Improving the Accuracy of Fractional Evergreen Forest Cover Estimation at Subpixel Scale in Cloudy and Rainy Areas by Harmonizing Landsat-8 and Sentinel-2 Time-Series Data

Taixia Wu, Yuting Zhao, Shudong Wang, Hongjun Su, Senior Member, IEEE, Yingying Yang, and Dongzhen Jia

Abstract—Evergreen forest provides essential ecosystem services such as maintaining the balance in carbon and oxygen cycles and air purification. However, under cloudy and rainy weather conditions, it is difficult to obtain optical remote sensing images with high spatial resolution and complete time series. In addition, surfaces underlying the forest canopy can be complex and fragmented. To solve this problem, we developed a new approach (NDVI_{CV}-LS) for mapping urban evergreen forest at the subpixel scale. In order to capture more accurate growth characteristics of evergreen forest, we harmonized Landsat-8 and Sentinel-2 images with cloud cover less than 10% acquired within 1 year to denser the time-series dataset. In view of the time series fluctuation stability of evergreen forest, the coefficient of variation (CV) of the normalized difference vegetation index (NDVI) was used to distinguish evergreen forest from other vegetation. Meanwhile, the annual minimum NDVI (NDVI_{ann-min}) was used as the parameter in a dimidiate pixel model for estimating fractional evergreen forest cover (FVC_{over}). Hefei, a cloudy and rainy subtropical city in China, was selected as a case study to evaluate the validity of the model. The verification results revealed that harmonizing Landsat-8 and Sentinel-2 time-series images to extract evergreen forest improved the overall accuracy by 8% compared with using Landsat-8 images alone, indicating that the NDVI_{CV}-LS model can improve the accuracy of FVC_{over} estimation, especially for areas with complex underlying surfaces under cloudy and rainy conditions.

Index Terms—Coefficient of variation (CV), fractional evergreen forest cover, Landsat-8, Sentinel-2, time series.

I. INTRODUCTION

VARIATION in forest types shape the structure of the forest landscape [1] which largely determines forest ecosystem services and functions. A good understanding of the spatial distribution of different forest types contributes to better forest ecosystem and land space management [2], which then facilitates the protection of wildlife habitats and biodiversity [3], [4]. Evergreen forest is an important part of the subtropical urban ecosystem. Efforts to protect the regional environment and promote sustainable development [5] have to take these forests into account. Evergreen trees contribute to air purification, cooling, humidification, water and soil conservation, and other ecological services [6]–[8]. They play an irreplaceable role in maintaining the carbon balance of urban ecosystems and protecting the urban ecological environment [9]. Compared with deciduous trees, evergreen trees have a larger leaf mass per area and longer leaf life; thus, they have a higher capacity for carbon sequestration and dust retention [10], [11]. Moreover, in evergreen forest, the relative humidity is approximately 7% higher than that on nonforested land in winter [12]. Evergreen forest also effectively mitigates urban heat-island effects [13]. Evergreen trees have high landscape aesthetic value because they can compensate for landscape gaps in areas where deciduous plants grow in winter. Therefore, it is vital to have a comprehensive understanding of evergreen forest distribution.

