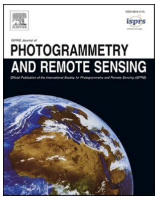
ISPRS Journal of Photogrammetry and Remote Sensing 218 (2024) 246–259

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs



A thin cloud blind correction method coupling a physical model with unsupervised deep learning for remote sensing imagery

Liying Xu [a], Huifang Li [a,*], Huanfeng Shen [a], Chi Zhang [b], Liangpei Zhang [c]

[a] School of Resource and Environmental Sciences, Wuhan University, Wuhan 430079, China
[b] Guangzhou Urban Planning & Design Survey Research Institute, Guangzhou 510800, China
[c] The State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan 430079, China

ARTICLE INFO

Keywords:
Thin cloud correction
Physical model
Unsupervised learning
Feature separation network
Remote sensing images

ABSTRACT

Thin cloud disturbs the observation of optical sensors, thus reducing the quality of optical remote sensing images and limiting the subsequent applications. However, the reliance of the existing thin cloud correction methods on the assistance of in-situ parameters, prior assumptions, massive paired data, or special bands severely limits their generalization. Moreover, due to the inadequate consideration of cloud characteristics, these methods struggle to obtain accurate results with complex degradations. To address the above two problems, a thin cloud blind correction (TC-BC) method coupling a cloudy image imaging model and a feature separation network (FSNet) module is proposed in this paper, based on an unsupervised self-training framework. Specifically, the FSNet module takes the independence and obscure boundary characteristics of the cloud into account to improve the correction accuracy with complex degradations. The FSNet module consists of an information interaction structure for exchanging the complementary features between cloud and ground, and a spatially adaptive structure for promoting the learning of the thin cloud distribution. Thin cloud correction experiments were conducted on an unpaired blind correction dataset (UBCSet) and the proposed TC-BC method was compared with three traditional methods. The visual results suggest that the proposed method shows obvious advantages in information recovery for thin cloud cover regions, and shows a superior global consistency between cloudy regions and clear regions. The TC-BC method also achieves the highest peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM). The FSNet module in the TC-BC method is also proven to be effective. The FSNet module can achieve a superior precision when compared with five other deep learning networks in cloud-ground separation performance. Finally, extra experimental results show that the TC-BC method can be applied to different cloud correction scenarios with varied cloud coverage, surface types, and image scales, demonstrating its generalizability. Code: https://github.com/Liying-Xu/TCBC.

1. Introduction

The atmosphere is the place where radiative transfer mainly occurs, and atmospheric conditions inevitably affect the observation of the Earth's surface from satellite platforms. Thin clouds, as one of the most common atmospheric phenomena, feature a thin optical thickness and high transparency, which allows ground features to be seen through them. Thin cloud include, but are not limited to, cirrus, stratiform, and partial cumulus clouds. In optical remote sensing images, thin clouds appear flat and lack clear boundaries, and can seriously degrade the data quality and thus limit subsequent applications, such as object recognition, target detection, and natural resource monitoring. This implies that thin cloud correction is a necessary preprocessing procedure. Over the past few decades, several thin cloud correction approaches have been proposed for use with optical remote sensing images, especially for the visible and near-infrared (VNIR) bands. The proposed thin cloud blind correction (TC-BC) method is aimed at thin cloud correction on coastal and VNIR bands. The existing thin cloud correction approaches can be grouped into three categories: 1) radiative transfer equation (RTE)-based methods; 2) data-based methods; and 3) hybrid methods, as shown in Fig. 1.

The classic RTE expresses the top-of-atmosphere (TOA) radiance for a flat, Lambertian surface under a horizontally homogeneous atmosphere (Gao et al., 1998), and involves many in-situ parameters about

---

* Corresponding author.
E-mail address: huifangli@whu.edu.cn (H. Li).

https://doi.org/10.1016/j.isprsjprs.2024.09.008
Received 31 July 2023; Received in revised form 8 March 2024; Accepted 11 September 2024

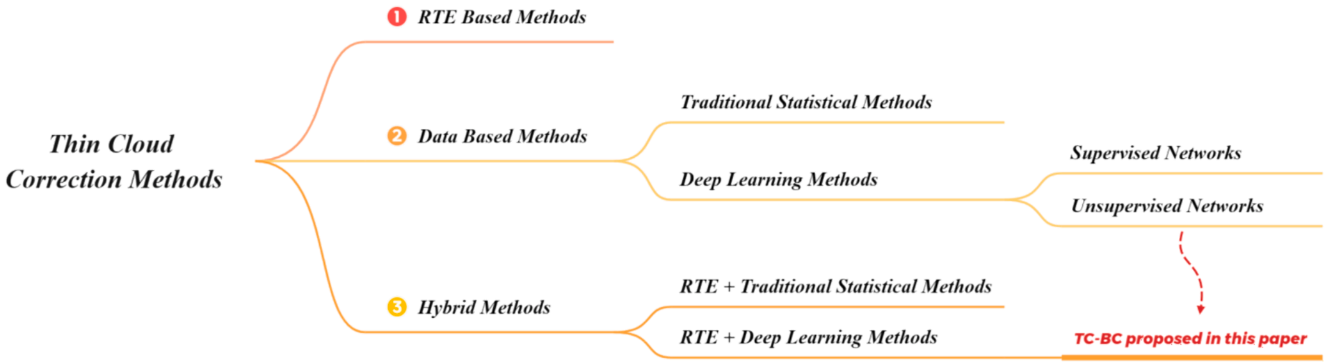


Fig. 1. The classification process of the thin cloud correction methods.

the atmosphere and the radiance. However, it is not practical to solve the RTE numerically for operational atmospheric correction. Furthermore, the basic RTE is not sufficient to solve spatially heterogeneous problems, such as aerosols and thin clouds. As a result, some researchers have tried to estimate the spatial distribution and optical depth of aerosols or thin clouds to retrieve the land reflectance by considering the spatial and spectral relationship between the clear and cloudy regions (Richter, 1996; Liang et al., 2001). MODTRAN (Guanter et al., 2009) is a commonly used model for building look-up tables (LUTs) to obtain the path radiance, transmittance, and other parameters in the RTE. Although this is an effective way to raise the calculation efficiency, the approximation can limit the accuracy. Furthermore, the inherent configurations of some specific atmospheric profiles and aerosol models can also interfere with the understanding of the true state of the atmosphere. As a result, the RTE-based methods usually fail to deal with the thin cloud problem in remote sensing images, as thin clouds are usually heterogeneous in both the horizontal and vertical dimensions.

The data-based methods are independent of in-situ parameters, but are usually reliant on statistical information derived from the images. According to the characteristics of thin clouds, several representative methods for correcting thin clouds have been developed. In the spatial domain, dark channel prior based methods (Chavez, 1988; Kaiming et al., 2009; Makarau et al., 2014; Shen et al., 2020) and color attenuation prior based methods (Zhu et al., 2015) have been developed based on the high reflectance of cloud. In the spectral domain (Ahmad et al., 2014), haze-optimized transform based methods (Zhang et al., 2002; He et al., 2010; Chen et al., 2016), tasseled cap transformation based methods (Crist and Cicone, 1984), and independent component analysis based methods (Shen et al., 2015) have been developed based on the correlation of cloud cover in different bands. In the frequency domain, homomorphic filtering based methods (Shen et al., 2014), wavelet transform based methods (Yong et al., 2002) and Gaussian low-pass filtering based methods (Liu et al., 2014) have been developed based on the spatial smoothness of cloud. Benefiting from the simplicity and effectiveness, these kinds of methods have been widely used. However, there are two deficiencies in these methods. Firstly, the traditional statistical modeling methods usually simplify the relationship between cloudy regions and clear regions as a linear relationship, whereas the actual relationship is often nonlinear. Secondly, when the prior hypothesis is invalid, the empirical models cannot remove the cloud completely.

As a research hotspot in recent years, deep learning has been demonstrated to be a very powerful tool in the statistical information mining of remote sensing images. In the field of thin cloud correction, deep learning networks have a strong advantage in terms of complex nonlinear feature learning and can obtain better results than the traditional methods. Scholars have studied convolutional neural networks (CNNs)and generative adversarial networks (GANs) and have proposed networks such as U-Net (Ronneberger et al., 2015), the residual neural network (ResNet) (He et al., 2016), and conditional generative adversarial networks (CGANs)(Mirza and Osindero, 2014). These networks and their variants have been widely used in thin cloud correction tasks(Jiang and Lu, 2018; Qin et al., 2018; Enomoto et al., 2017; Grohnfeldt et al., 2018; Toizumi et al., 2019). Moreover, some networks have been proposed specifically to remove thin clouds, such as the residual symmetrical concatenation network (RSCNet) (Li et al., 2019), and the spatial attention generative adversarial network (SPAGAN) model (Wang et al., 2019). However, as with the conventional end-to-end learning methods, these methods are supervised and need a lot of paired images of cloudy/clear data, which are not readily accessible. This is because the system of optical remote sensing image processing is complex (Povey and Grainger, 2015). The complexity lies in the uncertain interactions among the subsystems. Specifically, the random noise produced by the vibration of the satellite platform, the unknown environmental parameters that come from the reflection of the ground, and the scattering that occurs in the atmosphere all introduce nonlinear errors to the system. Accordingly, remote sensing images are unique under the same spatial and temporal conditions. As a result, there are no absolutely “cloudy/clear” paired images in the real world, which limits the effectiveness of these supervised methods. With the advancement of generative adversarial techniques, unsupervised GANs can achieve pixel-level estimation, which effectively simulates the influence of nonlinear and random errors. The cycle-consistent adversarial network (CycleGAN) (Zhu et al., 2017) makes it possible to train an unsupervised thin cloud correction network (Singh and Komodakis, 2018). Faced with the complex imaging process, unlike the update of supervised network parameters guided by a single data sample, the update of CycleGAN-based network parameters is conducted by back-propagation of a discriminative module, which considers all the images in the dataset. As a result, CycleGAN-based networks can effectively recover more realistic ground information from true cloudy images. However, these methods do not consider the random distribution of clouds, which results in low stability when dealing with complex cloudy images. In summary, the high cost of massive paired dataset construction and the low correction accuracy of the unsupervised algorithms limit the engineering value of the existing deep learning methods.

