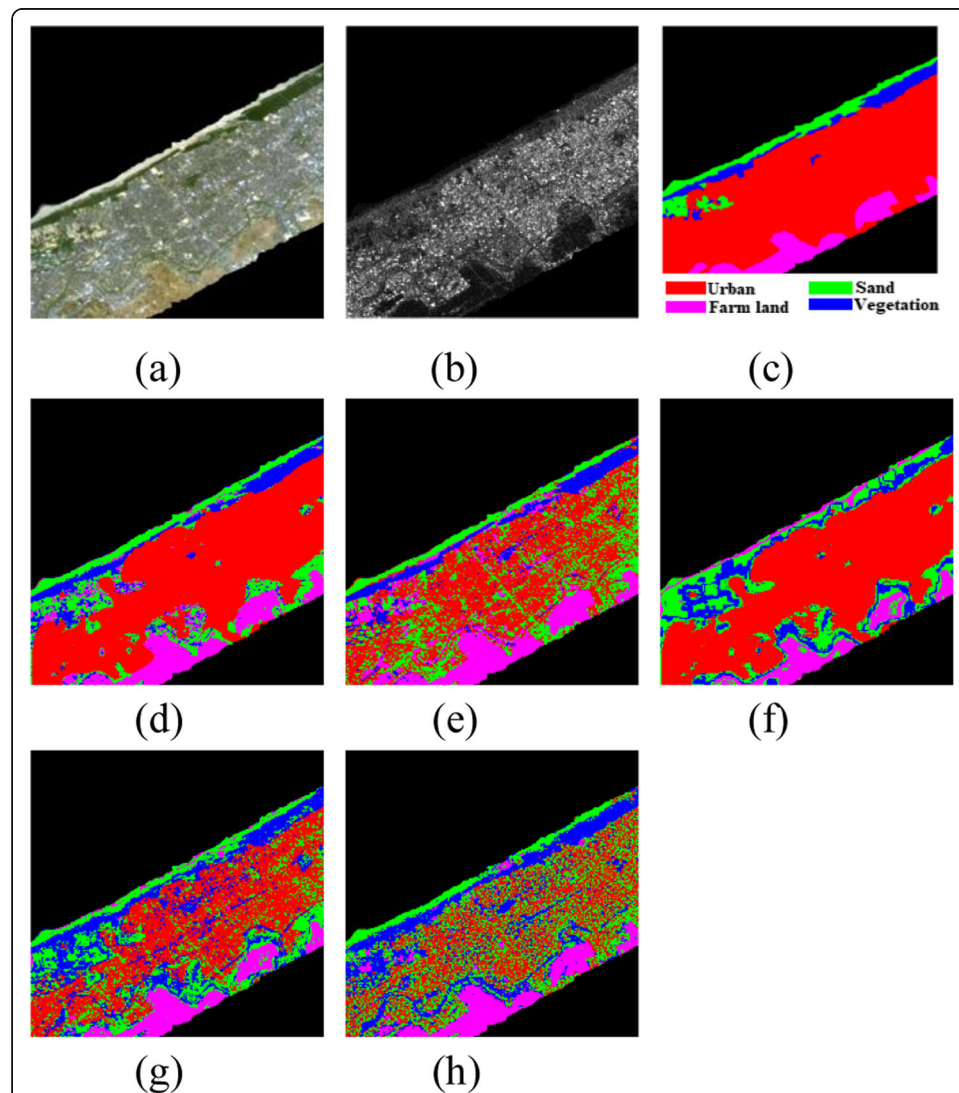

Fig. 6 Classification Results of Neiye, Japan. The study site in Neiye could be shown by **a** Landsat 8 image and **b** Sentinel-1 image. The ground truth (**c**). The classification results are provided, with the proposed MRF-ARF algorithm combing two sources (**d**), the result by optical source (**e**), the result by SAR source (**f**), and the results of MRF-RF and ATWT-EMD algorithms (**g**, **h**)

there are more surrounding vegetation areas. Beside the boundary between urban and farmland, some misclassifications have also appeared. This is because this area belongs to non-urban areas.

