Evaluation of Three Error-Correction Models Based on the Matched Pixel Scale Ground “Truth”: A Case Study of MCD43A3 V006 Over the Heihe River Basin, China

Jingping Wang, Xiaodan Wu, Rongqi Tang, Qicheng Zeng, Zheng Li, Jianguang Wen, and Qing Xiao

Abstract—Validation merely provides the “impression of accuracy” of satellite products but does not deal with their errors. There have been several error correction/adjustment models. Nevertheless, their performances have not been evaluated and compared. In this study, three typical models, namely random forests (RF), cumulative distribution function (CDF), and Kalman filter (KF) were comprehensively evaluated based on the pixel scale ground “truth” regarding their ability to correct errors of coarse-resolution satellite albedo products. Moderate-resolution imaging spectroradiometer albedo product (i.e., MCD43A3 V006) was utilized as an example due to its widespread use and application. These three models all show significant improvements regarding the accuracy of the corrected MCD43A3. Root-mean-square error (RMSEs) decreased from 0.037–0.020, 0.021, and 0.025 for RF, CDF, and KF, respectively. Biases were reduced from -0.018–0.004, -0.001, and -0.001 for these three models, respectively. And R² was increased from 0.585–0.849, 0.823, and 0.764 for RF, CDF, and KF, respectively. Generally, RF shows the best overall performance, followed by CDF and KF. These three models are more adept at handling the bias of MCD43A3 than their consistency with respect to the pixel scale ground “truth”, and the improvement is the most significant at the sites with large errors. Nevertheless, the performance of RF shows dependence on the number and representativeness of training samples. When these conditions were not satisfied, CDF performs best in this situation. Regarding the stability of their performance, RF performs better in reducing RMSE while CDF performs better in reducing Bias and improving consistency.

Index Terms—Cumulative distribution function, error correction, Kalman filter, MCD43A3, pixel scale ground “truth”, random forests.

I. INTRODUCTION

AND surface albedo controls the distribution of radiance energy between the earth surface and atmosphere, playing an important part in the surface-to-air energy exchange and climate change mechanism research [1]. Compared with traditional in situ observations, which are only conducted in small areas, remote sensing technology provides the only practical way to map the macro earth surface regularly. There have been many satellite albedo products, such as moderate-resolution imaging spectroradiometer (MODIS), global land surface satellites (GLASS), and Suomi National Polar-orbiting Partnership Visible Infrared Imaging Radiometer Suite. However, since the satellite-derived albedo is not a directly observed variable but is indirectly retrieved from the satellite-observed reflectance or radiance using albedo algorithms, they inevitably suffer from errors caused by the errors of satellite observations and the limitations of the retrieval algorithms.

Previous studies have made comprehensive assessments regarding the accuracy and uncertainty of satellite albedo products and demonstrated that these products can not meet the Global Climate Observing System accuracy requirement of 5% in some areas. Stroeve et al. [2] used in situ albedo measurements on the Greenland Ice Sheet to evaluate the MCD43 V005 product, and showed that albedo with solar zenith angles less than 75° presents a root-mean-square error (RMSE) of 0.067, and negative albedo anomalies appeared over western Greenland. Lelouch et al. [3] evaluated the VeGeTation (VGT) and European polar system Ten-day surface ALbedo albedo products using the ground observations, Meteosat second generation Ten-day ALbedo Reprocessed/Near Real-Time, and MODIS products. The results indicated that there was a great deviation of VGT when compared with MODIS, and the accuracy of the albedo over most land cover types performed badly. Lu et al. [4] evaluated the performance of four albedo products including GLASS, GlobAlbedo, quality assurance for essential climate variables (QA4ECV) project, and MCD43GF, and proved that the accuracy of albedo products decreased over the area with high aerosol optical depth (AOD), cloud cover, and sparse vegetation. Although these validation works provide the “impression of accuracy” of satellite albedo products, they did not deal with these errors embedded in satellite products.
In order to improve the application accuracy of satellite albedo products, it is necessary to correct the errors of satellite albedo products. Several commonly used methods have been applied to the error correction of various satellite products. The mean-based approaches including linear scaling [5] and local intensity scaling [6] are simple and feasible for the precipitation-related variables correction. The basic idea is to use the ratio between the monthly mean of the data to be corrected and reference value as the correction factor, which is then used to correct the daily data in the corresponding month. Nevertheless, such a method will eliminate the extreme value by averaging, and is not suitable for areas with a higher frequency of extreme events. The cumulative distribution function (CDF) based method adjusts the distribution of the data to make the overall range of the data close to the reference value. It has been widely used in the error correction of precipitation [7] and soil moisture [8]. The CDF-based method can not only correct the average deviation, but also correct the error of data distribution shape. KF was first proposed by mathematician Kalman in 1960 for state estimation of the stochastic process. It corrects the data based on the optimal estimation theory. Specifically, the minimum variance of analysis error is taken as the optimal standard, and the optimal result is achieved through recursive processing (i.e., prediction and correction) [9]. Machine learning algorithms have also played an active role in product correction in recent years [10], [11]. The basic idea is to correct the product by establishing the relationship between the independent variables (including the satellite product to be corrected and its related auxiliary variables) and the target variable (i.e., reference values corresponding to this satellite product). The key of this method is to select the most appropriate algorithm and auxiliary variables according to the characteristics of the variables to be corrected.

Although different error-correction models have been developed, their performances have not been comprehensively evaluated and compared. Furthermore, it is important to note that the reference value in the current error-correction works generally with the in situ observations or the interpolated information from in situ observations. Nevertheless, both of them cannot represent the pixel scale ground “truth” either because there is a significant scale difference between in situ and satellite-based observations, or because the interpolated data may not represent the real situation of the study region due to the widely distributed spatial heterogeneity [12], [13], [14]. Consequently, the error-correction models did not play the maximum benefits. Hence, this article aims to evaluate the error-correction models (i.e., RF, CDF, and KF) based on the pixel scale ground “truth”. The latter is achieved by upsampling in situ measurements from the “point” scale to the coarse pixel scale. MCD43A3 albedo product was used as an example of satellite products to be corrected due to its widespread use. The findings in this article shall provide a reference for the selection of the most appropriate model to correct the error of satellite albedo products.

II. STUDY AREA AND EXPERIMENTAL DATASETS

A. Study Area

Heihe River Basin (HRB) is the second-largest inland river basin in China, covering a 1.43 × 10^6 km^2 area within the scope between 37.84°N–38.98°N and 98.94°E–101.12°E. Complex ecological landscapes are distributed from upstream, midstream, including glaciers, alpine grasslands, alpine meadows, artificial oasis, deserts, and natural oasis, which present typical surface characteristics of inland river basins in arid areas of China. The integrated observation network deployed in HRB has provided long-time series in situ measurements of various surface parameters for the study of different earth system science researches [15]. In this article, 13 automatic weather stations (AWS) located in the upper and middle reaches have been involved. These stations are featured by different land cover types and climate characteristics, with elevations ranging from 1500–4000 m. These factors make the study area significantly heterogeneous. The spatial distribution of AWSs is shown in Fig. 1, and the details of each station are presented in Table I.