As mentioned above, the RTE-based methods focus on quantitative description and can accurately correct atmospheric degradation, but they require a lot of auxiliary parameters. The data-based methods are easily accessible, but the overly simple qualitative statistics result in miscorrection. The third type of method is the hybrid methods, which combine the advantages of the other two types of methods by simplifying the transfer model and coupling this with the statistical information in the images. However, to date, only a few hybrid methods have been developed. For example, Li et al. (2012) proposed a constraint based on band correlation to enhance the spatial details and improve thin cloud correction; Lv et al. (2016) used the linear statistical information derived from cloudy images to solve the unknown variables in the RTE empirically; and Zhang et al. (2021) demonstrated the feasibility of the use of scattering theory in describing the band variations of

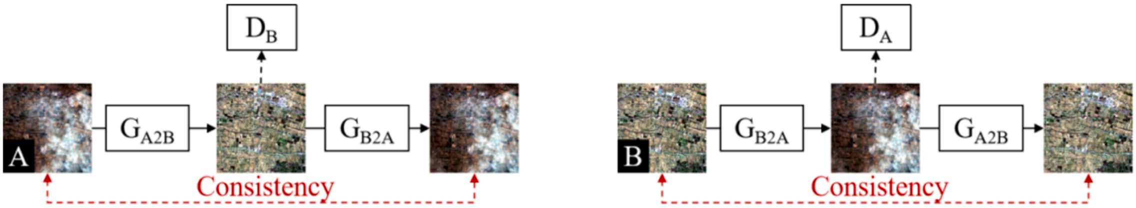


Fig. 2. The framework of CloudGAN.

thin cloud and then developed a cirrus correction method combining the statistical relationship between the coastal band and blue band. Cloud-related physical mechanisms have also been incorporated into deep learning networks to improve the accuracy of the correction, in methods such as Dehaze Net (Cai et al., 2016) and CR-GAN-PM (Li et al., 2020). These methods effectively improve the accuracy of cloud correction and have promoted the development of cloud correction of remote sensing images from qualitative processing to quantitative correction. However, this type of method still has a strong dependence on special bands or a huge paired training dataset. To sum up, there is room for the further development of hybrid methods.

Motivated by the above discussion, it is an important and significant challenge (Reichstein et al., 2019) for thin cloud correction studies to develop a high-precision method requiring less prior knowledge or training data. The unsupervised deep learning methods have strong applicability and can quantitatively learn the complex nonlinear correlation between images without auxiliary conditions. If the scattering law and imaging mechanism of the cloud can be taken into account in the unsupervised correction process, it should be possible to achieve a high-fidelity result without the prerequisites required in the afore-mentioned methods. Therefore, the TC-BC method coupling a physical model with generative adversarial learning is proposed in this paper. The TC-BC method achieves thin cloud correction by separating the cloud information and ground information in the cloudy images. The proposed approach is composed of three key components. Firstly, a self-training thin cloud correction framework is presented by integrating the RTE and a deep learning network to enhance the robustness and generalizability of the TC-BC method. Secondly, a feature separation network (FSNet) module in the TC-BC method is designed to mine and use the spectral and spatial characteristics of the imagery, to achieve a high correction accuracy. The FSNet module can separate mixed cloud-ground information completely and identify the boundaries of thin clouds accurately. Thirdly, a small unpaired blind correction dataset (UBCSet) containing a thin cloudy set, a ground set, and a cloud set was built to train the TC-BC method. The UBCSet considers atmospheric scattering and the spectral correlation of clouds to improve the training speed of the TC-BC method and the correction ability in complex cloudy regions.

The novelty and contributions of the TC-BC method can be summarized as follows:

1) The TC-BC method is a general algorithm that needs less prior information.
2) The TC-BC method can remove and detect thin clouds simultaneously, with a high degree of accuracy.
3) The training of the TC-BC method is based on an unpaired dataset, so that the cost of dataset construction is reduced.
4) The FSNet module shows a better performance than the existing networks and could likely be applied in similar domains beyond cloud correction, such as haze removal, shadow removal, and rain removal.

The rest of this paper is organized as follows. Section 2 introduces the related works. Section 3 explains the proposed method for thin cloud correction in detail. Section 4 introduces the datasets used in this study. Section 5 presents the experimental analysis and discussion, including the results obtained by the TC-BC method and the comparative methods. Finally, Section 6 presents our conclusions.

2. Related works

This section focuses on the imaging model for cloudy images and the unsupervised frameworks based on CycleGAN.

2.1. Imaging model for cloudy images

In order to eliminate the error introduced by the optical remote sensing imaging system, the experimental data used in this study were top-of-atmosphere (TOA) reflectance dat. In a cloudy region, the imaging model (Kneizys et al., 1980) can be expressed as shown in Eq.(1):

$$I^i(x,y) = I_c^i(x,y) + T{\downarrow}T{\uparrow}\frac{I_g^i(x,y)}{1 - SI_g^i(x,y)} \tag{1}$$

where $I_c^i(x,y)$ is the cloud component of the TOA value at pixel (x,y) in band i, $T{\downarrow}$(respectively $T{\uparrow}$) is the total transmission of the atmosphere on the path between the sun and the surface (respectively between the surface and the sensor), $I_g^i(x,y)$ is the land surface component of the TOA value at pixel (x,y) in band i, and S is the spherical albedo of the atmosphere. The term of S_c is closer to 0, and $T{\downarrow}T{\uparrow}$ is typically greater than 0.9. Eq. (1) can thus be simplified to an additive model and used in the TC-BC method:

$$I^i(x,y) = I_c^i(x,y) + I_g^i(x,y) \tag{2}$$

2.2. CycleGAN

CycleGAN-based networks are the prevailing methods in unsupervised training, and are made up of two generative modules (i.e.,G_{A2B} and G_{B2A}) and two discriminative modules (i.e.,D_A and D_B). In the earliest stage, CycleGAN (Zhu et al., 2017) was applied in the image-to-image translation task to transfer the style of two datasets (i.e., A and B). Singh and Komodakis (2018) proposed CloudGAN based on CycleGAN, which learns the mapping between the cloudy images (style A) and clear images (style B) to correct the thin cloud in the cloudy images. The unsupervised training process is implemented by cyclic consistency constraints, as shown in Fig. 2. Compared with the dataset of the traditional strongly supervised networks, the unpaired dataset of CloudGAN is easier to construct.

3. Methodology

In this section, the overall architecture and key structure of the proposed TC-BC method are introduced. Specifically, the self-training framework, FSNet module, and objective function are described in detail.

3.1. Self-training framework

It is a general consensus that massive data are a necessary

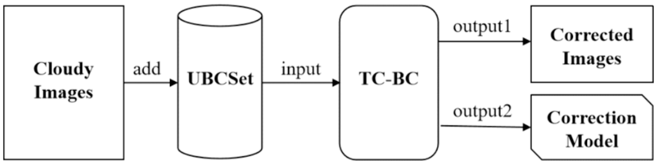


Fig. 3. The idea of "self-training". The unpaired blind correction dataset (UBCSet) is the dataset, and the thin cloud blind correction (TC-BC) method is the method proposed in this paper.

requirement for deep learning algorithms. The existing algorithms always adopt the idea of "training first, testing later", which means that the network is first trained on a massive dataset to obtain a generalized model. This model is then applied to different thin cloud correction scenarios during the testing process. As a result, the effect of thin cloud correction is dependent on the quantity and quality of the training data, but has no relationship with the test data. Moreover, there are still two problems in the construction of the training dataset. Firstly, there are no "cloudy/clear" paired remote sensing images under the same spatial and temporal conditions. Secondly, because of the complex cloud and surface cover, it is difficult to build a massive dataset containing all the possible ground objects.

Based on the above analysis, we introduce the concept of "self-training" and aim to achieve a superior correction effect in cloudy images. As shown in Fig. 3, the cloudy images are added to the UBCSet as the input and participate in the whole training process. Because of the coupling of a deep learning network and a physical model, the data utilization and information mining capabilities of the TC-BC method are improved. The TC-BC method can not only be trained without supervision, but can also obtain the thin cloud corrected images and correction model simultaneously. The corrected images are the optimal cloud removal result for the original cloudy images.

The thin cloudy image input into the self-training framework is separated into cloud and ground components by the FSNet module. For the image processing task, the cloud component is the thin cloud detection result for the cloudy image, and the ground component is the corrected result. These two components are then reconstructed as a new cloudy image by the imaging model shown in Eq.(2). The TC-BC method needs to minimize the difference between the reconstructed thin cloudy image and the input image to achieve the self-training process, as shown in Fig. 4. At the same time, in order to ensure the calculability of the training process, a symmetrical structure is introduced, where the cloud reference image and ground reference image are the input. The two FSNet modules in the self-training framework share the same training weights. In addition, a Markov discriminator is introduced to guide the parameter update of the FSNet module during training. Unlike the input of the Markov discriminator in CycleGAN, the combination of the cloudy image and its ground component is used as the input of this discriminator, which ensures the fidelity of the surface information after correction.

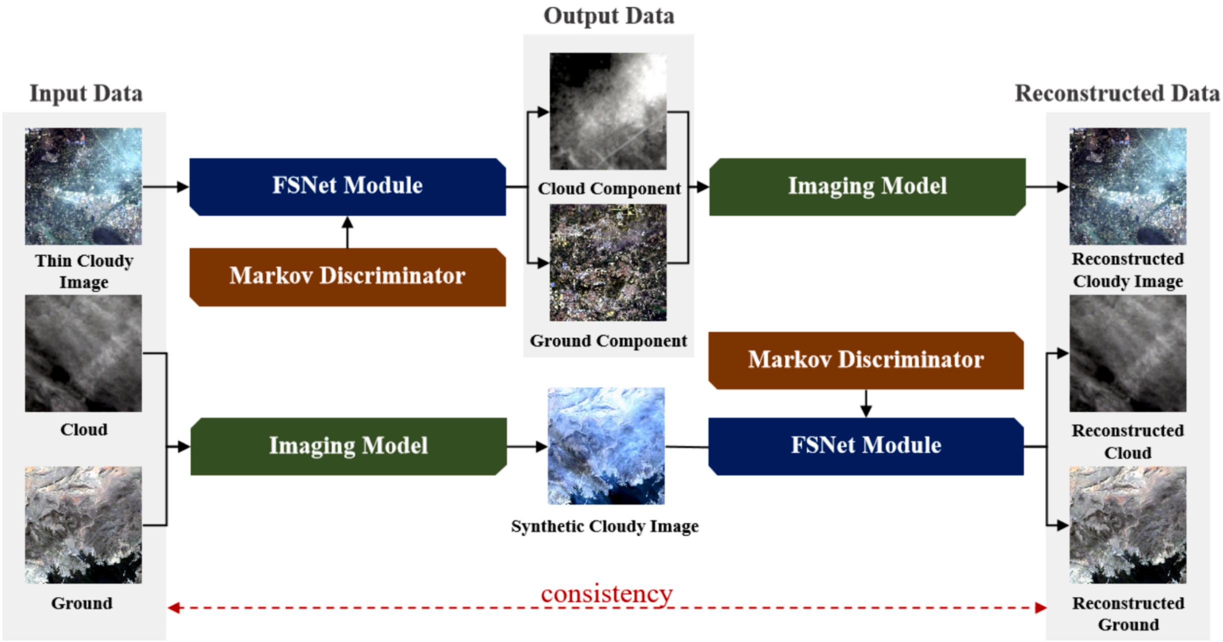


Fig. 4. The self-training framework of the TC-BC method.