When using only optical image for classification as shown in Fig. 6e, the classification of non-urban areas is very good, but the urban area will be misclassified, because this area is often mixed with shadows, black roads, and light gray plates together with vegetation.

When only SAR image classification is used as shown in Fig. 6f, the classification effect of non-urban areas is very poor, indicating that SAR images have less classification ability for areas without detailed information.

For the result of MRF-RF as shown in Fig. 6g, in the urban area, it is possible that the reliability factor of the optical image is smaller than that of the SAR image, showing that the optical image has a greater advantage in the classification, resulting in a misclassification.

For the result of ATWT-EMD as shown in Fig. 6h, there are still some points of misclassification in the urban area, namely, shadows, black roads, and light gray plates caused by vegetation.

To quantitatively assess the classification performance, product's accuracy (PA), user's accuracy (UA), overall accuracy (OA), and Kappa coefficient (KC) are utilized. The comparison of experimental results in Xiamen could be seen in Table 1. The overall accuracy and Kappa coefficient of MRF-RF are the largest, reaching 93.61% and 0.8717 respectively, indicating that the classification performance of the proposed algorithm is superior to comparison algorithms.

In terms of PA, compared with result of the SAR image classification only, the water and vegetation areas of the algorithm have been significantly improved, increasing by 48.04% and 48.60%, respectively. This is because when using only SAR image classification, the SAR image has little detailed information in the non-building area, causing it to less discriminating ability in this area. Compared with the optical image classification only, the classification accuracy of urban and vegetation has been improved by about 38.62% and 9.97%, respectively. The reason for the classification accuracy of urban area is that compared with the optical image classification only, MRF-ARF utilized the amendment reliability factor in SAR image, to distinguish the urban area. Therefore, the classification accuracy of the urban area is improved. The result of PA in MRF-RF

Table 1 Performance comparison of Xiamen, China

Method	Precision	Urban	water	Vegetation	Overall accuracy	Kappa
MRF-ARF	PA	97.87%	85.20%	87.47%	93.61%	0.8717
	UA	92.46%	87.96%	96.40%		
MRF-RF	PA	58.48%	88.81%	62.45%	61.00%	0.4209
	UA	97.99%	8.48%	84.01%		
ATWT-EMD	PA	66.48%	88.65%	60.02%	64.90%	0.3549
	UA	74.38%	29.93%	60.98%		
SAR only	PA	95.55%	37.16%	38.87%	72.75%	0.3938
	UA	70.52%	99.76%	81.07%		
Optical only	PA	59.25%	93.26%	77.50%	67.15%	0.4444
	UA	96.97%	25.77%	56.05%		

and ATWT-EMD algorithms have better performance in water than MRF-RF, and lower in urban and vegetation than MRF-ARF.

In terms of UA, MRF-ARF has best performance in water and vegetation, among three algorithms. MRF-RF has the best result in urban area.

MRF-RF and ATWT-EMD have the overall accuracy of less than 70%. The Kappa coefficient is less than 0.5, indicating MRF-ARF has the best performance in contrast with the other two algorithms, although the other two algorithms have some superiority in PA or UA with certain land covers.

The comparison of experimental results in Neiye could be seen in Table 2. The overall accuracy and Kappa coefficient of MRF-ARF are the largest, reaching 87.86% and 0.7219 respectively, indicating its superiority in classification. Compared with the overall accuracy and Kappa coefficient in Xiamen, these two indexes in Table 2 have decreased, because the land cover in this image has lower contrast and classification is more difficult. It can also be seen that in terms of PA, compared with the SAR image classification algorithm alone, the values by MRF-ARF at sand, vegetation, and farmland have been significantly improved, increasing by 14.35%, 15.88%, and 25.15%, respectively. The reason is that when using only SAR image classification, the SAR image has little detailed information in the non-building area, which shows less discriminative ability. When MRF-ARF is used in the non-building area, the main contribution of amendment reliability factor from the optical image is good, because the optical image has rich spectral characteristics in non-building areas. Thus, the PA of MRF-ARF in sandy, vegetation, and agricultural arable land has been significantly improved. Compared with the optical image classification alone, although the PA of MRF-ARF at vegetation and farmland decreased slightly, the PA of MRF-ARF at urban and sand regions has reached 31.94% and 9.04%, respectively. Compared with the optical image classification alone, MRF-ARF takes advantage of the SAR image, owning good discrimination ability for building cover. As to MRF-RF, PA at urban region is lower than that of MRF-ARF, following with ATWT-EMD, while PA of ATWT-EMD at vegetation and farmland is best.