B. Data

1) In Situ Site Observation: In situ-based albedo was collected from the 13 in situ sites, as shown in Fig. 1. Each site was equipped with two Kipp and Zonen Radiometers to measure the upward and downward shortwave radiations (300–2800 nm) every 10 min. And the surface albedo can be obtained by calculating the ratio of the upward and downward radiations. To keep consistent with the transit time of MODIS and obtain relatively stable radiation values, we calculated the average upward and downward radiations around local solar noon (+30 min) and make a ratio to generate the albedo on a daily basis. In this article, the data covering the period from January 1, 2013–December 31, 2014, were used to capture the dynamic variation of surface albedo over the whole seasonal cycle.

2) MCD43A3 Albedo Product: MCD43A3 (V006) is the most widely used surface albedo product. This product adopts the bidirectional reflectance distribution function model driven by the semiempirical linear kernel to retrieve the bidirectional reflectance under any solar irradiation and satellite observation conditions. Then the black sky albedo (BSA) can be obtained by integrating the bidirectional reflectance in the observation hemisphere, and the white sky albedo (WSA) can be obtained by double integrating the bidirectional reflectance in the incident hemisphere and the outgoing hemisphere. MCD43A3 used in this article has a spatial resolution of 500 m and a temporal resolution of daily. It provides the BSA and WSA of 1–7 bands at local solar noon, as well as three broad-band albedos of visible (300–700 nm), near-infrared (700–3000 nm), and shortwave (3000–5000 nm) bands. The shortwave albedo is employed in this article because its spectral range is comparable to the in situ-based albedo [1]. We extracted the quality-controlled BSA and WSA corresponding to the 13 AWSs from 2013 to 2014. Since in situ measurements provide blue sky albedo, the BSA, and WSA of MCD43A3 were linearly combined using the sky diffuse light ratio [16]

\[
\alpha = \alpha_{\text{WSA}} \times r + \alpha_{\text{BSA}} \times (1 - r)
\]

where \( \alpha \) is the MCD43A3 blue sky albedo; \( \alpha_{\text{WSA}} \) and \( \alpha_{\text{BSA}} \) represent WSA and BSA, respectively; \( r \) is the actual sky diffuse light ratio, which can be obtained using (2) as suggested by.
Fig. 1. Land cover map of the Heihe River Basin and the distribution of AWSs.

### TABLE I

| Scope    | Site    | Latitude(°) | Longitude(°) | Elevation(m) | Land Cover       |
|----------|---------|-------------|--------------|--------------|------------------|
| Upstream | ASP     | 38.0473     | 100.4643     | 3033         | Grasslands       |
|          | ASN     | 38.0898     | 100.5204     | 3529         | Grasslands       |
|          | ACO     | 37.9841     | 100.4108     | 3536         | Grasslands       |
|          | DSL     | 38.8399     | 98.9406      | 3739         | Grasslands       |
|          | EBO     | 37.9492     | 100.9151     | 3294         | Grasslands       |
|          | HZS     | 38.2254     | 100.1918     | 2612         | Croplands        |
|          | HCG     | 38.0033     | 100.7312     | 3137         | Grasslands       |
|          | JYL     | 37.8364     | 101.116      | 3730         | Grasslands       |
| Midstream| DMS     | 38.6555     | 100.3722     | 1556         | Croplands        |
|          | BJT     | 38.915      | 100.3042     | 1562         | Bare exposed rock|
|          | SSW     | 38.7892     | 100.4933     | 1594         | Bare exposed rock|
|          | ZYS     | 38.9751     | 100.4646     | 1460         | Croplands        |
|          | HZZ     | 38.7652     | 100.3186     | 1731         | Bare exposed rock|

Stokes and Schwartz [17]

\[ r = 0.122 + 0.85 \times \exp\left(-4.8 \times \cos \theta\right) \]  \hspace{1cm} (2)

where \( \theta \) is the solar zenith angle at local solar noon.

3) Auxiliary Datasets: The auxiliary data were mainly used to establish the relationship between the MCD43A3 and pixel scale ground “truth” when using the machine learning-based model. The normalized difference vegetation index (NDVI), solar elevation angle (SEA), digital elevation model (DEM), land cover types, temperature, soil moisture, and snow, which are related to albedo were introduced [18].

The NDVI information was extracted from the MOD09GA_006_NDVI dataset from the Google Earth Engine (GEE) platform. It was calculated based on the near-infrared band and red band of the MOD09GA surface reflectance product, with a spatial resolution of 500 m. SEA was calculated according to the longitude, latitude, altitude, and time of each in situ station. DEM was obtained from the NASA Shuttle Radar Topographic Mission digital elevation dataset, which provides high-quality elevation data at near-global scope with a spatial resolution of 30 m. The land cover types were extracted from the global land cover dataset FROM-GLC-2010, which divided the land surface into 9 first class and 25 second class on a 30 m resolution scale. The second class was used to describe the land cover type of each in situ station in this article. Temperature data was extracted from the ERA5 Daily aggregates dataset through GEE, which is the latest climate reanalysis product produced by ECMWF/Copernicus Climate Change Service. Soil moisture was sourced from the work of Chai et al. [19], which provided the daily soil moisture content covering the Qilian Mountains. Snow information was acquired from MOD10A1 V006, which has been widely used and validated. To make the spatial resolution of auxiliary data consistent with the MCD43A3 and pixel scale ground “truth”, the elevation data was aggregated...
from 30–500 m with the mean aggregation method; the land cover data was aggregated from 30–500 m with the nearest neighbor resampling method; the temperature and soil moisture data were resampled from 25000–500 m with the bilinear interpolation method. It is important to point out that although these auxiliary datasets inevitably suffer from uncertainties due to the limitation of the retrieval algorithm or the original satellite observation errors, such uncertainties can be reduced in the resampling process as the small scale uncertainties were canceled out in the aggregation process [20]. Furthermore, in the model training and application, the same data sources can also enable the accuracy and robustness of the final corrected MCD43A3 V006.

III. METHODOLOGY

The performances of three error-correction models, namely RF, CDF, and KF, were evaluated due to their wide applicability. To achieve this, 70% of the data (i.e., the matched MCD43A3, the pixel scale ground “truth”, and the auxiliary data) were used to train the error-correction models, which is defined as the training dataset. And the remaining 30% of the data were used to validate the error-correction models, which is defined as the validation dataset.