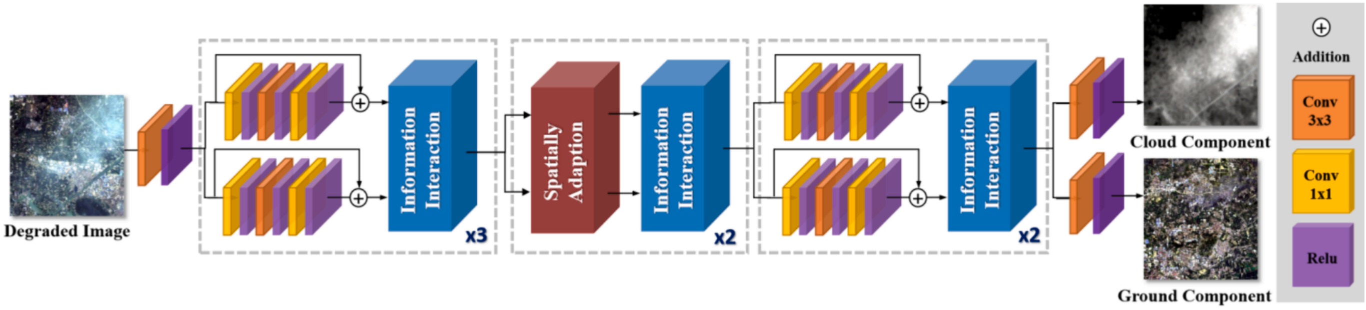


Fig. 5. Details of the proposed FSNet module.

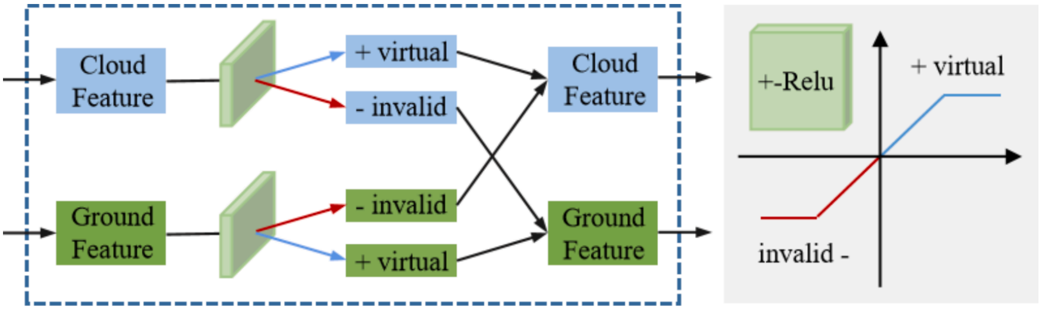


Fig. 6. The architecture of the information interaction structure.

3.2. FSNet module

Deep learning methods can remove most of the thin cloud in different scenarios, due to their strong nonlinear fitting capabilities. However, the correction results for some complex thin cloud covered regions where cloud information and ground information are mixed will not be perfect. While the traditional methods focus on the enhancement of the ground information, they overlook the cloud characteristics, which are important for the correction task, which makes it difficult to remove complex thin clouds completely. Therefore, considering both the cloud and ground information, the FSNet module is proposed to achieve high-accuracy thin cloud correction.

The FSNet module is a two-stream network, as shown in Fig. 5, the input of which is the thin cloud contaminated image. The input image is first encoded globally by a convolutional structure. The encoded features are then fed into the two streams with the same structure, i.e., the cloud stream and the ground stream. Finally, through convolutional decoding, the output of the cloud stream is the thin cloud detection result, and the output of the ground stream is the thin cloud correction result. The FSNet module contains two key structures—the information interaction structure and the spatially adaptive structure—which progressively capture the high-level semantics of the images. The information interaction structure is focused on the separation of the mixed cloud-ground information in the vertical space, while the spatially adaptive structure is focused on the identification of cloudy and clear regions in the horizontal space. The two structures are described in detail as follows.

3.2.1. Information interaction structure

In a cloudy region, the radiative characteristics of the cloud will be mixed in a complex manner with the radiative characteristics of the ground, because of the diversity of cloud shape, the high complexity of the land cover, and the large coverage of remote sensing imagery. According to the imaging model for cloudy images, all the information can be divided into either cloud information or ground information. The activation function layer determines the activation and inhibition of the features by retaining the positive signals and discarding the negative signals. Therefore, in the activation function layer, the features rejected by the cloud stream belong to the ground stream, and vice versa.

The information interaction structure is shown in Fig. 6. The global feature coding from the cloudy image is input into both the cloud stream and ground stream. An activation function named the positive–negative rectified linear unit (ReLU) (Hu and Guo, 2021) is introduced to continuously exchange the invalid signals from the cloud stream and ground stream in the coding process, to blindly separate these two features.

3.2.2. Spatially adaptive structure

Thin clouds are almost transparent and are largely distributed uniformly in the atmosphere. Meanwhile, a convolutional layer always codes the features in a local range. Although the receptive field of a CNN increases with the increase of the convolutional layers, the CNN will still lack the ability to learn global contextual information. Therefore, it is difficult to detect the boundaries of thin clouds only through convolutional layers. In order to better detect cloudy regions, a spatially adaptive structure is introduced, as shown in Fig. 7.

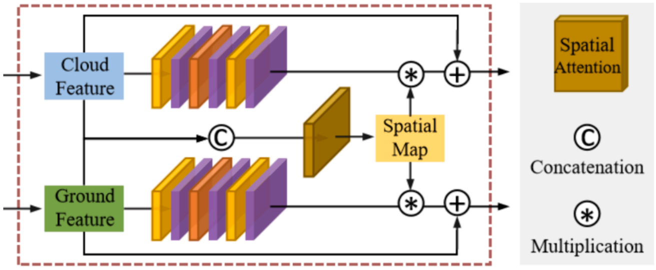


Fig. 7. The architecture of the spatially adaptive structure.

By weighting the features with the spatial map in the spatially adaptive structure, the FSNet module can improve the separation ability for cloud-ground information. The spatial map, which is produced by the attention block, can describe the uneven thin cloud distribution and intensity in space, as shown in Fig. 8.

Since the initialize recurrent neural network (IRNN) (Bell et al., 2016) pays more attention to the sequence of the input in nonlinear feature learning, it has advantages in aggregating spatial context information, compared to a CNN. Based on this ability, a four-direction sensing two-round spatial IRNN is introduced to learn the spatial map. The concatenation of the cloud and ground features is the input of the attention structure. A 1 × 1 convolutional layer is then applied to gain four-direction (left, right, up, and down) spatial context features, and a convolutional block is applied to gain four maps of the weights representing the learning process of the input features. The four maps of the weights are then multiplied with the spatial context features along the different directions in an element-wise manner. By repeating this multiplicative operation, the global spatial map is generated. In addition, to better preserve the edge features, the padding of this structure is reflection padding.

3.3. Objective function

An alternating iterative optimization approach is employed to train the generator and discriminator separately. By maximizing the discriminative loss in the Markov discriminator and minimizing the generative loss in the FSNet module, the TC-BC method can correct cloudy images.

The generative loss consists of consistency loss L_c, identity loss L_i, and adversarial loss L_a.

$$L_G = \omega_1 L_c + \omega_2 L_i + \omega_3 L_a \tag{3}$$

where ω represents the weighting coefficient and the values of $\omega_1, \omega_2, \omega_3$ are 10,5 and 1, respectively.

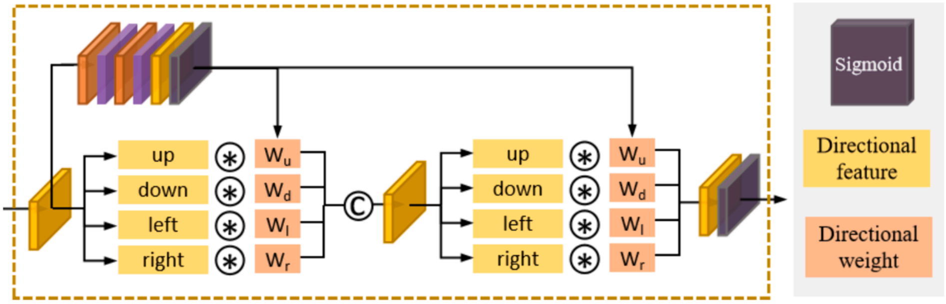


Fig. 8. The architecture of the attention structure.

The consistency loss is a key component of the TC-BC method that keeps the consistency between the reconstructed image and the input image at the pixel level. It accelerates the convergence rate and improves the robustness of the optimization process. In addition, the consistency loss is designed with the image features of the frequency domain and can be expressed as shown in Eq.(4).

$$L_c = \omega_4 \|I - I''\|_1 + \left\|I''_g - I_g\right\|_1 + \|I''_c - I_c\|_1 \tag{4}$$

where the value of ω_4 is 0.5,I represents the input cloudy image, I_c represents the input cloud reference, I_g represents the input ground reference, $\cdot''$ represents the corresponding reconstructed image, and $\|\cdot\|_1$ represents L_1 regulation.

The identity loss is designed to enhance the ability of the FSNet module to capture the cloud and ground features separately from cloudy images, which can be expressed as follows:

$$L_i = \left\|I^*_g - I_g\right\|_1 + \|I^*_c - I_c\|_1 \tag{5}$$

where $\cdot^*$ represents the feature separation result of the FSNet module in the TC-BC method.

The adversarial loss is used to make the distribution characteristics of the thin cloud correction results consistent with those of clear images.

$$L_a = \varepsilon_{I \rightarrow P_{data}(I)}[\log D(I, I'_g)] \tag{6}$$

The discriminative loss can be expressed as follows:

$$L_D = \varepsilon_{I_g \rightarrow P_{data}(I_g)}[\log D(I', I_g)] + \varepsilon_{I'_g \rightarrow P_{data}(I_g)}[\log(1 - D(I, I'_g))] \tag{7}$$

4. UBCSet dataset

In order to compare the thin cloud removal effect of different methods, the unpaired blind correction dataset (UBCSet) was constructed to train the TC-BC method. The target cloudy images can be included in the UBCSet. The UBCSet not only guides the training of the TC-BC method, but can also be corrected concurrently.