In terms of UA, MRF-RF is best at building region following by ATWT-EMD, and MRF-ARF. MRF-ARF has best results at sand and vegetation than the other two algorithms.

Table 2 Performance comparison of Neiye, Japan

Method	Precision	Urban	Sand	Vegetation	Farmland	Overall accuracy	Kappa
MRF-ARF	PA	90.67%	74.01%	69.16%	90.56%	87.86%	0.7219
	UA	96.45%	65.65%	65.12%	68.20%		
MRF-RF	PA	54.33%	74.26%	85.20%	81.38%	60.69%	0.3905
	UA	99.67%	22.32%	24.91%	86.23%		
ATWT-EMD	PA	42.38%	62.10%	91.54%	90.59%	52.05%	0.3168
	UA	98.13%	15.11%	26.65%	78.74%		
SAR only	PA	80.25%	59.66%	53.28%	65.41%	75.31%	0.5149
	UA	98.92%	26.16%	29.75%	75.92%		
Optical only	PA	58.73%	64.97%	87.74%	96.07%	64.89%	0.4250
	UA	96.91%	19.68%	59.68%	53.72%		

According to PA and UA at the sites of Xiamen and Neiye, these three algorithms have their own advantages at different land covers. For OA and Kappa coefficient, MRF-ARF shows best results following by MRF-RF and ATWT-EMD. The joint usage of MRF-ARF could provide better result than the results of single-source Landsat 8 or Sentinel-1, demonstrating the superiority of the amendment reliability factors with two sources.
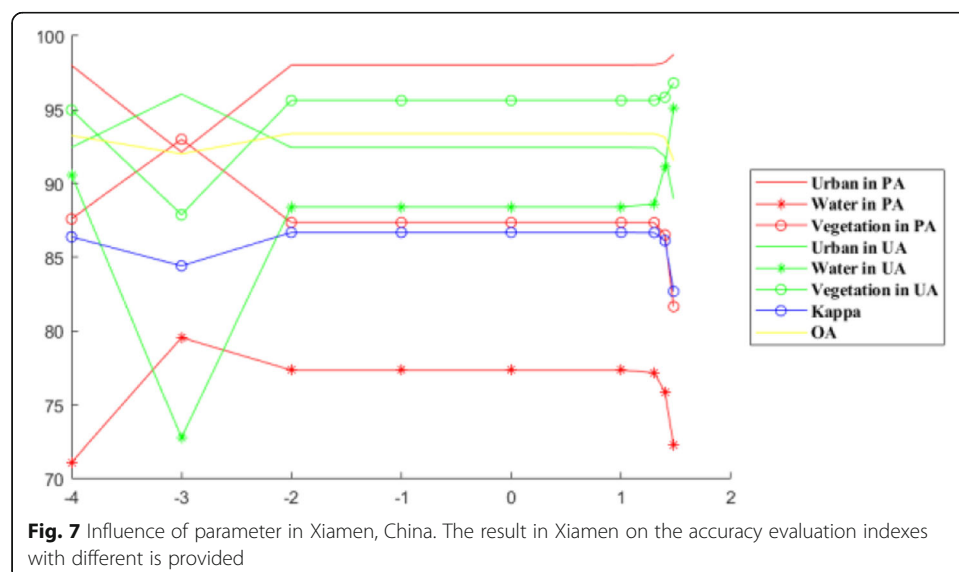
4 Discussions

To quantitatively analyze the sensitivity of parameters in MRF-ARF, the parameter β in regularity term of MRF-ARF, threshold, and grayscale symbiotic matrix window size are assessed according to PA, UA, OA, and Kappa coefficient at different land covers.