A. Acquisition of Pixel Scale Ground “Truth”

The pixel scale ground “truth” were obtained from the single in situ site measurements and the upscaling method proposed by Wu et al. [21]. The ability of the upscaling model had been confirmed to be effective, producing in situ-based coarse pixel albedo with an RMSE of about 0.01 and a correlation coefficient larger than 0.81 on the 1 km pixel scale. The basic idea of the upscaling method is to derive the upscaling coefficient using time series clear-sky high-resolution albedo maps by assuming that they can accurately capture the temporal and spatial variation characteristics of surface albedo within a coarse pixel scale. The high-resolution albedo was used to simulate the subpixel albedos with the coarse pixel. In this study, Huanjing (HJ) albedo maps from 2010–2012 are processed to derive the upscaling coefficients due to their high spatial resolution of 30 m and higher update frequency. First, the regression relationship between the HJ pixel albedo time series corresponding to the in situ site and the other HJ pixel albedo time series was established

\[
\begin{align*}
\theta_{HJ}(x, y, d) &= W(x, y)^T \theta_{HJ\_insitu}(d) \\
\theta_{HJ\_insitu}(d) &= \left[1, \theta_{HJ\_insitu}(d)\right]^T
\end{align*}
\]

where \(x\) and \(y\) represent the location of HJ pixel within the coarse pixel, and \(d\) is the date of the HJ albedo map. \(W\) denotes the vector of upscaling coefficients to be derived. Then, a cost function \(J\) was established by combining the data through the whole time series in order to enable a robust estimation

\[
J = \text{Minimize} \left\{ \sum_{d=1}^{L} \left[ \theta_{HJ}(x, y, d) - W(x, y)^T \theta_{HJ\_insitu}(d) \right]^2 \right\}
\]

(4)

where \(L\) denotes the length of HJ albedo time series. Then, \(W\) was obtained by minimizing the cost function through the ordinary least squares algorithm

\[
W(x, y) = \left( \theta_{HJ\_insitu}(d) \theta_{HJ\_insitu}^T \right)^{-1} \theta_{HJ\_insitu}(d) \theta_{HJ\_truth}(x, y, d)
\]

(5)

It is important to note that the \(W\) was calculated for each HJ pixel within the coarse pixel. Once it was determined, they can be applied to in situ site measurements (given as \(\theta_{in\_site}\)), and the in situ-based reporting albedo (denoted as \(\theta_{in\_site\_HJ}(x, y, d)\)) over the location corresponding to the HJ pixel on date \(d\) can be determined

\[
\theta_{in\_site\_HJ}(x, y, d) = W(x, y)^T \theta_{in\_site}(d).
\]

(6)

Then, the pixel scale ground “truth” (given as \(\theta_{in\_site\_truth}\)) can be calculated by aggregating \(\theta_{in\_site\_HJ}(x, y, d)\) within the coarse pixel

\[
\theta_{in\_site\_truth}(d) = \frac{\int \int f_{PSF}(x, y) \theta_{in\_site\_HJ}(x, y, d)}{\int \int f_{PSF}(x, y)}
\]

(7)

where \(D\) denotes the spatial extent of the coarse pixel, and \(f_{PSF}\) denotes the point spread function (PSF) of the coarse pixel satellite albedo, which is characterized by Peng et al. [22].

It is undeniable that the \(\theta_{in\_site\_truth}\) is not the absolute truth on the pixel scale due to the limitations of the upscaling method. But it is the best approximation of the absolute truth, which is much better than using in situ observations or the interpolated information from in situ observations. Because in situ measurements cannot represent the pixel scale ground truth due to the spatial scale mismatch and spatial heterogeneity. It is noteworthy that it was the same datasets of the pixel scale ground “truth” and MCD43A3 were used for the evaluation of the three error-correction models. The consistency of the datasets can enable that the difference in the accuracy indicators was merely attributed to the different performances of these models.

B. Error-Correction Models

1) Random Forest Algorithm: The random forest (RF) algorithm is selected because it has the advantage of strong generalization ability, high speed, and precision. It has been widely used in remote sensing domains and proved to be efficient and robust. It is an ensemble of large sets of classification and regression trees [23]. The construction of the forest is mainly as follows. First, \(n\) samples are randomly selected with the bootstrap sampling method from the original sample dataset, and \(k\) sample subsets are obtained by \(k\) times of sampling. Second, for each sample subset, the optimal attribute is selected from the randomly \(m\) selected attributes for splitting, and a single tree is set up. Finally, the combination of the \(k\) regression trees established in the previous step is the random forest. The prediction result of RF is the averaging of outputs from all trees,
which will smooth the fluctuations between different trees and yield more accurate prediction results.

The RF-based error-correction model was carried out as follows. First, the pixel scale ground “truth”, the auxiliary datasets (i.e., NDVI, SEA, DEM, land cover, temperature, soil moisture, and snow), and MCD43A3 were input to the RF model. The pixel scale ground “truth” was used as the dependent variables, and the auxiliary datasets and MCD43A3 were used as independent variables. In order to strengthen the robustness of the RF model, 90% of the training dataset was used to model the relationship (defined as subtrain dataset) and the remaining 10% was used to test the performance of the trained models (denoted as subtest dataset). It should be noted that the numerical ranges of different variables are quite different, which may affect the sample distribution. In order to eliminate such an influence, the auxiliary variables were normalized using the Z-score normalization method [24]. Second, the performance of the RF-trained model was evaluated with the subtest dataset. It is important to note that such an evaluation is only a preliminary test of the RF model. The performance of the three error-correction models will be evaluated and compared based on the validation dataset.

2) CDF Matching: CDF matching method is an enhanced nonlinear technology and is generally used to correct the deviation between different datasets [25]. It was first proposed by Calheiros and Zawadzki [26], who had proved its feasibility in obtaining the probability relationship between reflectance and rainfall rate. Currently, this method has been widely used in the correction and fusion of different datasets of the same variable such as soil moisture or precipitation. The basic principle of the CDF matching is as follows: CDF refers to the sum of the probabilities of random variables falling in a certain interval of sample space, and it is the integral of the probability density function (8).

Then, the cumulative distribution curve of a dataset is adjusted to match that of the reference dataset with the algorithm of Yao et al. [8]. After the CDF adjustment, the discrepancies in their accuracies can be corrected

\[ F(m) = \Phi(x \leq m) \]  

where \( F(m) \) represents the sum of the probabilities of all values less than or equal to \( m \) for the random variable \( x \).

In this article, we perform the CDF correction of MCD43A3 albedo using the pixel scale ground “truth” of each site in the training dataset. Owing to the limited number of training data during the experimental period over each in situ site, MCD43A3 and pixel scale ground “truth” are not segmented in this article, and the polynomial fitting function was adopted. By adjusting the degree of polynomial from 1–4 (with an interval step of 1), the optimal degree of the polynomial will be determined according to the coincidence of the CDF curves between the corrected MCD43A3 and the pixel scale ground “truth”. And the optimal polynomial coefficients derived will then be applied to the validation dataset for each in situ site. Finally, the CDF-based corrected MCD43A3 can be calculated.