Remote sensing can acquire real-time, large-scale data; thus, it is commonly employed to obtain basic maps such as maps of land-cover types at regional and global scales [14]. Time-series remote sensing images from satellites are important data sources for monitoring vegetation characteristics. Based on phenological analyses, vegetation characteristics are extracted [15]–[18]. For example, the growth characteristics of evergreen forest remain almost unchanged throughout the year; therefore, they have been extracted over large areas using time-series indices such as the enhanced vegetation index, land surface water index, and normalized difference vegetation index (NDVI) [19]–[21]. Previous studies mostly used time-series data from the moderate resolution imaging spectroradiometer and advanced very high resolution radiometer sensors [22]. These sensors provide the needed temporal resolution for observing growth dynamics. However, the low spatial resolution will cause a lot of errors by mixing multiple objects with different spectral curves and
phenomenological characteristics in one pixel [23], especially in spatially heterogeneous areas which have fragmented landscapes [24], [25]. In recent years, the satellites with moderate to high spatial resolution (e.g., Landsat and Sentinel-2) have developed rapidly and can well express the detailed features of the surface [26]. Medium-resolution satellites can reflect the growth status of vegetation throughout the growing period, if the time series is complete enough. However, evergreen species are mostly distributed in tropical and subtropical regions where cloudy and rainy weather conditions occur frequently. It is hard to acquire images with high spatial resolution and complete time series under cloudy and rainy weather conditions, making it difficult to conduct broad-scale, high-precision dynamic monitoring. Most satellite sensors cannot take high temporal resolution and high spatial resolution images at the same time [27]. Vegetation extraction from urban areas with complex underlying surface requires high spatial resolution images to reduce the influence of mixed pixels. However, the temporal resolution of medium- to high-resolution images cannot meet the requirement of constructing continuous time series under cloudy and rainy weather conditions.

Multisource images fusion is one of the most effective solutions to this problem [28]–[30]. The method not only helps to resolve the contradictions between high temporal resolution and high spatial resolution imagery datasets [31], but also promotes the utilization of current earth observation data [32]. Fusion of datasets with similar characteristics, such as Landsat-8 operational land imager (OLI) and Sentinel-2 multispectral instrument (MSI) images, provides the potential to observe landscape dynamics at a spatial resolution of 30 m. As an example, the harmonized Landsat-8 OLI and Sentinel-2 MSI datasets were used to assess winter wheat yields at regional scale [33], estimate the number and timing of mowing events in Central Europe grasslands [34], and detect irrigated areas in Southern Italy [35]. In the very past few years, despite the growing attention in harmonizing Landsat-8 and Sentinel-2 images, there is a lack of papers on their use for evergreen forest mapping of subtropical urban areas.

The underlying surface in subtropical urban areas is complex and fragmented. Mixed pixels constitute a serious problem. To improve the accuracy of evergreen forest extraction, it is necessary to decompose the mixed pixels. So far, a variety of pixel unmixing models have been developed, including linear, probabilistic, geometric-optical, stochastic, and fuzzy models [36]. The dimidiate pixel model, a linear mixed model, is widely used to estimate fractional vegetation cover (FVC) at the sub-pixel scale [37]–[39].

In this article, we developed a novel approach (NDVI_CVLS) for improving the accuracy of evergreen forest extraction in cloudy and rainy areas at subpixel scale. Landsat-8 and Sentinel-2 data were harmonized to construct a time-series dataset to capture more accurate growth characteristics of evergreen forest. In particular, the time series fluctuation of evergreen forest is relatively stable compared with other vegetation types (i.e., deciduous forests and farmland). We used the coefficient of variation (CV) of NDVI to separate evergreen forest form continuous crop and deciduous forest, which were extremely admixed with low coverage evergreen forest. Then, the annual minimum NDVI (NDVI_{min-min}) was used as a parameter in the dimidiate pixel model to estimate fractional evergreen forest cover (FVC_{ever}). Given the spatial heterogeneity, the study area was divided into areas with and without an impervious surface by the normalized difference built-up index (NDBI). These two regions were assigned different parameters to extract evergreen forest at subpixel scale. A typical subtropical city, Hefei, China, was selected as the study area to evaluate the validity of the model. The results indicated that our model enabled higher accuracy of FVC_{ever} estimation, making it possible to conduct evergreen forest monitoring in subtropical urban areas with heavy cloud cover and rainfall.

II. STUDY AREA AND DATA

A. Study Area

The study area is located in Hefei City, Anhui Province, China, which has a central latitude and longitude of 31.82°N and 117.23°E (Fig. 1). Hefei covers an area of 11 445 km². The terrain in this area is dominated by hills with an altitude of approximately 15–80 m, and most of the slopes in the study area are less than 2°. This area has a typical subtropical humid monsoon climate with four distinct seasons, long frost-free periods and abundant rain. The average annual temperature is approximately 15.7°C, the average annual relative humidity is about 77% and the annual precipitation is approximately 1000 mm. Climatic conditions are suitable for growing evergreen trees. Common native evergreen broad-leaved trees include magnolia, camphor, and holly, whereas evergreen coniferous trees include mason pine and Cryptomeria fortunei, and the main evergreen shrub is osmanthus. However, frequent occurrence of clouds and rain makes it difficult to acquire remote-sensing images with high spatial resolution and complete time series. Taking into account the high degree of urbanization in Hefei, the underlying surface is complex and fragmented. Therefore, the study area is highly representative and applicable to most similar situations.