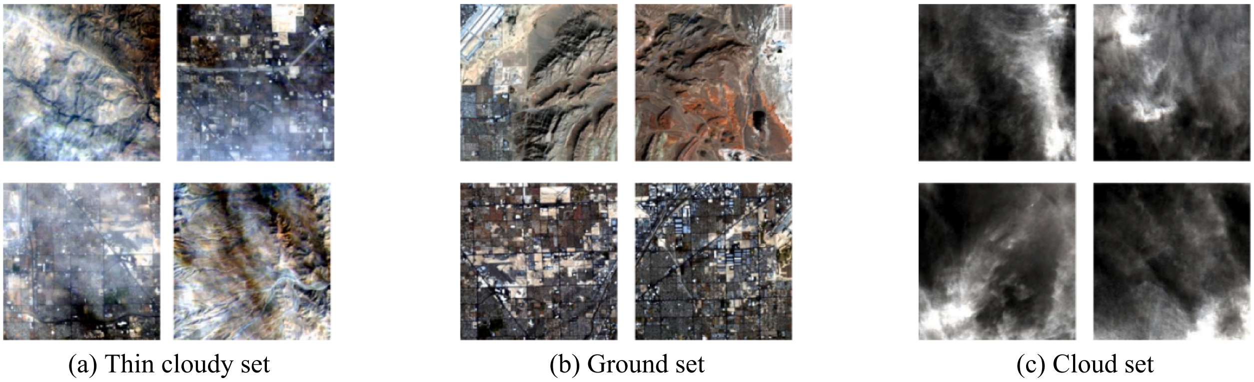
4.1. Construction of the UBCSet dataset

The Landsat-8/9 images acquired by the Operational Land Imager (OLI) instrument with high quality and wide coverage have been widely used, and are available from the United States Geological Survey. In this study, 30 remote sensing images were collected globally between 2011 and 2022, featuring various land-cover types, such as forests, bare soil, and urban areas. Because of their wide use in remote sensing application research, the coastal and VNIR bands were selected as the study objects. The geographic positions and acquisition times of the original images are provided in the Appendix A.

The UBCSet is made up of three sets: a thin cloudy set, a ground set, and a cloud set, as shown in Fig. 9. After radiometric calibration, the original images were clipped into patches with the size of 5 × 256 × 256 pixels. The patches in the thin cloudy set and ground set were picked from cloudy regions and clear regions of the clipped patches directly. The other patches with other degradation influences, such as thick clouds or cloud shadows, were discarded. The cloudy patches degraded by thin clouds were selected for correction. The clear patches were used to provide ground reference information for training the TC-BC method. There are 2580 images in the thin cloudy set and 2354 images in the ground set. The cloud set contains 977 cloud images, each with five bands. These cloud images were constructed from the cirrus band via the scattering model (Zhang et al., 2021), which is expressed by Eq.(8). The patches of the cloud set provide the cloud reference information.

$$I^i_c = \left(\frac{\lambda_{cirrus}}{\lambda_i}\right)^{\gamma} \cdot I^{cirrus}_c \tag{8}$$

where the subscript *cirrus* denotes the cirrus band of the multispectral image;λ represents the wavelength; and γ is a scattering-related parameter, whose value can be calculated from the degraded images.



(a) Thin cloudy set (b) Ground set (c) Cloud set

Fig. 9. Sample patches of the UBCSet.

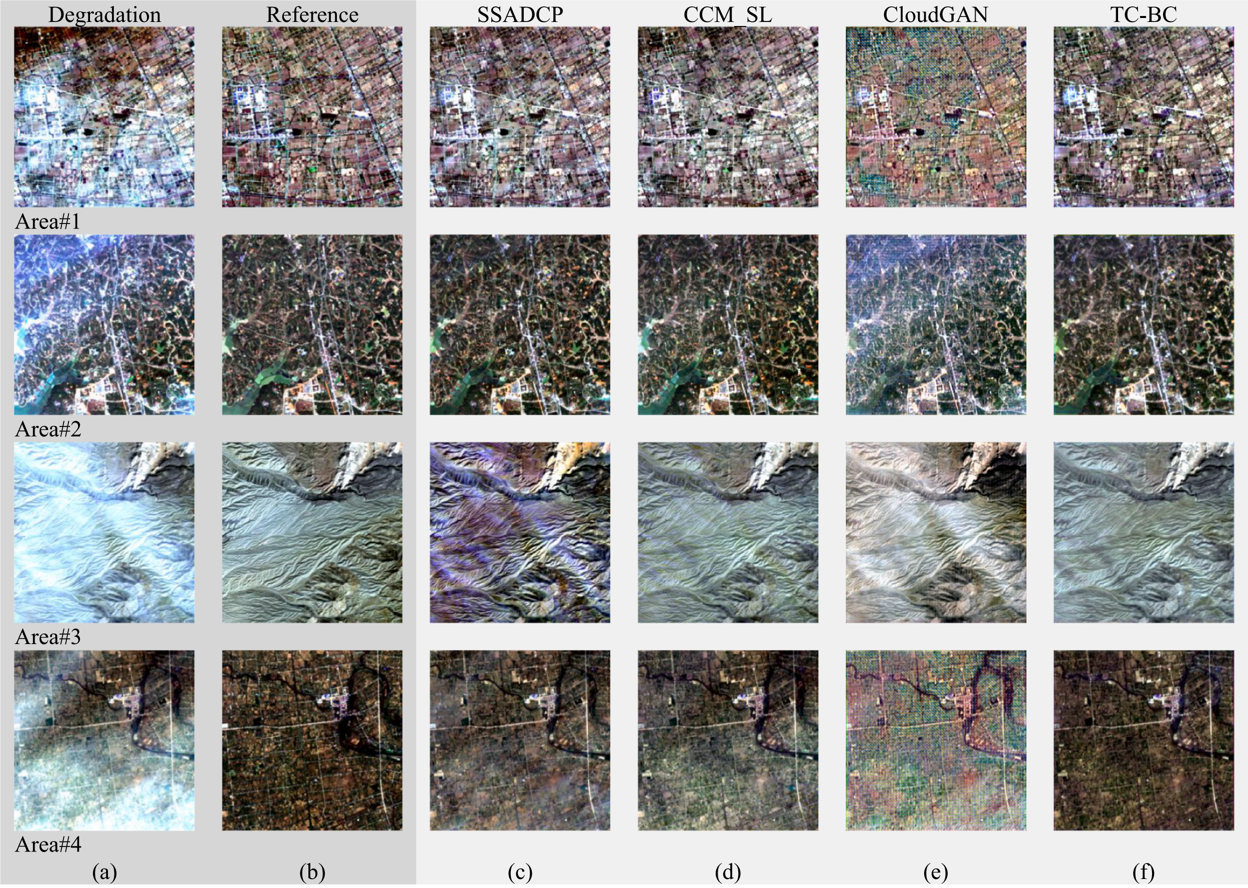


Fig. 10. Comparison of the thin cloud correction results of the different methods. (a) Cloudy image. (b) Ground reference. (c) SSADCP. (d) CCM_SL. (e) CloudGAN. (f) TC-BC.

4.2. *Expansion of the UBCSet dataset*

The patches of the UBCSet can be continuously added during thin cloud correction tasks. In addition, the UBCSet can be expanded by synthesizing new data. In order to compare the performance between the FSNet module and other networks, a traditional paired dataset named Ex_UBCSet was expanded from the UBCSet. Through the imaging model of the cloudy images, two images respectively coming from the ground set and cloud set in the UBCSet can be synthesized into a new cloudy patch. In the Ex_UBCSet, there are 1139 paired patches in the training set and 325 paired patches in the test set. Because of the same ground information in the synthesized cloudy patches and paired ground patches, the Ex_UBCSet is suitable for guiding supervised network training and appraising the thin cloud correction ability by evaluating the difference between the correction result and the ground patch.

5. Experiments and discussion

In this section, the thin cloud correction effectiveness of TC-BC is first discussed. We then describe how the superiority of the FSNet module was verified with five comparative deep learning networks. The cloud-ground separation results and the validation results of the TC-BC method for specific applications are then provided. Qualitative and quantitative comparisons are both employed in this paper to evaluate the proposed TC-BC method. To analyze the comparative methods more completely, two widely used quantitative metrics are used, i.e., the peak signal-to-noise ratio (PSNR) and the structural similarity index measure (SSIM). The larger the value of the metrics, the better the performance of the method.

5.1. *Thin cloud correction with the TC-BC method*

The TC-BC method was trained on the UBCSet for the thin cloud correction task and compared with three advanced cloud removal methods, i.e., the spatial-spectral adaptive dark channel prior (SSADCP) method (Shen et al., 2020), the cirrus correction method with scattering law (CCM_SL)(Zhang et al., 2021), and CloudGAN (Singh and Komodakis, 2018). To comprehensively present the thin cloud correction results of the different methods, four remote sensing images with varying degrees of thin cloud degradation and different land cover were selected, as shown in Fig. 10(a). The ground references are illustrated in Fig. 10(b). All the images are from real scenes, with a 16-day interval of acquisition time. Fig. 10(c)–(f) show the cloud correction results obtained using the different methods, i.e., SSADCP, CCM_SL, CloudGAN, and the proposed TC-BC method.

It can be seen that Area #1 and Area#2 feature a few thin clouds with a homogenous spatial distribution, while Area#3 and Area#4 feature large areas of thin clouds with strong spatial heterogeneity. The main land-cover types of the four images are city, vegetation, soil, and water. As shown in the results of SSADCP, the degradation of the most homogeneous thin clouds can be effectively corrected to restore the ground-truth information. However, in the images with large areas of thin clouds and a wide cloud distribution, there are still residual clouds and spectral distortion in the correction results, as shown in Fig. 10(c) Area#3 and Area#4. It can be found that SSADCP is not suitable for the correction task when the interference of thin clouds is too serious and the surface information loss exceeds 50 %. Since most thin clouds belong to cirrus clouds, high-precision correction of thin clouds can be achieved in most of the images, as shown in the results of CCM_SL. However, the junctions in some of the correction results between the cloud region and

Table 1
The metric scores for the different images in the thin cloud correction task.