4.1 Analysis of parameter β on the performance

Figures 7 and 8 show the experimental results of the influence of parameters β on the classification performance in Xiamen and Neiye, respectively. The abscissa values are the values of different β in logarithm and the vertical coordinate values are the indexes for PA, UA, OA, and Kappa coefficients at different land covers. As shown in Fig. 7, when the SAR image threshold and the gray level symbiotic matrix window size in Xiamen are fixed at 0.68 and 13, respectively, the Kappa coefficient (blue line) and the overall accuracy (yellow line) have good stable performance at [0.01, 0.1, 1, 10], with the abscissa values shown as [− 2, − 1, 0, 1]. When the parameter β is distributed in the interval as [0.01, 10], the classification performance of MRF-ARF is stable. Thus, 0.01 is chosen for the site of Xiamen. From Fig. 8, the Kappa coefficient (blue line) and the overall accuracy (yellow line) are stable at the interval of [0.0001, 10]. Thus, 0.001 is chosen for the site of Neiye.

4.2 Analysis of SAR threshold on the performance

Figures 9 and 10 show experimental results of threshold values on above four indexes in Xiamen and Neiye, respectively. For the site of Xiamen, the threshold values locate

Fig. 7 Influence of parameter in Xiamen, China. The result in Xiamen on the accuracy evaluation indexes with different is provided

Fig. 8 Influence of parameter in Neiye, Japan. The result in Neiye on the accuracy evaluation indexes with different is provided
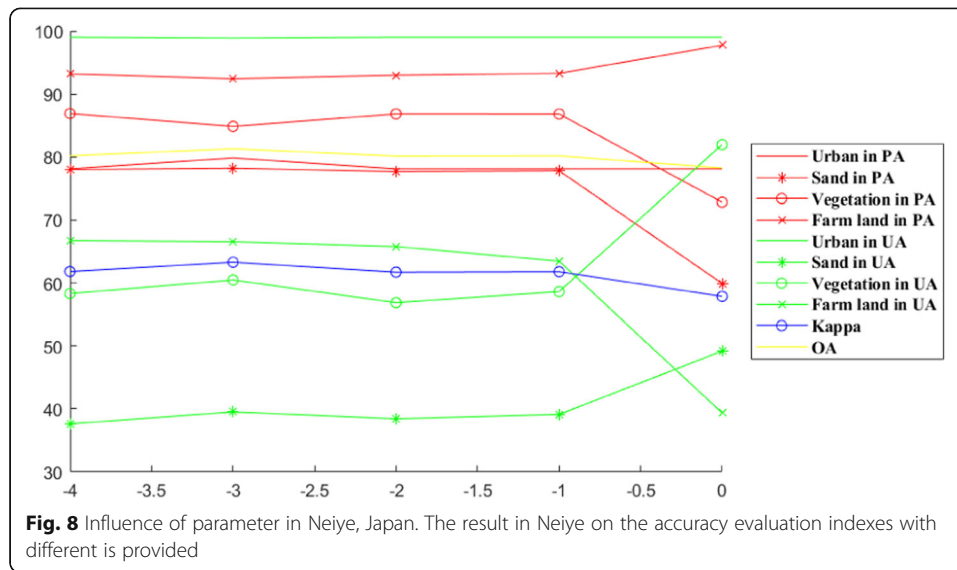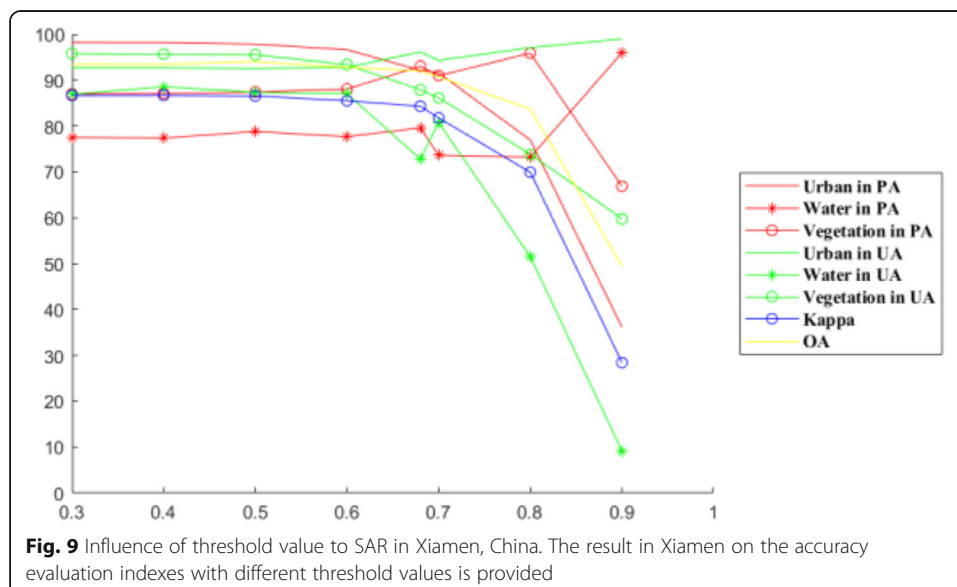
at the interval of [0.3, 0.7]. The Kappa coefficient shown in Fig. 9 for the blue line, and the overall accuracy shown in Fig. 10 for the yellow line, could be stable. When the threshold value of SAR image is greater than 0.7, the performance obviously begins to decrease.