B. Data Sources and Preprocessing

Images used in this article were obtained from Landsat-8 OLI and Sentinel-2 MSI sensors on days with cloud cover less than 10%. The spatial resolution for Landsat-8 images is 30 m, whereas that for Sentinel-2 is 10 and 20 m (for bands 2, 3, 4, and 8A only; Table I). Sentinel-2 consists of two satellites—A and B. Each satellite has a revisiting period of 10 days, and as a combination, the two satellites have a revisiting period of 5 days, which is shorter than that of Landsat-8 (16 days). All data were downloaded from the United States Geological Survey Global Visualization Viewer website (http://glovis.usgs.gov/). All images were acquired in 2019, with no valid data obtained in February due to cloud cover (Fig. 2).

Given the biases between the two sensors in terms of different spectral band ranges and view geometries, a series of preprocessing steps were applied to obtain a consistent surface reflectance record for the reconstruction of a 30-m time-series dataset [40].
The following is the preprocessing process of reconstructing the time-series dataset.

1) The same atmospheric correction algorithm, LaSRC [41] was applied to both Landsat-8 (L1T) and Sentinel-2 (L1C) images.

2) The cloud and cloud-shadow masking for Landsat-8 was performed by LaSRC internally, and for Sentinel-2 by an adapted version of Fmask [42].

3) To improve the spatial coregistration between the two sensors, geometric resampling and geographic registration were performed by automated registration and orthorectification package [43].

4) In view of the differing solar and view angles between the two sensors, the bidirectional reflectance distribution function effects were corrected [44].

5) A bandpass adjustment (based on a linear fit between equivalent spectral bands) is applied to Sentinel-2 images to match the Landsat-8 spectral responses [35].

III. METHODOLOGY

The workflow for our method is schematized in Fig. 3. First, we reconstructed a time-series dataset through harmonized Landsat-8 and Sentinel-2 images. Next, we constructed an evergreen forest extraction model (NDVI_{CV-LS}). The CV of
Fig. 3. Flowchart representing the process of fractional evergreen forest cover estimation and accuracy assessment. NDVI, normalized difference vegetation index; MVC, maximum value composite; CV, coefficient of variation; GF-1, Gaofen-1.

### TABLE I
**CHARACTERISTICS OF LANDSAT-8 OLI AND SENTINEL-2 MSI BANDS**

| Sensor   | Bands | Wavelength (nm) | Resolution (m) |
|----------|-------|-----------------|----------------|
| Landsat-8 OLI | B2 (B) | 450–515         | 30             |
|          | B3 (G) | 525–600         | 30             |
|          | B4 (R) | 630–680         | 30             |
|          | B5 (NIR) | 845–885    | 30             |
| Sentinel-2 MSI | B2 (B) | 458–523         | 10             |
|          | B3 (G) | 543–578         | 10             |
|          | B4 (R) | 650–680         | 10             |
|          | B8A (NIRn2) | 855–875    | 20             |

B, blue; G, green; R, red; NIR, near-infrared.

There is no table provided in the text.

time-series NDVI was used to separate evergreen forest from other vegetation. $\text{NDVI}_{\text{ann-min}}$ was used as a parameter in the dimidiate pixel model to accurately extract evergreen forest at subpixel scale. We ran a model simulation with high-resolution images and field samples to determine a reasonable threshold value. Finally, root mean square error (RMSE) and mean absolute error (MAE) were computed to explain errors associated with the results. Confusion matrices of different products were also established to compare the accuracy of the NDVI$_{CV}$-LS model with that of other methods.