Area	Evaluation indicator	SSADCP	CCM_SL	CloudGAN	TC-BC
#1	PSNR	35.00	42.26	31.51	**42.78**
	SSIM	0.9203	0.9589	0.9298	**0.9640**
#2	PSNR	33.88	36.01	26.63	**42.26**
	SSIM	0.9270	0.9596	0.9379	**0.9731**
#3	PSNR	28.95	37.99	22.18	**39.69**
	SSIM	0.7557	0.9413	0.9033	**0.9466**
#4	PSNR	35.28	40.24	32.47	**40.87**
	SSIM	0.8870	0.9350	0.8122	**0.9546**

the clear region are unnatural, with obvious spectral distortion, as shown in Fig. 10(d) Area#4. This is because the correction process of CCL_SM is only for a single pixel, without considering the global features of the correction results. In the results of CloudGAN, as shown in the fifth column of Fig. 10, two serious problems are prominent: the checkerboard effect and residual clouds. These problems can be attributed to the unstable training of CloudGAN, as it only focuses on how to enhance the ground features, and does not take into account the relationship between the cloud and the ground. The weak constraint of the network leads to non-unique or unstable solutions. The TC-BC method can respond to this challenge well and achieve a superior performance, even when faced with complex cloud contamination, owing to its adequate consideration of the spatial and spectral characteristics of clouds. Compared with the above three methods, the proposed TC-BC method uses an information interaction block to divide the cloud and ground information in the cloud coverage area to achieve complete removal of thin clouds. The spatially adaptive structure aggregates the global information of the image, which is helpful for accurately identifying cloudy and clear regions and achieving a natural transition of spectral features. Furthermore, the discriminator enhances the authenticity of the correction results. As a result, the TC-BC method achieves the best correction effect on all the images in this experiment. The quantitative comparison results are listed in Table 1. The results suggest that the TC-BC method shows obvious advantages, with the highest PSNR and SSIM.

5.2. Performance of the FSNet module

In order to prove the effectiveness of the FSNet module for separating surface information, five representative methods were utilized for comparison, based on the Ex_UBCSet: two traditional networks, i.e., U-Net (Ronneberger et al., 2015; Zheng et al., 2021) and ResNet (He et al., 2016; Meraner et al., 2020); two thin cloud correction networks, i.e., RSCNet (Li et al., 2019) and SPAGAN(Wang et al., 2019; Pan, 2020); and also an interactive dual-stream strategy called “your trash is my treasure” (YTMT) (Hu and Guo, 2021). It should be clarified that Ex_UBCSet is a paired synthesized dataset, which was specially used for evaluating the ground generation ability of the compared networks.

For the six methods, it can be observed that all of the methods can

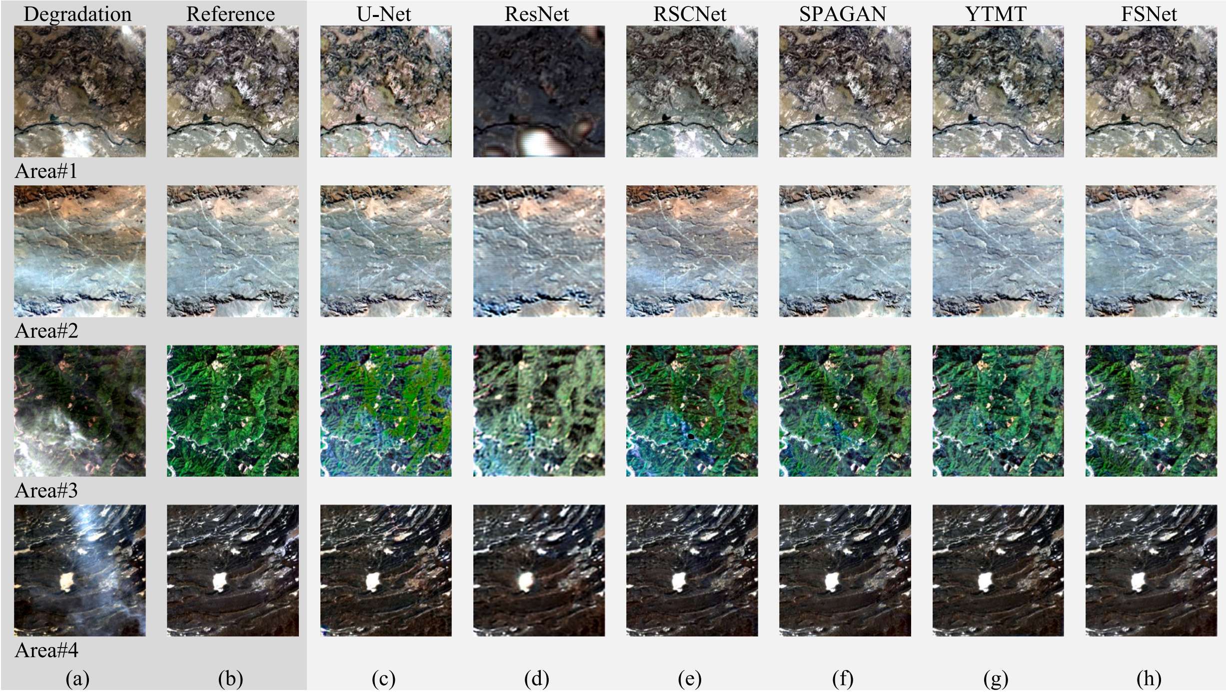


Fig. 11. Comparison of the thin cloud correction results of the different networks. (a) Cloudy image. (b) Ground reference. (c) U-Net. (d) ResNet. (e) RSCNet. (f) SPAGAN. (g) YTMT. (h) FSNet.

Table 2
The metric scores for the different images in the thin cloud correction task.

Area	Evaluation indicator	U-Net	Res_Net	RSC_Net	SPAGAN	YTMT	FSNet
#1	PSNR	38.55	29.92	41.23	45.34	43.49	**48.88**
	SSIM	0.9828	0.8326	0.9866	0.9952	0.9929	**0.9973**
#2	PSNR	46.26	43.47	41.43	53.38	51.05	**53.56**
	SSIM	0.9931	0.9770	0.9894	0.9978	0.9962	**0.9979**
#3	PSNR	41.59	37.65	39.12	44.98	45.90	**46.02**
	SSIM	0.9959	0.9885	0.9926	0.9983	0.9983	**0.9987**
#4	PSNR	41.34	37.96	41.38	43.65	43.65	**44.07**
	SSIM	0.9795	0.9534	0.9889	0.9917	0.9894	**0.9953**

Table 3
The metric scores with the Ex_UBCset.

Evaluation indicator	U-Net	Res_Net	RSC_Net	SPAGAN	YTMT	FSNet
PSNR	39.78	37.09	38.87	43.06	41.94	**44.28**
SSIM	0.9837	0.9423	0.9715	0.9915	0.9889	**0.9931**

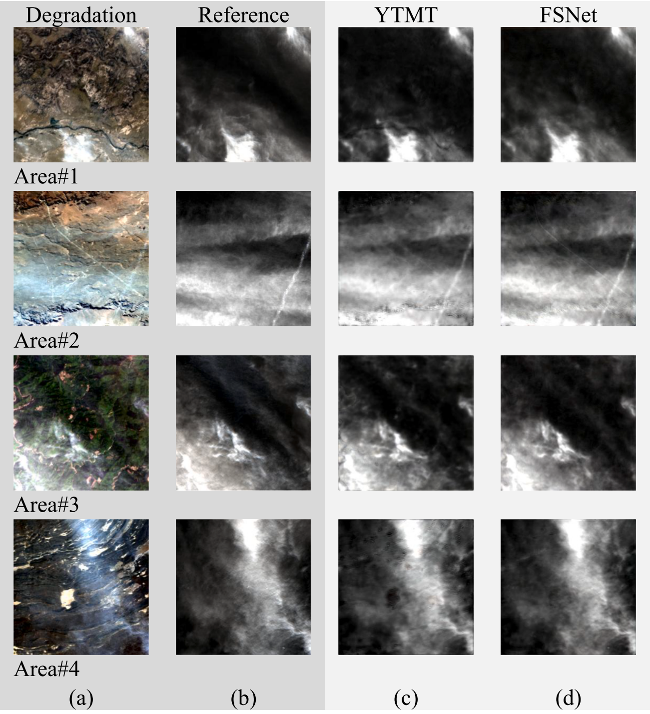


Fig. 12. Comparison of the thin cloud detection results obtained by the different networks. (a) Cloudy image. (b) Cloud reference. (c) YTMT. (d) FSNet.

remove most of the thin clouds in the degraded images, but the details of the correction results are always different. Four representative thin cloud contaminated images are presented to compare the performance of the different methods in Fig. 11, and a quantitative comparison of the results for the four images is provided in Table 2. These images contain different land cover and cloud cover. The experimental results show that the FSNet module can remove thin clouds more completely in cloud regions and maintain the color better in clear regions, which suggests that the FSNet module has a superior thin cloud correction ability.

Furthermore, the average thin cloud correction accuracy of the different networks with the Ex_UBCSet is provided in Table 3. The quantitative metrics for the traditional networks are clearly lower than those for the FSNet module with the Ex_UBCSet, which further demonstrates the superiority of the proposed network. Furthermore, experimental verification on a previously available dataset (Li et al., 2019) yielded comparable results to those obtained with the Ex_UBCSet.

In addition, YTMT and FSNet can simultaneously detect clouds while also correcting clouds. The cloud detection results for the cloudy images are given in Fig. 12. The experimental results show that the detection result of the FSNet module is better than that of YTMT. In Fig. 12(c) Area#1 and Area#2, YTMT cannot separate the cloud cover on the dark surface, resulting in unsatisfactory cloud detection results that are not continuous enough to match the true cloud distribution. In Area#3, the bright surface part is identified as clouds by YTMT, and the cloud is identified as a bright surface in Area#4 by YTMT. However, these phenomena never occur with the FSNet module. This is because the FSNet module uses a spatial adaptive structure considering the global distribution of the thin clouds.

An ablation study was conducted for each component in the FSNet module, i.e., the information interaction structure, the spatially adaptive structure, and the attention structure, to analyze their accuracy contribution. The specific results listed in Table 4 show that the introduction of each structure significantly improves the quantitative indicators, compared to the baseline.

Table 4
The metric scores for the ablation study with the Ex_UBCset.

Structure				**Evaluation indicator**	
Baseline	IIS	SAS	AS	PSNR	SSIM
√				38.38	0.9772
√	√			39.68	0.9784
√	√	√		39.79	0.9802
√		√	√	43.01	0.9895
√	√	√	√	44.28	0.9931

*IIS: information interaction structure. SAS: spatial adaptive structure (without attention structure). AS: attention structure.