From Fig. 10, the Kappa coefficient (blue line) and overall accuracy (yellow line) could be stable at interval of [0.5, 0.7]. When the threshold value is greater than 0.7, the performance begins to decrease. Thus, 0.6 is chosen for the sites of Xiamen and Neiye in the experiments.

4.3 Analysis of grayscale co-occurrence matrix window size on the performance

Figure 11 shows the experimental results with the window size of grayscale co-occurrence matrix (GLCM) on above four indexes in Xiamen. It could be seen that the
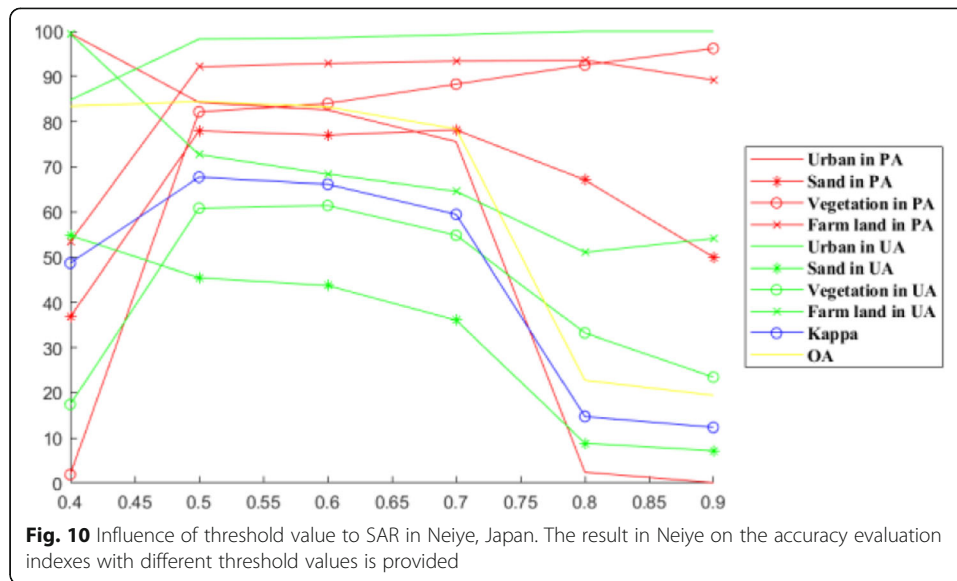