3) Kalman Filtering: The basic idea of the KF method is to update the estimation of state variable by using the estimated value of the previous time and the observed value of the current time, and then obtain the estimated value of the current time. Through an iterative process, the optimal estimated value can be obtained using the minimum mean square error as the best estimation criterion. Before Kalman filtering, we need to calculate the mean field bias (denote as X) between the MCD43A3 and pixel scale ground “truth” according to (9), which can be used as the correction factor of MCD43A3. If MCD43A3 and pixel scale ground “truth” are both accurate, X should be constant and be equal to 1. However, X always changes with time due to the influence of various interference factors. Therefore, X should be updated with the KF method

\[ \hat{X}_t = \hat{X}_{t-1} \]  
\[ P_t = P_{t-1} + Q \]  

where \( X_t \) and \( S_t \) represent the pixel scale ground “truth” and MCD43A3 albedo, respectively.

The two main steps are time update and measurement update, which are also known as the prediction and correction. The time update process [(10), (11)] was to establish a priori estimation for the current state, i.e., calculating the prediction value of X and the corresponding error covariance. In this way, the priori estimation for the next time state can be constructed

\[ \hat{X}_t = \hat{X}_{t-1} \]  
\[ P_t = P_{t-1} + Q \]  

where \( \hat{X}_t \) denotes the priori estimation of the predicted variable of X at time t; \( P_t \) represents the variance of the priori estimation error at time t; and \( Q \) indicates the process error variance.

Then, the measurement update equations [(12)–(14)] were used to obtain the posterior estimation of the current state, i.e., improving the prior estimation of X and its estimation error

\[ K_t = P_t / (P_t + R) \]  
\[ \hat{X}_t = \hat{X}_{t-1} + K_t \times (X_t - \hat{X}_t) \]  
\[ P_t = P_t - K_t \times P_t \]

where \( K_t \) denotes the Kalman gain factor at time t; \( R \) is the measurement error variance; \( \hat{X}_t \) represents the optimal estimation at time t, namely, the optimized correction factor we need; and \( P_t \) is the variance of the posteriori estimation error at time t.

During the process of KF-based correction, the two parameters, namely the process error variance (Q) and the measurement error variance (R) were adjusted to achieve the best filtering result. After obtaining the \( \hat{X}_t \) corrected by KF, we arithmetically average the \( \hat{X}_t \) to act as the correction factor for each in situ site to correct the MCD43A3 albedo on the validation dataset.

C. Performance Evaluation

When the error-correction models based on three algorithms were determined, they can be applied to MCD43A3 in the validation dataset, and the corrected MCD43A3 will be compared to the matched pixel scale ground “truth” in order to check the performance of the three error-correction models. Four statistical indicators including RMSE, RRMS (relative RMSE), Bias (mean bias), and R² (coefficient of determination) are introduced to check the closeness and consistency of the correction results.
with respect to the pixel scale ground “truth”

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N} (C_i - R_i)^2}{N}}$$  \hspace{1cm} (15)$$

$$\text{RRMSE} = \frac{\sqrt{\sum_{i=1}^{N} (C_i - R_i)^2 / N}}{R} \times 100\%$$ \hspace{1cm} (16)$$

$$\text{Bias} = \sum_{i=1}^{N} \frac{(C_i - R_i)}{N}$$ \hspace{1cm} (17)$$

$$R^2 = \frac{\left[ \sum_{i=1}^{N} (C_i - C) (R_i - R) \right]^2}{\sum_{i=1}^{N} (C_i - C)^2 \sum_{i=1}^{N} (R_i - R)^2}$$ \hspace{1cm} (18)$$

where $C_i$ and $R_i$ are the corrected MCD43A3 with three error-correction models and the pixel scale ground “truth”, respectively. And $C$ and $R$ are their averaged values, respectively. $N$ is the number of the total samples.

The overall performance of the three error-correction models was first evaluated by combining the results of all in situ sites in the validation dataset. Then, the performances of these models over each site are compared in order to investigate the stability of the performance of these models. The performance of the three error-correction models will be evaluated from two aspects. First, a preliminary evaluation of the RF-based error-correction model was based on the subtest dataset and the performance of the CDF- and KF-based error-correction models was preliminarily checked on the training dataset. Second, for a fair comparison of the performance of these three error-correction models, we have adopted the unified validation dataset.

### IV. RESULTS AND DISCUSSION

#### A. Overall Performance of the Three Error-Correction Models

The training of the RF-based error-correction model was based on the subtraining dataset. The well-trained model was first evaluated using the subtest dataset. Fig. 2 shows the scatter plot between the RF-corrected MCD43A3 and the pixel scale ground “truth”. It can be seen that the RF-corrected MCD43A3 agrees well with the pixel scale ground “truth” with a high $R^2$ of 0.835. Most data points fall around the 1:1 line with a small RMSE of 0.025 and Bias of -0.003. Hence, it can be considered that the RF-based error-correction model is reasonable in correcting the Bias of MCD43A3. Nevertheless, it is important to note that there is one outlier that deviates far from the 1:1 line. This may be caused by the overcorrection or the insufficient correction of the RF-based error-correction model. This is understandable since the performance of the RF model is influenced by many factors such as the selection of machine learning models, the selection of auxiliary variables, the source and accuracy of the variable datasets, and the selection of training samples [27]. To further explore the relative importance of these albedo-related auxiliary variables in correcting the errors of MCD43A3, the increase of mean square error (MSE) provided by the RF-based error-correction model was adopted as the indicator. As shown in Fig. 3, the elevation, MCD43A3, and NDVI were more important than the others in establishing the error-correction model. The largest influence of elevation on MCD43A3 correction can be attributed to the broad distribution scope of these AWSs, with large ranges of elevation from about 1500–4000 m. The important role of MCD43A3 is understandable because it provides the albedo values close to the pixel scale ground “truth”. The important role of NDVI can be explained by the wide range of NDVI values in the training datasets given that the in situ sites (see Fig. 1) were covered by grassland, cropland, and bared exposed rock. The role of soil moisture in establishing the MCD43A3 error-correction model is almost as important as temperature. The importance of land cover types and snow is not as much as we expected. This can be attributed to the fact that the number of land cover types is very few, as shown in Table I, and that the sample size with snow cover in the training dataset is very small. Hence, the results presented here may be not transferable to other regions given the limited number of in situ sites.