5.3. *Cloud and ground separation*

To further verify the effect of the proposed TC-BC method in the thin cloud detection and correction task in the real world, four degraded images with different cloud amounts and different degradations of the ground surface were selected to demonstrate the separation effectiveness for cloud and ground information. The results are shown in Fig. 13. The experimental results indicate that the cloud components exhibit a high degree of visual similarity to the actual cloud, while the ground components accurately preserve both the spatial details and spectral characteristics.

A real thin cloudy image was obtained on February 19, 2023, as shown in Fig. 14(a), with the center coordinates 118°01′40″E and 31°59′07″N. Visual observation confirms a significant presence of thin clouds in this cloudy image. However, the corresponding cirrus band exhibited in Fig. 14(b) shows very few clouds. Therefore, the observed degradation is mainly caused by other thin clouds, besides cirrus clouds. The clear reference image was acquired on February 3, 2023, and is shown in Fig. 14(c). Fig. 14(d) and Fig. 14(e) are the cloud and ground components separated by the TC-BC method. As the figure shows, the TC-BC method effectively captures the intricate boundaries of the mixed thin clouds and accurately distinguishes thin clouds from the ground. This suggests the robustness of the TC-BC method in detecting and separating thin clouds of various forms and characteristics.

5.4. *Applications of TC-BC*

In this section, we describe how the effectiveness and applicability of the TC-BC method were validated in large scenes and different satellite sensors.

5.4.1. *Validation on large complex scenes*

In order to validate the effectiveness of the TC-BC method on large scenes of remote sensing data, two remote sensing images covering extensive land areas were selected for the cloud-ground separation experiment. The results are shown in Fig. 15. Fig. 15(a) is a thin cloud covered OLI image centered at 114°38′55″ W and 36°13′12″ N, which was acquired on February 17, 2016. The main land-cover types of this image are rock, water, and buildings, and the cloud in this image is thin but unevenly distributed. It is worth emphasizing that the rocks and

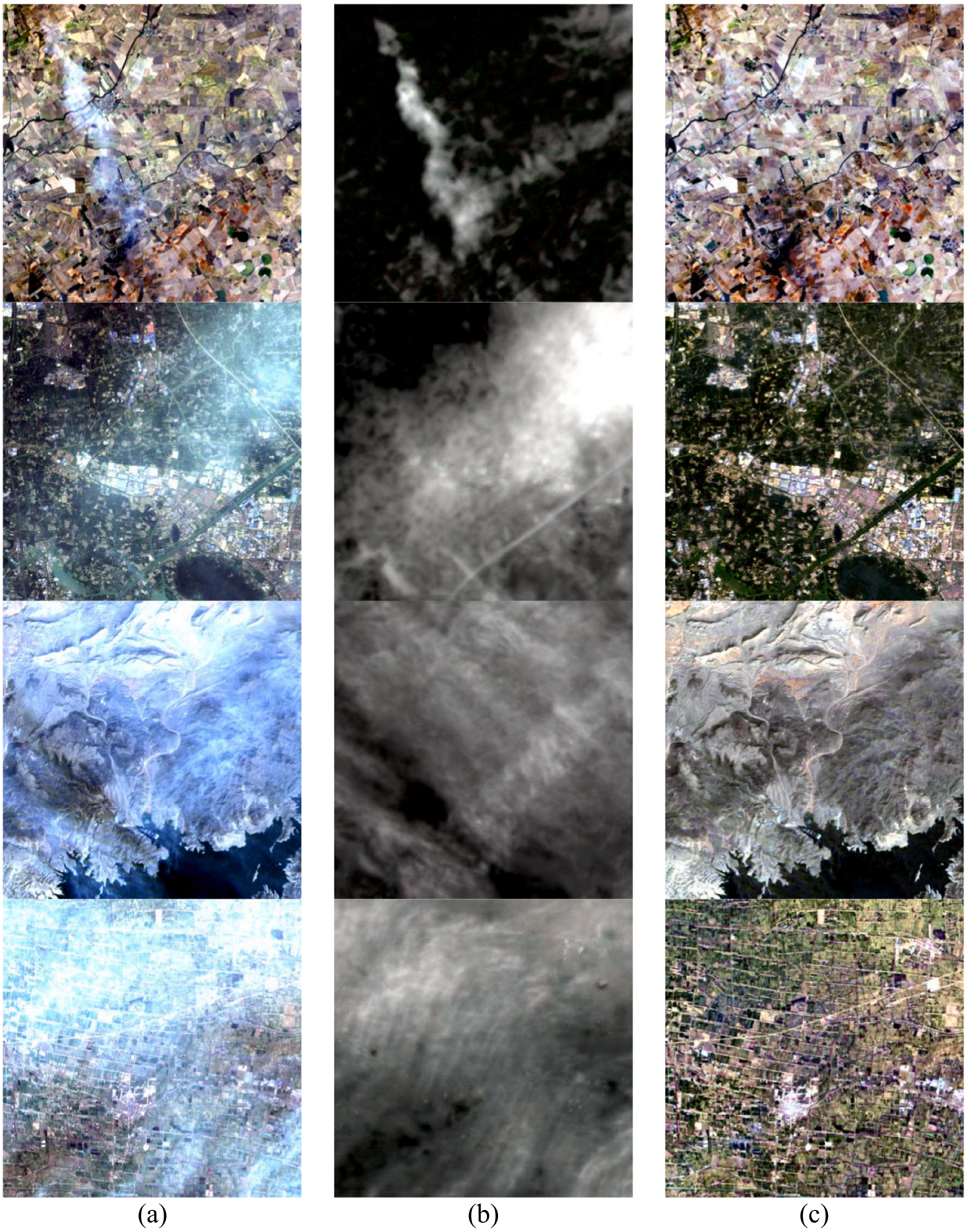


Fig. 13. Cloud-ground separation results of TC-BC. (a) Cloudy image. (b) Cloud component. (c) Ground component.

buildings in this image exhibit high brightness, which makes them more challenging to separate from the clouds. The reference clear image is shown in Fig. 15(b), which was acquired on March 4, 2016. Fig. 15(c) shows the thin cloud detection image, and Fig. 15(d) shows the correction results of the TC-BC method. It can be found that the TC-BC method is suitable for the large-scale data and can achieve excellent separation of cloud and ground information, even under the condition of high thin cloud coverage. Fig. 15(e) is a cloudy image centered at 112°3′1.15″E and 33°4′25.62″N, which was acquired on February 9, 2021. The main land-cover types of this image are vegetation, water, soil, and buildings. The ground reference is shown in Fig. 15(f), which was acquired on January 8, 2021. Fig. 15(g) shows the thin cloud detection image, and Fig. 15(h) shows the correction results of the TC-BC method. It can be found that the TC-BC method can achieve excellent separation of the cloud and ground information, even in the condition of various mixed land-cover types. The corrected result is visually greener than the reference, which can be attributed to the phenological difference of the forest between the two dates and the different lighting conditions.

5.4.2. *Validation on Sentinel-2 data*

Two cloudy Sentinel-2 Multi-Spectral Instrument (MSI) images were collected to investigate the compatibility of the proposed method with different satellite sensors. The results are shown in Fig. 16. The first cloudy image in Fig. 16(a) centered at 113°20′37″E and 32°52′18″N was acquired on June 26, 2023. Fig. 16(b) shows its reference clear image, which was acquired on June 6, 2023. The second image centered at 111°20′34.04″E and 33°51′22.14″N was captured on June 9, 2023, as shown in Fig. 16(e). Its reference clear image was captured on June 24, 2023, as shown in Fig. 16(f). There was little surface change between the cloudy images and their reference images. Fig. 16(c) and Fig. 16(g) show thin cloud detection results. Correction results are shown in Fig. 16(d)

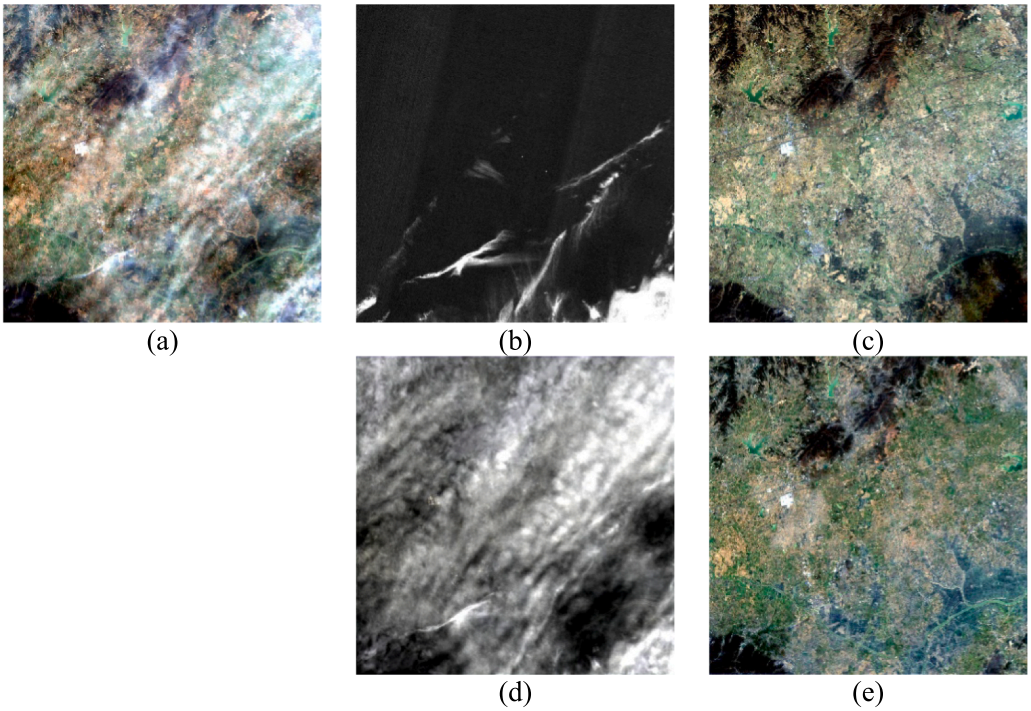


Fig. 14. Cloud-ground separation result obtained by TC-BC for the non-cirrus thin cloudy image. (a) Cloudy image. (b) The cirrus band of the cloudy image. (c) Ground reference. (d) Cloud component. (e) Ground component.