Fig. 9 Influence of threshold value to SAR in Xiamen, China. The result in Xiamen on the accuracy evaluation indexes with different threshold values is provided

Fig. 10 Influence of threshold value to SAR in Neiye, Japan. The result in Neiye on the accuracy evaluation indexes with different threshold values is provided

Kappa index (blue line) and the overall accuracy (yellow line) reach the maximum value when the window size is 9, and the values are 93.70% and 0.8733, respectively. The PA and UA of the building fluctuate greatly, indicating that the window size of GLCM has a great influence on the performance of the classification. The window size of 9 is used for Xiamen.

Figure 12 shows the results of the window size of GLCM on above four indexes in Neiye. The grayscale symbiotic matrix window size range is [4, 49]. It could be seen that the Kappa index (blue line) and the overall accuracy (yellow line) reach the maximum when the window size is 33, and the values are 87.86% and 0.7219, respectively. In fact, if the window size of GLCM is too large, non-edge points could possibly be judged as edge points according to the experiments. If the window size of GLCM is too
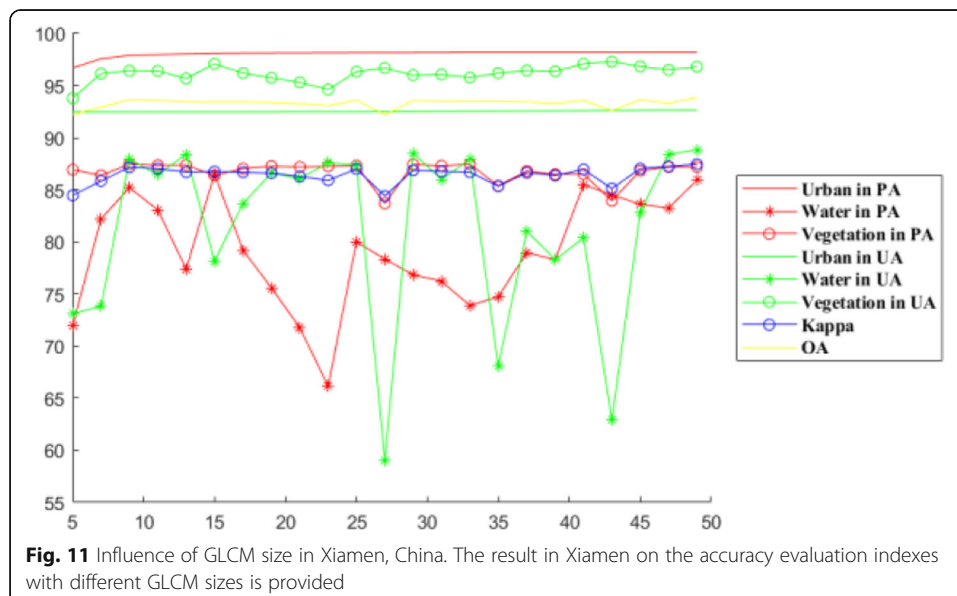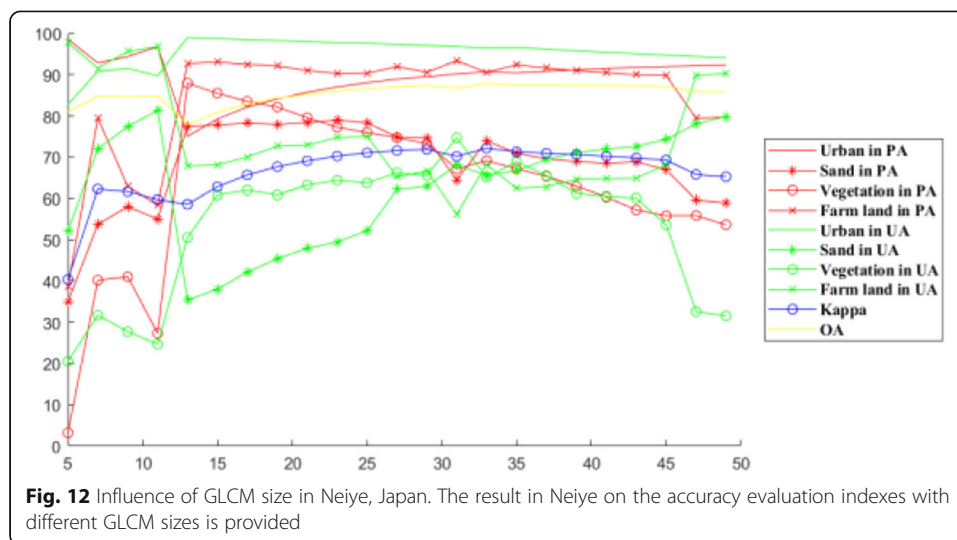