The CDF curves of the MCD43A3, the CDF-corrected MCD43A3, and the pixel scale ground “truth” are presented in Fig. 4. In order to test the influence of the degree of the polynomial, the results with 1 to 4 polynomial degrees were shown. Here, we only present the results of EBO for conciseness. It can be seen that the CDF curves of MCD43A3 and pixel scale ground “truth” show large difference, indicating the necessity of conducting such an error correction. After the CDF
TABLE II
SUMMARY STATISTICS (RMSE, RRMSE, AND R²) OF THE CORRECTED MCD43A3 BASED ON THE THREE ERROR-CORRECTION MODELS IN THE LOW, MEDIUM, AND HIGH ALBEDO AREA, RESPECTIVELY

| Range  | N | MCD43A3         | RF-corrected | CDF-corrected | KF-corrected |
|--------|---|-----------------|--------------|---------------|--------------|
|        |   | RMSE      | RRMSE | R²   | RMSE      | RRMSE | R²   | RMSE      | RRMSE | R²   |
| low    | 276 | 0.017     | 10.0%  | 0.370 | 0.019     | 10.6%  | 0.500 | 0.017     | 10.0%  | 0.518 | 0.018 | 10.2%  | 0.479 |
| median | 447 | 0.041     | 16.5%  | 0.293 | 0.017     | 6.8%   | 0.739 | 0.017     | 6.8%   | 0.758 | 0.021 | 8.3%   | 0.622 |
| high   | 6   | 0.175     | 37.0%  | 0.763 | 0.111     | 23.4%  | 0.640 | 0.141     | 29.8%  | 0.860 | 0.163 | 34.5%  | 0.925 |
| Overall| 729 | 0.037     | 16.8%  | 0.585 | 0.020     | 9.1%   | 0.849 | 0.021     | 9.6%   | 0.823 | 0.025 | 11.0%  | 0.764 |

The comparison of the performance of these three error-correction models (i.e., RF, CDF, and KF) based on the validation dataset is shown in Fig. 6. For comparison purposes, the scatter plot between MCD43A3 and the pixel scale ground “truth” is also shown [see Fig. 6(a)]. It can be found that the MCD43A3 does not agree with pixel scale ground “truth” very well. The spatial-temporal variation of surface albedo was not captured by MCD43A3 very well and it is especially true within the median range of albedo (≈0.2–0.4), indicated by the low R² of 0.585. Furthermore, MCD43A3 deviates from the pixel scale ground “truth” with a large RMSE of 0.037 and a large Bias of -0.018. And the errors of MCD43A3 are nonignorable since the RRMSE reaches 16.8% (see Table II). This is another confirmation of the necessity of correcting the satellite albedo product for better application.

After error correction, the corrected-MCD43A3 based on three models all presents a great improvement regarding its closeness and consistency with the pixel scale ground “truth” [see Fig. 6(b), (c), and (d)]. The RMSE decreased from 0.037–0.020, 0.021, and 0.025 for RF, CDF, and KF-based error-correction models, respectively. And R² increased from 0.585–0.849, 0.823, and 0.764 for RF, CDF, and KF-based error-correction models, respectively. It is noteworthy that, compared to MCD43A3, the most significant change of the corrected MCD43A3 is the Bias, which decreases from -0.018–0.004, -0.001, and -0.001 for RF, CDF, and KF-based error-correction models, respectively. These results suggest that the three error-correction models generally work well in correcting MCD43A3 based on the pixel scale ground “truth”.

When it comes to the comparison of the overall performance between these three models, the RF-based error-correction model shows the best performance with the smallest RMSE and the highest R². The CDF-based error-correction model follows, and the KF-based error-correction model ranks the last indicated by the largest RMSE and lowest R². Nevertheless, it is important to point out that the performance of the three error-correction models depends on the situation. Table II provides the statistical indicators in the low, median, and high albedo areas. In the high albedo area (≥0.4), KF-corrected MCD43A3 is the most consistent with the pixel scale ground “truth” with the highest R² of 0.925, followed by the CDF- and RF-corrected MCD43A3, whose R² are 0.860 and 0.640, respectively. RF-corrected MCD43A3 performs best regarding the deviation from the pixel scale ground “truth”, indicated

adjustment, the CDF-corrected MCD43A3 shows a similar CDF curve to that of pixel scale ground “truth”, demonstrating the capability of the CDF-based error-correction model in dealing with the discrepancy between MCD43A3 and pixel scale ground “truth”. It is noteworthy that there is a slight difference between the CDF-corrected MCD43A3 with different degrees of the polynomial, indicating the non-negligible effect of the degree of the polynomial. In this article, the polynomial with the best agreement between CDF-corrected MCD43A3 and pixel scale ground “truth” will be employed to correct MCD43A3.

Fig. 5(a) presents the scatter plot between the pixel scale ground “truth” and MCD43A3. It can be seen that these data points are scattered, with an RMSE of 0.035, Bias of -0.017, and R² of 0.646. Although such an overall accuracy is reasonable in some applications, there is a considerable number of data points with large errors. For instance, in relatively higher albedo areas (larger than 0.3), the difference between them can even reach larger than 0.3. After MCD43A3 was corrected with the Kalman filter, the KF-corrected MCD43A3 agrees with pixel scale ground “truth” very well [see Fig. 5(b)]. The RMSE decreased from 0.035–0.011 and R² increased from 0.646–0.958. Particularly, the Bias decreased from -0.017–0. Therefore, it can be concluded that the KF-based error-correction model can significantly improve the accuracy of MCD43A3, and it is especially true in high albedo area given the fact that almost all points are closer to the 1:1 line.

Fig. 4. CDF curves of MCD43A3 (yellow curve), the CDF-corrected MCD43A3 (pink, purple, green, and red curves), and the pixel scale ground “truth” (blue curve). (1)–(4) in the legend indicate the results with polynomial degrees from 1 to 4, respectively.
Fig. 5. Scatter plot between the MCD43A3 and pixel scale ground “truth” (a), as well as the scatter plot between KF-corrected MCD43A3 and pixel scale ground “truth” (b). This result is obtained using the training dataset.

Fig. 6. Scatterplots between MCD43A3 (a), RF-corrected (b), CDF-corrected (c), and KF-corrected (d) MCD43A3 and the pixel scale ground “truth” based on the validation dataset.

by the smallest RMSE of 0.111, CDF-corrected MCD43A3 follows with the RMSE of 0.141, and KF-corrected MCD43A3 ranks last with the RMSE of 0.163. In the median albedo area (∼0.2–0.4), RF- and CDF-corrected MCD43A3 show nearly equivalent performance with the same RMSE of 0.017 and quite close R² of 0.739 and 0.758, respectively. KF-corrected MCD43A3 performs the worst with the largest RMSE of 0.021 and lowest R² of 0.622. When it comes to the low albedo area (<0.2), the CDF-based error-correction model performs slightly better than the other two models indicated by the highest R² of 0.518. Moreover, the improvements in the median albedo area are much greater than those in the high and low albedo areas, as the RRMSE of the three corrected albedos decreased from 16.5% to 6.8%, 6.8%, and 8.3%, which was more than twice as much as those in high and low albedo areas. This phenomenon indicates that the three models are not ideal for capturing the maximum and minimum of surface albedos. As a whole, the RF-based error-correction model should be given priority if the overall accuracy of corrected results is focused. Nevertheless, the CDF-based error-correction model slightly outperforms RF- and KF-based error-correction models in capturing the minimum albedo.
TABLE III
SUMMARY STATISTICS (RMSE, BIAS, AND $R^2$) OF MCD43A3 AND CORRECTED MCD43A3 BASED ON THE THREE ERROR-CORRECTION MODELS OVER 13 AWSS