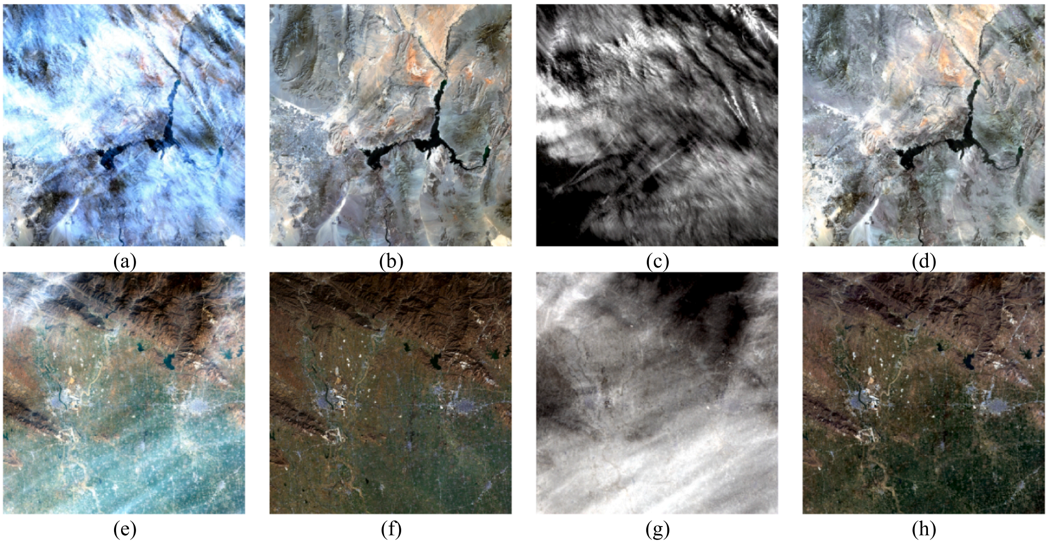


Fig. 15. Cloud-ground separation results obtained by TC-BC with large-scale data in Landsat-8/9.

and Fig. 16(h). The Sentinel-2 data have a varying resolution across the different bands, with the VNIR bands used for correction having a resolution of 10 m while the cirrus band has a resolution of 60 m. Therefore, we used the 30 m resolution cloudy images generated from Landsat-8/9 to guide the TC-BC method. Despite the resolution differences between the two sensors, the TC-BC method still achieves effective cloud detection and removal in Sentinel-2 data. This further demonstrates the applicability of the proposed TC-BC method across different sensors.

6. Conclusion

In this paper, the novel thin cloud blind correction (TC-BC) method under an unsupervised self-training framework has been proposed to solve the thin cloud contamination problem in optical remote sensing images. The TC-BC method is both more universal and can obtain a higher thin cloud correction accuracy than the previous methods. The key idea is the coupling of a physical mechanism and a data-driven method. Specifically, a self-training framework connecting the FSNet module and the imaging model of cloudy images makes unsupervised training possible. The design of the FSNet module takes the cloud distribution characteristics of cloudy images into account, which helps to improve the correction accuracy of the TC-BC method at the pixel level. In addition, the scattering law of the cloud was considered while constructing the UBCSet.

The TC-BC method, as proposed in this paper, has strong

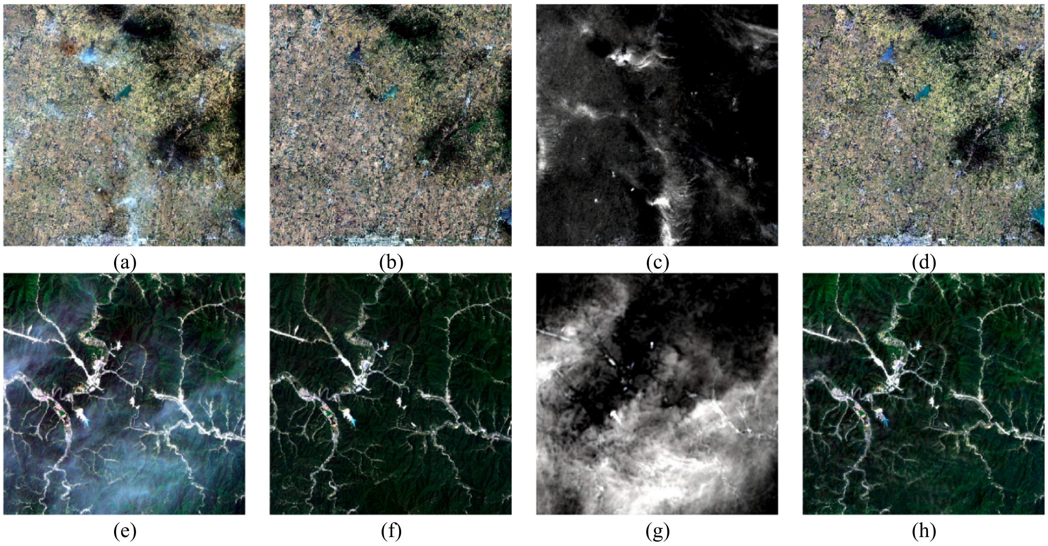


Fig. 16. Cloud-ground separation results obtained by the TC-BC method with large-scale Sentinel-2 data.

practicability because of the low cost of the dataset construction, which not only gets rid of the need for massive paired data but also overcomes the dependence on specific bands or prior knowledge. The qualitative and quantitative analyses demonstrated that the proposed TC-BC method can achieve a superior thin cloud correction performance, compared with three advanced cloud correction methods. The correction results of the TC-BC method were superior in thin cloud removal, ground fidelity, and global consistency, all of which indicate that the proposed TC-BC method can adapt to images with complex thin cloud contamination. Moreover, the FSNet module performed better than the other five deep learning networks in both cloud correction and cloud detection.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was funded by the National Key Research and Development Program of China (2022YFF1301103), the National Natural Science Foundation Excellent Young Scientists Fund Program (42422109) and the Fundamental Research Funds for the Central Universities (2042023kfyq04). The authors would like to thank all the researchers who kindly shared the codes used in this study.

Appendix

Table A1
Original data information for the UBCSet.

Data	Top left corner		Bottom right corner		Data	Top left corner		Bottom right corner	
2011/11/26	031°00′28.17″	N	029°52′22.89″	N	2018/4/12	038°12′29.19″	S	039°35′32.78″	S
	113°18′03.26″	E	114°32′40.60″	E		070°07′00.45″	W	068°21′00.73″	W
2013/12/23	035°14′44.76″	N	034°54′03.52″	N	2018/8/7	042°21′43.86″	N	042°01′33.46″	N
	119°12′17.04″	W	118°46′47.82″	W		005°11′19.03″	W	004°43′08.76″	W
2013/12/23	034°29′54.59″	N	034°09′28.95″	N	2018/8/7	004°15′95.84″	N	041°27′35.58″	N
	119°08′31.53″	W	118°43′17.76″	W		005°39′49.59″	W	004°54′35.01″	W
2013/12/23	034°15′56.96″	N	033°55′32.88″	N	2018/8/7	041°25′36.83″	N	041°05′49.68″	N
	117°55′28.37″	W	117°30′36.19″	W		006°06′11.31″	W	005°38′03.59″	W
2014/6/12	041°44′48.48″	N	041°03′45.92″	N	2018/8/23	042°21′43.86″	N	042°01′33.46″	N
	085°09′11.93″	E	086°33′04.44″	E		005°11′19.03″	W	004°43′08.76″	W
2014/6/12	041°08′49.10″	N	040°48′08.22″	N	2018/8/23	004°15′95.84″	N	041°27′35.58″	N
	086°35′11.99″	E	087°02′38.02″	E		005°39′49.59″	W	004°54′35.01″	W
2015/9/23	044°03′39.99″	N	043°21′55.39″	N	2018/8/23	041°25′36.83″	N	041°05′49.68″	N
	080°59′06.99″	W	080°02′14.45″	W		006°06′11.31″	W	005°38′03.59″	W
2016/2/17	036°47′20.58″	N	035°39′38.52″	N	2018/8/23	041°57′58.61″	N	041°17′19.80″	N
	115°19′54.64″	W	114°00′00.85″	W		004°52′44.64″	W	003°56′34.76″	W
2016/3/4	036°47′20.58″	N	035°39′38.52″	N	2018/9/26	046°53′48.12″	N	046°13′18.29″	N
	115°19′54.64″	W	114°00′00.85″	W		072°56′29.19″	E	073°57′41.41″	E
2016/3/4	036°14′59.64″	N	035°15′50.95″	N	2018/9/26	046°48′40.15″	N	046°28′10.42″	N
	115°33′16.87″	W	114°23′30.77″	W		073°57′31.26″	E	074°27′54.79″	E
2016/4/16	035°07′48.75″	N	034°46′54.26″	N	2022/7/11(L9)	032°23′09.23″	N	031°09′22.34″	N
	135°41′34.40″	E	136°06′34.24″	E		120°14′33.06″	E	121°09′38.06″	E
2016/4/16	032°52′34.96″	N	032°18′27.64″	N	2022/7/11(L9)	032°34′22.97″	N	032°08′51.90″	N

(*continued on next page*)

Table A1 (*continued*)

Data	Top left corner		Bottom right corner		Data	Top left corner		Bottom right corner	
	119°11′33.94″	E	119°59′30.33″	E		118°18′25.99″	E	119°15′07.27″	E
2016/5/2	034°51′06.25″	N	033°52′24.36″	N	2022/7/28	032°09′22.39″	N	031°49′49.15″	N
	135°38′11.73″	E	136°47′23.16″	E		119°32′32.32″	E	120°21′51.46″	E
2016/5/2	035°07′48.75″	N	034°46′54.26″	N	2022/7/28	031°52′10.73″	N	031°31′53.04″	N
	135°41′34.40″	E	136°06′34.24″	E		120°27′30.28″	E	120°52′05.87″	E
2016/5/2	033°14′00.47″	N	032°26′26.91″	N	2022/08/05(L9)	032°23′09.23″	N	031°09′22.34″	N
	118°24′47.96″	E	120°02′16.68″	E		120°14′33.06″	E	121°09′38.06″	E
2017/2/27	033°45′15.46″	N	033°02′44.89″	N	2022/08/05(L9)	032°34′22.97″	N	032°08′51.90″	N
	119°24′06.96″	E	120°12′14.66″	E		118°18′25.99″	E	119°15′07.27″	E
2018/4/8	031°02′39.00″	N	030°15′16.77″	N	2022/8/13	032°09′22.39″	N	031°49′49.15″	N
	113°11′02.01″	E	114°36′08.54″	E		119°32′32.32″	E	120°21′51.46″	E

*(L9) indicates Landsat-9 data.