Fig. 11 Influence of GLCM size in Xiamen, China. The result in Xiamen on the accuracy evaluation indexes with different GLCM sizes is provided

Fig. 12 Influence of GLCM size in Neiye, Japan. The result in Neiye on the accuracy evaluation indexes with different GLCM sizes is provided

small, only part of the target could possibly be detected. For Neiye, the window size of GLCM is 33.

5 Conclusion

In this paper, a classification algorithm based on MRF with amendment reliability factors is proposed. Based on the coarse urban label field, the additional controlling factors are involved in reliability factors to construct the amendment reliability factors. The amendment reliability factors could fully utilize the advantage of Landsat 8 and Sentinel-1 to balance the contribution of weight in the data term of MRF. Xiamen and Neiye are chosen as the testing sites. According to the experimental comparison, the proposed MRF-ARF shows the superiority of at least 20% in OA and at least 0.2 in Kappa coefficient in contrast with those of comparison algorithms. Although PA and UA of these algorithms have their own advantages, this paper still provides a way for land cover classification with the joint use of optical and SAR images.

Abbreviations
MRF: Markov random field; SAR: Synthetic aperture radar; MRF-ARF: Amendment reliability factors; SVM: Support vector machine; GLCM: Gray level co-occurrence matrix; OLI: Operational Land Imager; IW: Interferometric Wide Swath; NASA: National Aeronautics and Space Administration; PA: Product's accuracy; UA: User accuracy; OA: Overall accuracy; ATWT: À trous wavelet transform; EMD: Empirical mode decomposition

Author details
[1]Information Science and Technology College, Dalian Maritime University, Dalian 86-116026, Liaoning, China. [2]College of Information Engineering, Dalian University, Dalian 86-116622, Liaoning, China.