| Sites | N  | MCD43A3 RMSE | Bias | $R^2$ | RMSE | Bias | $R^2$ | RMSE | Bias | $R^2$ |
|-------|----|--------------|------|-------|------|------|-------|------|------|-------|
| ASP   | 91 | 0.036        | -0.034 | 0.876 | 0.024 | 0.006 | 0.720 | 0.014 | -0.003 | 0.859 | 0.013 | -0.003 | 0.876 |
| ASN   | 18 | 0.022        | 0.019  | 0.760 | 0.011 | 0.008 | 0.656 | 0.008 | 0.002 | 0.894 | 0.013 | 0.006 | 0.760 |
| ACO   | 11 | 0.048        | -0.047 | 0.020 | 0.012 | 0.009 | 0.550 | 0.008 | 0.001 | 0.072 | 0.012 | -0.006 | 0.020 |
| DSL   | 33 | 0.012        | 0.005  | 0.906 | 0.033 | 0.026 | 0.027 | 0.003 | 0.000 | 0.963 | 0.010 | 0.003 | 0.906 |
| EBO   | 73 | 0.027        | 0.000  | 0.733 | 0.013 | 0.005 | 0.975 | 0.026 | -0.002 | 0.747 | 0.027 | -0.001 | 0.733 |
| HZS   | 72 | 0.039        | -0.034 | 0.749 | 0.014 | 0.003 | 0.860 | 0.020 | -0.002 | 0.738 | 0.019 | -0.001 | 0.749 |
| HCG   | 73 | 0.038        | -0.019 | 0.738 | 0.022 | 0.000 | 0.851 | 0.031 | -0.003 | 0.717 | 0.032 | -0.004 | 0.738 |
| JYL   | 18 | 0.106        | -0.085 | 0.587 | 0.062 | -0.023 | 0.339 | 0.041 | 0.009 | 0.661 | 0.062 | -0.010 | 0.587 |
| DMS   | 81 | 0.021        | 0.004  | 0.668 | 0.021 | 0.005 | 0.675 | 0.015 | -0.002 | 0.801 | 0.021 | 0.002 | 0.668 |
| BJT   | 74 | 0.020        | -0.007 | 0.513 | 0.016 | 0.008 | 0.702 | 0.019 | 0.000 | 0.494 | 0.018 | 0.000 | 0.513 |
| SSW   | 87 | 0.017        | -0.007 | 0.874 | 0.009 | 0.003 | 0.889 | 0.009 | 0.000 | 0.901 | 0.015 | -0.001 | 0.874 |
| ZYS   | 49 | 0.049        | -0.028 | 0.438 | 0.012 | -0.002 | 0.939 | 0.039 | 0.003 | 0.422 | 0.040 | -0.006 | 0.438 |
| HZZ   | 49 | 0.061        | -0.060 | 0.941 | 0.009 | 0.001 | 0.981 | 0.015 | 0.000 | 0.936 | 0.026 | 0.003 | 0.941 |
| All   | 729| 0.037        | -0.018 | 0.585 | 0.020 | 0.004 | 0.849 | 0.021 | -0.001 | 0.823 | 0.025 | -0.001 | 0.764 |

B. Performance of These Error-Correction Models Over Different Sites

In order to further explore the characteristics of these error-correction models, the performances of these models over each in situ site were evaluated. Their quantitative accuracy indicators are summarized in Table III and their histograms were shown in Fig. 7.

The accuracy of MCD43A3 varies with the in situ sites, which is the worst at JYL, with the largest RMSE of 0.106 and Bias of -0.085 [see Fig. 7(a) and (b)]. After the error correction, the RMSEs of the RF-, CDF-, and KF-corrected MCD43A3 reduced, with the value of 0.062, 0.041, and 0.062. And the Bias of the three corrected MCD43A3 has been greatly improved, which decreased to -0.023, 0.009, and -0.010, respectively. In terms of the consistency with the pixel scale ground “truth”, the CDF-corrected MCD43A3 shows the greatest improvement, with $R^2$ increasing from 0.587–0.661. The $R^2$ of KF-corrected MCD43A3 remains the same. Nevertheless, the RF-based error-correction model even plays a negative role regarding the consistency, with $R^2$ reducing from 0.587–0.339.

And similar phenomenon can be found at ASP, ASN, and DSL, with $R^2$ decreasing from 0.876–0.720, from 0.760–0.656, and from 0.906–0.027, respectively. Particularly, the accuracy of the RF-corrected MCD43A3 shows an obvious decline at DSL, with the RMSE increasing from 0.012–0.033 and Bias increasing from 0.005–0.026. This can be attributed to the fact that the samples in the subtrain dataset may be not representative of the situation at this site. This demonstrates that the performance of the RF-based error-correction model is highly dependent on the amount and the representativeness of the samples in the training dataset. When the training sample is not representative or not sufficient, RF does not work well, and CDF performs best in this situation, with the RMSE decreasing from 0.012–0.003 and Bias decreasing from 0.005–0.010 (see Table III).

Over the in situ sites such as ACO, ZYS, and HZZ where MCD43A3 presents relatively large errors, the accuracy of the corrected MCD43A3 has been significantly improved. And the improvement is more significant for Bias than RMSE. The Bias reduced from -0.047–0.009, 0.001, and -0.006 at ACO, from -0.028–0.002, 0.003, and -0.006 at ZYS, and from -0.060–0.001, 0, and 0.003 at HZZ for RF-, CDF-, and KF-corrected MCD43A3, respectively. While the RMSE decreased from 0.048–0.012, 0.008, and 0.012 at ACO, from 0.049–0.012, 0.039, and 0.040 at ZYS, from 0.061–0.009, 0.015, and 0.026 at HZZ for three corrected albedos, respectively. Nevertheless, the improvement of $R^2$ is challenging for these error-correction models. Over ACO and ZYS sites where MCD43A3 shows poor consistency with pixel scale ground “truth”, only the RF-based error-correction model presents a positive role, with the $R^2$ increasing from 0.02–0.55 and from 0.438 to 0.939 at ACO and ZYS, respectively. By contrast, the $R^2$ of CDF- and KF-corrected MCD43A3 is almost unchanged. Over the HZZ where MCD43A3 shows a high consistency with pixel scale ground “truth”, the improvement of the three error-correction models is not obvious.