References

Ahmad, A., Abd Ghani, M.K., Razali, S.N., Sakidin, H., Hashim, N., 2014. Haze reduction from remotely sensed data, Appl. Math. Sci. 1755–1762, 10.12988/ams.2014.4289, 2014.

Bell, S., Zitnick, C. L., Bala, K., Girshick, R., 2016. Inside-outside net: detecting objects in context with skip pooling and recurrent neural networks. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 27-30 June 2016, 2874–2883, 10.1109/CVPR.2016.314.

Cai, B., Xu, X., Jia, K., Qing, C., Tao, D., 2016. DehazeNet: an end-to-end system for single image haze removal. IEEE Trans. Image Process. 25, 5187–5198. https://doi.org/10.1109/TIP.2016.2598681.

Chavez, P.S., 1988. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. Remote Sens. Environ. 24, 459–479. https://doi.org/10.1016/0034-4257(88)90019-3.

Chen, S., Chen, X., Chen, J., Jia, P., Cao, X., Liu, C., 2016. An iterative haze optimized transformation for automatic cloud/haze detection of landsat imagery. IEEE Trans. Geosci. Remote Sens. 54, 2682–2694. https://doi.org/10.1109/TGRS.2015.2504369.

Crist, E.P., Cicone, R.C., 1984. A physically-based transformation of thematic mapper data—the TM Tasseled cap. IEEE Trans. Geosci. Remote Sens. GE-22, 256–263. https://doi.org/10.1109/TGRS.1984.350619.

Enomoto, K., Sakurada, K., Wang, W., Fukui, H., Matsuoka, M., Nakamura, R., Kawaguchi, N., 2017. Filmy cloud removal on satellite imagery with multispectral conditional generative adversarial nets. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 21-26 July 2017, 1533-1541, 10.1109/CVPRW.2017.197.

Gao, B.C., Kaufman, Y.J., Han, W., Wiscombe, W., 1998. J.: Correction of thin cirrus path radiances in the 0.4–1.0 μm spectral region using the sensitive 1.375 μm cirrus detecting channel. J. Geophys. Res.: Atmos. 103, 32169–32176.

Grohnfeldt, C., Schmitt, M., and Zhu, X.: A Conditional Generative Adversarial Network to Fuse Sar And Multispectral Optical Data For Cloud Removal From Sentinel-2 Images, IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium, 22-27 July 2018, 1726-1729, 10.1109/IGARSS.2018.8519215.

Guanter, L., Richter, R., Kaufmann, H., 2009. On the application of the MODTRAN4 atmospheric radiative transfer code to optical remote sensing. Int. J. Remote Sens. 30, 1407–1424. https://doi.org/10.1080/01431160802438555.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep Residual Learning for Image Recognition, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 27-30 June 2016, 770-778, 10.1109/CVPR.2016.90.

He, X.Y., Hu, J.B., Chen, W., Li, X.Y., 2010. Haze removal based on advanced haze-optimized transformation (AHOT) for multispectral imagery. Int. J. Remote Sens. 31, 5331–5348. https://doi.org/10.1080/01431160903369600.

Hu, Q., Guo, X., 2021. Trash or treasure? An interactive dual-stream strategy for single image reflection separation. CoRR abs/2110.10546.

Jiang, H., Lu, N., 2018. Multi-scale residual convolutional neural network for haze removal of remote sensing images, 10.3390/rs10060945, 2018.

Kaiming, H., Jian, S., Xiaoou, T., 2009. Single image haze removal using dark channel prior. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition, 20-25 June 2009, 1956-1963, 10.1109/CVPR.2009.5206515.

Kneizys, F., Shettle, E., Gallery, W., Jr, J., Abreu, L., McClatchey, R., Fenn, R., Selby, J., 1980. Atmospheric transmittance/radiance: Computer code LOWTRAN 5, Unknown, 1980.

Li, H., Zhang, L., Shen, H., Li, P., 2012. A variational gradient-based fusion method for visible and SWIR imagery. Photogram. Eng. Remote Sens. 78, 947–958, 10.14358/PERS.78.9.947, 2012.

Li, W., Li, Y., Chen, D., Chan, J.-C.-W., 2019. Thin cloud removal with residual symmetrical concatenation network. ISPRS J. Photogram. Remote Sens. 153, 137–150. https://doi.org/10.1016/j.isprsjprs.2019.05.003.

Li, J., Wu, Z., Hu, Z., Zhang, J., Li, M., Mo, L., Molinier, M., 2020. Thin cloud removal in optical remote sensing images based on generative adversarial networks and physical model of cloud distortion. ISPRS J. Photogram. Remote Sens. 166, 373–389. https://doi.org/10.1016/j.isprsjprs.2020.06.021.

Liang, S., Fang, H., Chen, M., 2001. Atmospheric correction of Landsat ETM+ land surface imagery. I. Methods. IEEE Trans. Geosci. Remote Sens. 39, 2490–2498. https://doi.org/10.1109/36.964986.

Liu, J., Wang, X., Chen, M., Liu, S., Zhou, X., Shao, Z., Liu, P., 2014. Thin cloud removal from single satellite images. Optics Express 22, 618–632. https://doi.org/10.1364/oe.22.000618.

Lv, H., Wang, Y., Shen, Y., 2016. An empirical and radiative transfer model based algorithm to remove thin clouds in visible bands. Remote Sens. Environ. 179, 183–195. https://doi.org/10.1016/j.rse.2016.03.034.

Makarau, A., Richter, R., Müller, R., Reinartz, P., 2014. Haze detection and removal in remotely sensed multispectral imagery. IEEE Trans. Geosci. Remote Sens. 52, 5895–5905. https://doi.org/10.1109/TGRS.2013.2293662.

Meraner, A., Ebel, P., Zhu, X.X., Schmitt, M., 2020. Cloud removal in Sentinel-2 imagery using a deep residual neural network and SAR-optical data fusion. ISPRS J. Photogram. Remote Sensing: Off. Publicat. Int. Soc. Photogram. Remote Sens. (ISPRS) 166, 333–346. https://doi.org/10.1016/j.isprsjprs.2020.05.013.

Mirza, M., Osindero, S., 2014. Conditional generative adversarial nets, arXiv preprint arXiv:1411.1784, 2014.

Pan, H., 2020. Cloud Removal for Remote Sensing Imagery via Spatial Attention Generative Adversarial Network, 2020.

Povey, A.C., Grainger, R.G., 2015. Known and unknown unknowns: uncertainty estimation in satellite remote sensing. Atmos. Meas. Tech. 8, 4699–4718. https://doi.org/10.5194/amt-8-4699-2015.

Qin, M., Xie, F., Li, W., Shi, Z., Zhang, H., 2018. Dehazing for multispectral remote sensing images based on a convolutional neural network with the residual architecture. IEEE J. Select. Top. Appl. Earth Observat. Remote Sens. 11, 1645–1655. https://doi.org/10.1109/JSTARS.2018.2812726.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., Prabhat, 2019. Deep learning and process understanding for data-driven Earth system science. Nature, 566 (2019) 195–204, 10.1038/s41586-019-0912-1.

Richter, R., 1996. A spatially adaptive fast atmospheric correction algorithm. Int. J. Rem. Sens. 17, 1201–1214. https://doi.org/10.1080/01431169608949077.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional networks for biomedical image segmentation, medical image computing and computer-assisted intervention – MICCAI 2015, Cham; 2015. p. 234–241.

Shen, H., Li, H., Qian, Y., Zhang, L., Yuan, Q., 2014. An effective thin cloud removal procedure for visible remote sensing images. ISPRS J. Photogram. Remote Sens. 96, 224–235. https://doi.org/10.1016/j.isprsjprs.2014.06.011.

Shen, Y., Wang, Y., Lv, H., Qian, J., 2015. Removal of thin clouds in Landsat-8 OLI data with independent component analysis. Remote. Sens. 7, 11481–11500.

Shen, H., Zhang, C., Li, H., Yuan, Q., Zhang, L., 2020. A spatial-spectral adaptive haze removal method for visible remote sensing images. IEEE Trans. Geosci. Remote Sens. 58, 6168–6180. https://doi.org/10.1109/TGRS.2020.2974807.

Singh, P., Komodakis, N., 2018. Cloud-gan: cloud removal for Sentinel-2 imagery using a cyclic consistent generative adversarial networks. In: IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium, 22-27 July 2018, 1772-1775, 10.1109/IGARSS.2018.8519033.

Toizumi, T., Zini, S., Sagi, K., Kaneko, E., Tsukada, M., Schettini, R., 2019. Artifact-Free Thin Cloud Removal Using Gans, 2019 IEEE International Conference on Image Processing (ICIP), 22-25 Sept. 2019, 3596-3600, 10.1109/ICIP.2019.8803652.

Wang, T., Yang, X., Xu, K., Chen, S., Zhang, Q., Lau, R.W.H., 2019. Spatial attentive single-image deraining with a high quality real rain dataset. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 15-20 June 2019, 12262-12271, 10.1109/CVPR.2019.01255.

Yong, D., Guindon, B., Cihlar, J., 2002. Haze detection and removal in high resolution satellite image with wavelet analysis. IEEE Trans. Geosci. Remote Sens. 40, 210–217. https://doi.org/10.1109/36.981363.

Zhang, Y., Guindon, B., Cihlar, J., 2002. An image transform to characterize and compensate for spatial variations in thin cloud contamination of Landsat images. Remote Sens. Environ. 82, 173–187. https://doi.org/10.1016/S0034-4257(02)00034-2.

Zhang, C., Li, H., Shen, H., 2021. A scattering law based cirrus correction method for Landsat 8 OLI visible and near-infrared images. Remote Sens. Environ. 253, 112202. https://doi.org/10.1016/j.rse.2020.112202.

Zheng, J., Liu, X.Y., Wang, X., 2021. Single image cloud removal using U-net and generative adversarial networks. IEEE Trans. Geosci. Remote Sens. 59, 6371–6385. https://doi.org/10.1109/TGRS.2020.3027819.

Zhu, J.Y., Park, T., Isola, P., Efros, A.A., 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: 2017 IEEE International Conference on Computer Vision (ICCV), 22-29 Oct. 2017, 2242-2251, 10.1109/ICCV.2017.244.

Zhu, Q., Mai, J., Shao, L., 2015. A fast single image haze removal algorithm using color attenuation prior. IEEE Trans. Image Process. 24, 3522–3533. https://doi.org/10.1109/TIP.2015.2446191.