References
1. R. Touati, M. Mignotte, M. Dahmane, Multimodal change detection in remote sensing images using an unsupervised pixel pairwise-based Markov random field model. IEEE Trans. Image Process. **29**, 757–767 (2020)
2. Z. Na, Y.Y. Wang, X.T. Li, et al., Subcarrier allocation based simultaneous wireless information and power transfer algorithm in 5G cooperative OFDM communication systems. Physical Communication **29**, 164–170 (2018)
3. M.C. Hansen, T.R. Loveland, A review of large area monitoring of land cover change using Landsat data. Remote Sens. Environ. **122**, 66–74 (2012)
4. J. Xia, Intelligent secure communication for internet of things with statistical channel state information of attacker. IEEE Access **7**, 144481–144488 (2019)
5. G. Liu, Deep learning based channel prediction for edge computing networks towards intelligent connected vehicles. IEEE Access **7**, 114487–114495 (2019)
6. P. Gong, J. Wang, L. Yu, et al., Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM+data. Int. J. Remote Sens. **34**(7), 2607–2654 (2013)
7. Q. Wang, J. Lin, Y. Yuan, Salient band selection for hyperspectral image classification via manifold ranking. IEEE Transaction on Neural Networks and Learning Systems **27**(6), 1279–1288 (2016)
8. Z. Na, J. Lv, M. Zhang, et al., GFDM based wireless powered communication for cooperative relay system. IEEE Access **7**, 50971–50979 (2019)
9. D.P. Shrestha, A. Saepuloh, F.V.D. Meer, Land cover classification in the tropics, solving the problem of cloud covered areas using topographic parameters. Int. J. Appl. Earth Obs. Geoinf. **77**, 84–93 (2019)
10. D.S. Lu, P. Mausel, M. Batistella, E. Moran, Comparison of land-cover classification methods in the Brazilian Amazon Basin. Photogramm. Eng. Remote. Sens. **70**(6), 723–731 (2004)
11. C. Oliver, S. Quegan, *Understanding Synthetic Aperture Radar Images* (Artech House, Boston, 1998)
12. N. Joshi, M. Baumann, A. Ehammer, et al., A review of the application of optical and radar remote sensing data fusion to land use mapping and monitoring. Remote Sens. **8**(70), 1–23 (2016)
13. G.C. Luis, T. Devis, G. Moser, et al., Multimodal classification of remote sensing images: a review and future directions. Proc. IEEE **103**(9), 1–25 (2015)
14. J.S. Xia, J. Chanussot, P.J. Du, X.Y. He, Spectral–spatial classification for hyperspectral data using rotation forests with local feature extraction and Markov random fields. IEEE Trans. Geosci. Remote Sens. **53**(5), 2532–2546 (2015)
15. G. Moser, S.B. Serpico, Combining support vector machines and Markov random fields in an integrated framework for contextual image classification. IEEE Trans. Geosci. Remote Sens. **51**(5), 2734–2752 (2013)
16. I. Hedhli, G. Moser, S.B. Serpico, et al., Classification of multisensor and multiresolution remote sensing images through hierarchical markov random fields. IEEE Geosci. Remote Sens. Lett. **14**(12), 2448–2452 (2017)
17. Y. Tarabalka, M. Fauvel, J. Chanussot, et al., SVM- and MRF-based method for accurate classification of hyperspectral images. IEEE Geosci. Remote Sens. Lett. **7**(4), 736–740 (2010)
18. A. Solberg, T. Taxt, A. Jain, A Markov random field model for classification of multisource satellite imagery. IEEE Trans. Geosci. Remote Sens. **34**(1), 100–113 (1996)
19. B. Waske, S.V.D. Linden, Classifying multilevel imagery from sar and optical sensors by decision fusion. IEEE Trans. Geosci. Remote Sens. **46**(5), 1457–1466 (2008)
20. B. Jeon, D.A. Landgrebe, Classification with spatio-temporal interpixel class dependency contexts. IEEE Trans. Geosci. Remote Sens. **30**(4), 663–672 (1992)
21. C. Zheng, L.G. Wang, X.H. Chen, A hybrid markov random field model with multi-granularity information for semantic segmentation of remote sensing imagery. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **12**(8), 2728–2740 (2019)
22. S. Chen, R. Zhang, H.J. Su, et al., SAR and multispectral image fusion using generalized IHS transform based on à trous wavelet and EMD decompositions. IEEE Journal of Sensors **10**(3), 737–774 (2010)
23. G. Lehureau, M. Campedel, F. Tupin, et al., *Combining SAR and optical features in a SVM classifier for man-made structures detection* (IEEE International Geoscience and Remote Sensing Symposium, Cape Town, 2009), pp. 873–876
24. L.J. Zhao, Y.L. Qin, Y.G. Gao, G. Yan, Detection of high-resolution SAR image building area using GLCM texture analysis. Journal of Remote Sensing **13**(3), 483–490 (2009)
25. Y. Byun, J. Choi, Y. Han, An area-based image fusion scheme for the integration of SAR and optical satellite imagery. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **6**(5), 2212–2220 (2013)
26. A. Baraldi, F. Parmiggiani, An investigation of the textural characteristics associated with gray level co-occurrence matrix statistical parameters. IEEE Trans. Geosci. Remote Sens. **33**(2), 293–304 (1995)

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.