When it comes to the in situ sites (e.g., DMS, BJT, and SSW) where MCD43A3 presents good accuracy, neither RMSE nor $R^2$ has been improved significantly by these three error-correction models. Such a phenomenon can also be found in Fig. 8, which shows that the improvement of accuracy presents an upward trend with the RMSE of MCD43A3. This demonstrates that the effectiveness of these error-correction models depends on the situation. The improvement is most significant over the sites with large errors of MCD43A3, but relatively small over the sites with good accuracy of MCD43A3. Therefore, the correction models are particularly useful for regions with large errors in satellite products. However, since the accuracy of MCD43A3 is affected by various factors such as the errors of satellite observations, the preprocessing of satellite data, and the limitations and assumptions of the retrieval algorithms, there is no conclusive result about the regions where MCD43A3 suffers from large errors. But these regions can be identified through comprehensive validation works. On the whole, the performance of the three error-correction models varies from site to site. Compared with the other two models, the RF-based error-correction model presents obvious superiority at ZYS, followed by HZZ, EBO,
HCG, and HZS. And the CDF-based error-correction model achieves the largest improvement at DSL, followed by DMS, ASN, and ACO. When it comes to the KF-based error-correction model, the RMSE of the corrected MCD43A3 is almost equal to or greater than the other two corrected MCD43A3 except for ASP. This demonstrates that KF is not as good as the other two models in improving the accuracy of MCD43A3. What’s more, the $R^2$ of KF-corrected MCD43A3 at each site is almost equal to that of MCD43A3, indicating the incapability of the KF-based error-correction model in improving the consistency of MCD43A3 with pixel scale ground “truth”.

In summary, the three models have the most significant improvement in Bias, followed by RMSE, and $R^2$ ranks last. Despite the overall improvement of $R^2$ of these three models, it may be reduced by these models at some sites (e.g., ASP, ASN, DSL, and JYL for the RF model, and ASP, HZS, HCG,
BJT, ZYS, and HZZ for the CDF model). Therefore, these error-correction models should be utilized with caution if the time variation trends rather than the albedo values themselves are focused.

Fig. 9 shows the boxplots of RMSEs, Biases, and $R^2$ of the four kinds of albedo. From Fig. 9(a), it is clear that the RMSE of RF, CDF, and KF-corrected MCD43A3 is significantly reduced. The reduction of RMSE is the most significant for RF-corrected MCD43A3, followed by CDF-corrected MCD43A3, and KF-corrected MCD43A3 ranks last. Furthermore, it is noteworthy that the RMSEs of RF-corrected MCD43A3 are more centralized, indicating the stability of the performance of the RF-based error-correction model in improving the accuracy of MCD43A3. Although the CDF-corrected MCD43A3 presents a smaller RMSE than that of KF, the stability of the CDF-based error-correction model is inferior to that of KF.

The three error-correction models all show significant improvements in Bias [see Fig. 9(b)]. Nevertheless, the RF-corrected MCD43A3 seems to be slightly overestimated while the KF-corrected MCD43A3 is slightly underestimated. The CDF-corrected MCD43A3 shows an obvious advantage over the other two correction models given that its Bias is almost all concentrated at about 0. Furthermore, the CDF presents the most stable behavior in correcting Bias, indicated by the narrowest boxplot. The improvement of $R^2$ is not significant for the three error-correction models, and it is particularly true for the KF-based error-correction model [see Fig. 9(c)]. The CDF-corrected MCD43A3 presents a slightly higher $R^2$ than the RF-corrected MCD43A3. In summary, the RF-based error-correction model shows an advantage over the other two models in reducing the RMSE of MCD43A3. But CDF performs the best in reducing the Bias of MCD43A3. KF is not as good as the other two models in correcting MCD43A3.

As is well-known, RF is a data-driven machine learning algorithm, and its performance is highly dependent on many other factors (e.g., the amount and representativeness of training data). However, the experimental data in this article only covers the time period of two years. The amount of the training data may not fully establish the relationship between the pixel scale ground “truth” and satellite products. As a result, the RF-corrected MCD43A3 may not reach its optimized point. Even so, the RF-based error-correction model still achieves the lowest RMSE in most of the sites. It is believed that RF will obtain better results with the support of more abundant data. Although CDF shows good performance in correcting Bias of MCD43A3, better correction effect can be achieved if more data is available to support segmented matching. The inferiority of the KF-based error-correction model is mainly due to the fact that its results are highly dependent on the determination of process and measurement error variance.

V. CONCLUSION

To further improve the quantitative application level of remote sensing and make satellite albedo products a reliable information source to meet the severe challenges such as global change,
it is urgent to correct the error of albedo products. Although lots of studies have done error correction of different satellite products such as precipitation and AOD, the performance of different error-correction models has not been evaluated and compared. Furthermore, most studies are based on in situ measurements rather than pixel scale ground “truth”, and spatial scale mismatch between in situ and satellite measurements will introduce considerable uncertainties to the results. In response to this challenge, this article conducts a comprehensive evaluation of three widely used error-correction models (i.e., RF, CDF, and KF) based on the pixel scale ground “truth”, which is derived by upsampling single in situ site measurements to the coarse pixel scale. MCD43A3 was employed in this article as an example to demonstrate the performance of the three error-correction models.

By using 70% of the matched data as the training dataset and 30% of the matched data as the validation dataset, it has been proved that the corrected-MCD43A3 based on three models all presents a great improvement regarding its closeness and consistency with the pixel scale ground “truth”. The overall RMSE decreased from 0.037–0.020, 0.021, and 0.025 for RF, CDF, and KF-based error-correction models, respectively. And R² increased from 0.585–0.849, 0.823, and 0.764 for RF, CDF, and KF-based error-correction models, respectively. The improvement is the most significant in Bias, with values close to 0 after correction. Generally, the overall performance is the best for the RF-based error-correction model, with the smallest RMSE and the highest R². The CDF-based error-correction model follows, and the KF-based error-correction model ranks last. And all three corrected MCD43A3 perform much better in the median albedo area than in the high and low albedo areas.

Nevertheless, it is important to note that the specific performance of these error-correction models depends on the situation. The performance of the RF-based error-correction model is highly dependent on the amount and the representativeness of the samples in the training dataset. It does not work well when the training sample is not representative or not sufficient. In this case, the CDF-based error-correction model performs best. The three models are more adept at improving Bias, followed by RMSE, and R² ranks last. Furthermore, it is found that these error-correction models are particularly useful for the region with large errors of satellite products, with significant improvement of RMSE and R². But the improvements are very slight over the sites with good accuracy. Despite the overall improvement of R² of these three models, it may be reduced by these models at some sites. Hence, these error-correction models should be utilized with caution if the time variation trends rather than the albedo values themselves are focused. What is more, when the stability of the performance was considered, the RF-based error-correction model shows an advantage over the other two models in reducing the RMSE of MCD43A3, and the CDF-based error-correction model performs the best in reducing the Bias of MCD43A3. KF is not as good as the other two models in correcting MCD43A3.

The findings in this article shall provide important guidance in the selection of error-correction models to correct the errors of satellite products. Nevertheless, due to the limited number of in situ sites and the limited experimental period, it is an open issue whether the presented conclusions are transferable to other regions or time periods. Furthermore, we would like to point out that the focus of this article is the comparison of the performance of three error-correction models based on pixel scale ground “truth” rather than the error correction of satellite products on the global scale. And this will be the focus of our future work by obtaining the pixel scale ground “truth” based on the in situ measurements worldwide and then dividing the matched data into different groups according to their land cover types, terrain features, and climatic characteristics in order to train more robust error-correction models.