#### A. Time-Series NDVI Dataset

Although cloud masking had been applied to the harmonized Landsat-8 and Sentinel-2 dataset during preprocessing, it performed poorly under surface and atmospheric states such as bright targets, thin clouds, cloud edges, and hazy conditions [45]. Additional operations were required to reduce cloud and shadow effects.

We applied an interpolation algorithm to the images of adjacent months, that is, we computed their average value, to supplement the missing images caused by cloud cover (e.g., images for February). For images with little cloud cover, interpolation results from adjacent months were used to replace the cloudy parts [25], [46]. During the noncritical growth period (i.e., from July to September), the vegetation characteristics did not change much. The maximum value composite method was used to synthesize the data from these months [47] in order to remove the influence of cloud cover, which could ensure that the CV values were not overly affected. Then, the NDVI was calculated as follows [48]:

$$\text{NDVI} = \frac{\text{NIR} - R}{\text{NIR} + R}$$  \hspace{1cm} (1)

where NIR is the near-infrared band reflectance record, and $R$ is the red band reflectance record. Through the abovementioned steps, we obtained a reconstructed NDVI time-series dataset. Fig. 4 shows that compared with Landsat-8 time-series images, the reconstructed NDVI time series is more complete and can reflect more growth status of vegetation throughout the growing period.

#### B. Coefficient of Variation

The CV is expressed as the ratio of the standard deviation to the mean. It reflects the differences among individual datapoints
in a dataset. The CV of time-series NDVI is sensitive to fluctuations in the phenological cycle of vegetation. For example, deciduous trees sprout in spring and wither in autumn, and crops are sown and harvested at specific dates. CV has the ability to capture these growth fluctuations in the phenological cycle; thus, it can be used to characterize phenological attributes of evergreen forest. In this article, we took the ratio of the standard deviation to the mean of NDVI time series as the CV for each pixel \[49\]. The standard deviation \(\sigma\) and mean \(\mu\) are calculated as follows:

\[
\mu = \frac{1}{n} \sum_{m=1}^{n} (\text{NDVI}_m)
\]

\[
\sigma = \sqrt{\frac{1}{n-1} \left[ \sum_{m=1}^{n} (\text{NDVI}_m - \mu) \right]^2}
\]

where \(\text{NDVI}_m\) is the NDVI associated with image \(m\) and \(n\) is the number of images in the time series. The CV of the pixel in row \(i\) and column \(j\) of the NDVI time series is then calculated as follows:

\[
\text{CV}_{i,j} = \frac{\sigma_{i,j}}{\mu_{i,j}}.
\]

Evergreen forest remains relatively stable throughout their phenological cycle. Compared with farmland and deciduous vegetation, the CV value of a pure evergreen forest is small [Fig. 5(a)]. However, when the coverage of evergreen forest in a certain pixel is low, its \(\text{NDVI}_{\text{ann-min}}\) [around 0.3, Fig. 7(b) and (c)] is almost close to continuous crop and deciduous forest [Fig. 5(b)]. Compared with \(\text{NDVI}_{\text{ann-min}}\), CV is sensitive to small changes [even 10%, Fig. 7(a)] in the proportion of evergreen forest; therefore, it can separate the low-coverage evergreen forest from other vegetation.

### C. Dimidiate Pixel Model

To accurately extract evergreen forest at subpixel scale, we used the dimidiate pixel model to estimate \(\text{FVC}_{\text{ever}}\). The dimidiate pixel model is a linear mixing model \[36\]. The model assumes that two components, background and vegetation, contribute to a pixel signal in terms of \(\text{FVC}\) estimation. The resulting signal, \(S\), as observed by a remote sensor can be expressed as \[50\]

\[
S = S_{\text{veg}} \times \text{FVC} + S_{\text{back}} \times (1 - \text{FVC})
\]

where \(S_{\text{veg}}\) is the signal contribution from the vegetation component and \(S_{\text{back}}\) is the contribution from the background component. Then, FVC is derived as follows:

\[
\text{FVC} = \frac{S - S_{\text{back}}}{S_{\text{veg}} - S_{\text{back}}}.
\]