References

[1] X. Wu et al., “Accuracy assessment on MODIS (V006), GLASS and MuSyQ land-surface albedo products: A case study in the Heihe River Basin, China,” Remote Sens., vol. 10, no. 12, Dec. 2018, Art. no. 2045.
[2] J. Stroeve et al., “Re-evaluation of MODIS MCD43 Greenland albedo accuracy and trends,” Remote Sens. Environ., vol. 138, pp. 199–214, Nov. 2013.
[3] G. Lellouch et al., “Evaluation of two global land surface albedo datasets distributed by the copernicus climate change service and the EUMETSAT LSA-SAF,” Remote Sens., vol. 12, no. 11, Jun. 2020, Art. no. 1888.
[4] Y. Lu et al., “Evaluation of satellite land surface albedo products over China using ground-measurements,” Int. J. Digit. Earth, vol. 14, no. 11, pp. 1493–1513, Jun. 2021.
[5] J. Teng et al., “How does bias correction of regional climate model precipitation affect modelled runoff?” Hydrol. Earth Syst. Sci., vol. 19, no. 2, pp. 711–728, Feb. 2015.
[6] M. Minville et al., “Adaptation to climate change in the management of a Canadian water-resources system exploited for hydropower,” Water Resour. Manag., vol. 23, no. 14, pp. 2965–2986, Mar. 2009.
[7] X. Ji et al., “Evaluation of bias correction methods for APHRODITE data to improve hydrologic simulation in a large Himalayan basin,” Atmos. Res., vol. 242, Sep. 2020, Art. no. 104964.
[8] Y. Xiao, J. Yu, and W. Sun, “Continuous fusion algorithm analysis for multisource remote sensing soil moisture data based on cumulative distribution fusion,” Trans. Chin. Soc. Agriculture Eng., vol. 35, no. 1, pp. 131–137, 2019.
[9] Y. Xu, “Application of Kalman Filter in precipitation estimation by radar,” J. Arid Meteorol., vol. 26, no. 1, pp. 78–82, 2008.
[10] L. Han et al., “A deep learning method for bias correction of ECMWF 24-240 h forecasts,” Adv. Atmos. Sci., vol. 38, no. 9, pp. 1444–1459, Jul. 2021.
[11] H. Chen, L. Sun, R. Cicelli, and P. Xie, “Deep learning for bias correction of satellite retrievals of Orographic precipitation,” IEEE Trans. Geosci. Remote Sens., vol. 60, Aug. 2021, Art. no. 4104611.
[12] X. Wu et al., “Spatial heterogeneity of albedo at subpixel satellite scales and its effect in validation: Airborne remote sensing results from HiWATER,” IEEE Trans. Geosci. Remote Sens., vol. 60, Oct. 2021, Art. no. 4407114.
[13] J. Wen et al., “Characterizing the effect of spatial heterogeneity and the deployment of sampled plots on the uncertainty of ground ‘truth’ on a coarse grid scale: Case study for near-infrared (NIR) surface reflectance,” J. Geophys. Res., vol. 127, no. 11, Jun. 2022, Art. no. e2022JD036779.
[14] X. Wu et al., “Quantification of the uncertainty caused by geometric registration errors in multiscale validation of satellite products,” IEEE Trans. Geosci. Remote Sens. Lett., vol. 19, Aug. 2021, Art. no. 8017905.
[15] X. Li et al., “HIWATER: An integrated remote sensing experiment on hydrological and ecological processes in the Heihe River Basin,” Adv. Earth Sci., vol. 27, no. 5, pp. 481–498, May 2012.
[16] B. Pinty et al., “Coupling diffuse sky radiation and surface albedo,” J. Atmos. Sci., vol. 62, no. 7, pp. 2580–2591, Jul. 2005.
[17] G. M. Stokes and S. E. Schwartz, “The atmospheric radiation measurement (ARM) program: Programmatic background and design of the cloud and radiation test bed,” Bull. Amer. Meteorol. Soc., vol. 75, no. 7, pp. 1201–1222, Jul. 1994.
[18] D. Xiao, F. Tao, and J. P. Moiwo, “Research progress on surface albedo under global change,” Adv. Earth Sci., vol. 26, no. 11, pp. 1217–1224, Nov. 2011.
Jingping Wang received the B.S. degree in geographic information science from Anhui Normal University, Wuhu, China, in 2020. She is currently working toward the M.S. degree in cartography and geographic information system with the College of Earth and Environmental Sciences, Lanzhou University, Lanzhou, China.

Her research interest includes scaling land surface parameters using machine learning models.

Xiaodan Wu received the B.S. degree in geographic information science from the China University of Petroleum, Beijing, China, in 2012, and the Ph.D. degree in cartography and geographic information system from the Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing, China, in 2017.

She is currently a Professor with the College of Earth and Environmental Sciences, Lanzhou University, Lanzhou, China. Her research interests include validation of quantitative remote sensing products over heterogeneous land surfaces.

Rongqi Tang received the B.S. degree in remote sensing science and technology from the Chengdu University of Information Technology, Chengdu, China, in 2021. He is currently working toward the M.S. degree in cartography and geographic information system with the College of Earth and Environmental Sciences, Lanzhou University, Lanzhou, China.

His research interest includes SIF retrieval by models.

Qicheng Zeng received the B.S. degree in geographic information science from Qinghai Normal University, Xining, China, in 2021. He is currently working toward the M.S. degree in cartography and geographic information system with the College of Earth and Environmental Sciences, Lanzhou University, Lanzhou, China.

His research interest includes soil moisture retrieval using active microwave technology.

Zheng Li received the B.S. degree in geographic information science in 2012 from Lanzhou University, Lanzhou, China, where he is currently working toward the M.S. degree in cartography and geographic information system with the College of Earth and Environmental Sciences.

His research interest includes vegetation remote sensing.

Jianguang Wen received the B.S. degree in geographic information system and the M.S. degree in cartography and geographic information system from the Jilin University of China, Changchun, China, in 2002 and 2005, respectively, and the Ph.D. degree in quantitative remote sensing from the Institute of Remote Sensing Applications, Chinese Academy of Sciences, Beijing, China, in 2008.

He is currently a Researcher with the Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing. His research interests include remote sensing radiative transfer mechanism and inversion, and remote sensing experiment in land surface parameters validation.

Qing Xiao received the B.S. degree in applied geophysics and the M.S. degree in cartography and geographic information system from Jilin University, Jilin, China, in 1993 and 1996, respectively, and the Ph.D. degree in quantitative remote sensing from the Institute of Remote Sensing Applications, Chinese Academy of Sciences, Beijing, China, in 2002.

He is currently a Researcher with the Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China. His main research interests include remote sensing radiative transfer mechanism and inversion, and remote sensing experiment in land surface parameters validation.