Because less information is reflected in a single band, NDVI is usually taken as the “\(S\)” to estimate FVC. NDVI is linearly related to vegetation density distribution, which can well reflect the actual situation of vegetation growth status \[51\]. As shown in Fig. 5(b), the \(\text{NDVI}_{\text{ann-min}}\) of evergreen forest was significantly higher than those of the other three types of ground objects.
On this basis, we adopted NDVI$_{ann-min}$ as the signal “S” in the model instead of NDVI, to extract evergreen forest at subpixel scale. Then, FVC$_{ever}$ can be expressed as

$$FVC_{ever} = \frac{NDVI_{ann-min} - NDVI_{ann-min_{back}}}{NDVI_{ann-min_{veg}} - NDVI_{ann-min_{back}}}$$

where NDVI$_{ann-min_{back}}$ is the NDVI$_{ann-min}$ when the pixel does not contain evergreen forest, and NDVI$_{ann-min_{veg}}$ is the NDVI$_{ann-min}$ when the pixel contains only evergreen forest. With this equation, evergreen forest was extracted at subpixel scale.

D. Thresholds Analysis

The key to model performance lies in the reasonable setting of parameters. Compared with deciduous vegetation and farmland, the annual NDVI fluctuation curve of evergreen forest is most stable [52], which means that evergreen forest has the smallest CV (Fig. 5). By setting an appropriate CV threshold, all pixels containing evergreen forest can be filtered out. The evergreen forest classification results obtained by CV contained a large number of mixed pixels, which needed to be accurately extracted at the subpixel level. The CV associated with built-up areas does not differ much from that associated with evergreen forest (Fig. 4). Consequently, classification results may be inaccurate when evergreen forest is extracted using only CV. The NDVI$_{ann-min}$ associated with evergreen forest is significantly larger than that associated with built-up areas. We set NDVI$_{ann-min}$ as the driving variable of the dimidiate pixel model to extract FVC$_{ever}$ from mixed pixels.

To obtain more accurate FVC$_{ever}$ values, the NDBI was used to divide the study area into areas with impervious and nonimpervious surfaces (Fig. 6), to which different parameters of the dimidiate pixel model were assigned. NDBI was calculated as follows [53]:

$$NDBI = \frac{MIR - NIR}{MIR + NIR}$$

where NIR represents the reflectance value of the NIR band and MIR represents the reflectance value of the mid-infrared band.

The Gaofen-1 (GF-1) images used in the study were acquired on November 20, 2019 and December 27, 2019, with a spatial resolution of 8 m in the multispectral bands and 2 m in the panchromatic bands. From the GF-1 images, 200 samples were randomly and uniformly selected. Each sample covered 8 $\times$ 8 pixels with a resolution of 30 m to reduce the registration deviation between GF-1, Landsat-8, and Sentinel-2 images.
The proportion of evergreen forest in each sample was estimated, as well as the average CV and NDVI$_{\text{ann-min}}$. Simulation analysis was carried out in combination with 29 field samples and visual interpretation via Google Earth. Thresholds of three model parameters—CV, NDVI$_{\text{ann-min,back}}$, and NDVI$_{\text{ann-min,veg}}$—were determined by mixed pixels with different proportions of evergreen forest.

E. Accuracy Assessment and Model Comparison

To verify the effectiveness of the NDVI$_{\text{CV,LS}}$ model, we compared its performance with that of the following three evergreen forest extraction methods: 1) NDVI$_{\text{min}}$, evergreen forest is extracted using only the NDVI$_{\text{ann-min}}$ without considering the constraints from the CV, with harmonized Landsat-8 and Sentinel-2 images; 2) NDVI$_{\text{win}}$, evergreen forest is extracted from single-scene NDVI images in winter, using Landsat-8 images acquired on January 23, 2019; and 3) NDVI$_{\text{CV}}$-L, evergreen forest is extracted using only Landsat-8 time-series data. The preprocessing for harmonizing Landsat-8 and Sentinel-2 images differs from standard Landsat-8 and Sentinel-2 products; thus, standard Landsat-8, Sentinel-2, and the harmonized Landsat-8 and Sentinel-2 dataset are not directly comparable. From here on, we will use the terms “Landsat” and “Sentinel-2” to refer to their respective subsets of the harmonized Landsat-8 and Sentinel-2 dataset. The thresholds of the parameters used in the above three models were consistent with those used in the NDVI$_{\text{CV,LS}}$ model. To assess the accuracy of the experimental results, RMSE and MAE were used as verification standards [54] and calculated as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (X_{\text{obs},i} - X_{\text{mod},i})^2} \quad (9)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |X_{\text{obs},i} - X_{\text{mod},i}| \quad (10)$$

where $n$ is the number of samples, $X_{\text{mod},i}$ is the FVC$_{\text{ever}}$ estimated using the NDVI$_{\text{CV,LS}}$ model, and $X_{\text{obs},i}$ is the FVC$_{\text{ever}}$ estimated from visual interpretation of high-resolution images.

IV. RESULTS

A. Analysis Results of Model Parameters

The relationships between CV, NDVI$_{\text{ann-min}}$, and the proportion of evergreen forest in mixed pixels are shown in Fig. 7. From the results of simulation analysis, it can be checked that the CV decreases with increasing evergreen forest proportion, whereas the NDVI$_{\text{ann-min}}$ increases with increasing evergreen forest proportion. When the evergreen forest proportion was greater than 0, the CV was less than 0.22. When the underlying surface types in the mixed pixels differed, the NDVI$_{\text{ann-min}}$ values corresponding to the same evergreen forest proportion were different. With an evergreen forest proportion of 1, the NDVI$_{\text{ann-min}}$ associated with impervious areas was 0.51, whereas that associated with nonimpervious areas was 0.54. With an evergreen forest proportion of 0, the NDVI$_{\text{ann-min}}$ associated with impervious and nonimpervious areas was 0.2 and 0.34, respectively. After considering the different underlying surface types contained in the mixed pixel, the following final threshold values for the NDVI$_{\text{CV,LS}}$ model were determined: CV throughout the study area, 0.22; NDVI$_{\text{ann-min,back}}$ and NDVI$_{\text{ann-min,veg}}$ in areas with impervious surfaces, 0.20 and 0.51, respectively; NDVI$_{\text{ann-min,back}}$ and NDVI$_{\text{ann-min,veg}}$ in other areas, 0.34 and 0.54, respectively.

B. Estimation of FVC$_{\text{ever}}$

Fig. 8 shows the spatial distribution of evergreen forest in the study area in 2019, estimated based on CV and NDVI$_{\text{ann-min}}$ thresholds obtained through prior analyses. Hefei contains 430.86 km$^2$ of evergreen forest, which are mainly distributed near mountainous areas. Sporadic evergreen tree stands are also scattered throughout the urban area and were likely grown for the purposes of afforestation and aesthetics. In the flat main crop-growing areas, there are few evergreen trees. The estimated FVC$_{\text{ever}}$ results almost match the actual ground situation. In particular, our method has certain advantages in extracting evergreen forest with low coverage in urban built-up areas [Fig. 8(a)].

C. Model Validation and Comparison

The extraction accuracy of the model was verified using field samples, Google Earth and GF-1 high-resolution images. The overall classification accuracy of the NDVI$_{\text{CV,LS}}$ model was higher than that of other models (Table II). The user accuracy of our model was 91% and producer accuracy was 86% in terms of evergreen forest extraction. For the study area, harmonizing Landsat-8 and Sentinel-2 time-series images to extract evergreen forest improved the overall accuracy by 8% compared with using Landsat-8 images alone. Meanwhile, the RMSE (0.11) and MAE (0.09) of NDVI$_{\text{CV,LS}}$ model were both lower than that of NDVI$_{\text{CV,L}}$ model (0.24 and 0.17, respectively). The overall classification accuracy of the NDVI$_{\text{min}}$ model was approximately 15% lower than that of the NDVI$_{\text{CV,LS}}$ model. The overall accuracy, user accuracy, and producer accuracy of the single-phase NDVI$_{\text{win}}$ model were all low at approximately 60%.

Although the NDVI$_{\text{CV,L}}$ model was based on time-series images and took CV and NDVI$_{\text{ann-min}}$ constraints into account, the latter two parameters could not reflect real phenological characteristics of evergreen forest. Because some images in phenology periods may not be available due to the influence of clouds and rain, which led to a lower classification accuracy. The NDVI$_{\text{min}}$ model extracted evergreen forest based on the minimum NDVI value in the time series, which was derived from a single image. The NDVI$_{\text{win}}$ model was also based on single-phase images. Thus, it is expected that the classification accuracy of these two models will not be high. In conclusion, the NDVI$_{\text{CV,LS}}$ model is the most effective and stable model.

We also compared the results of the NDVI$_{\text{CV,LS}}$ model with that of supervised classification methods to validate our model. We chose three classifiers—maximum likelihood classification, minimum distance classification, and support vector machine
Fig. 8. Distribution of FVC_{\text{ever}} in the study area in 2019 based on the NDVI_{CV}-LS model. (a) FVC_{\text{ever}} in the urban area (red box in the main map) where land cover is fragmented and FVC_{\text{ever}} is low. (b) FVC_{\text{ever}} according to evergreen forest area.

TABLE II

| Model          | User Accuracy | Producer Accuracy | Overall Accuracy |
|----------------|---------------|-------------------|-----------------|
|                | Evergreen Forest | Others | Evergreen Forest | Others |              |
| NDVI_{CV}-LS   | 91.1%          | 94.6%  | 86.2%          | 95.8%  | 93.0%        |
| NDVI_{CV}-L    | 85.0%          | 87.4%  | 82.3%          | 86.6%  | 85.3%        |
| NDVI_{min}     | 78.0%          | 80.4%  | 77.3%          | 81.6%  | 78.4%        |
| NDVI_{max}     | 67.1%          | 56.3%  | 65.2%          | 59.5%  | 66.7%        |

Others, objects in the confusion matrix other than evergreen forest.

classification [55]–[57]. Ten classification operations were repeated for each classifier, and the training samples used for each classification routine were randomly generated. The same training and verification samples were used for all three classifiers to ensure that the outputs were comparable. The commission and omission errors of the different classifiers in distinguishing evergreen forest at the pixel scale were analyzed (Fig. 9). The commission and omission errors of the NDVI_{CV}-LS model were 0.09 and 0.14, significantly smaller than those of the three supervised classification methods, indicating that our model can effectively extract evergreen forest.

V. DISCUSSION

Mapping evergreen forest using time-series images faces some difficulties, for instance, incomplete time series and low
Fig. 10. Variation in the annual NDVI associated with evergreen trees, deciduous trees, farmland, and built-up areas. Values were obtained from sample points in satellite images. Some key phenology periods (i.e., March, June, and November) were shown in red boxes.

Fig. 11. Consistency in the identification of forest (left column), farmland (middle column), and urban area (right column) using Landsat-8 and Sentinel-2 images based on the red band (top row), near-infrared band (middle row), and NDVI (bottom row).

Spatial resolution. There could be a great improvement, as shown in our study, by means of the NDVI$_{CV}$-LS model. Based on harmonizing Landsat-8 and Sentinel-2 images, we improved the accuracy of FVC$_{ever}$ estimation in subtropical urban areas that frequently experience cloudy and rainy weather conditions.

A. Necessity and Feasibility of Data Harmonization

Satellites that capture medium-to-high spatial resolution images tend to have a long revisiting period, such as 16 days for Landsat. For example, only 23 Landsat images of Hefei were taken in 2019, of which only nine had little or no cloud cover.
Combined with the growth characteristics of evergreen forest, a continuous time series of images is needed as a data source for mapping. Given the weather conditions in Subtropics of the Northern Hemisphere, not many usable images can be captured during the plum rain season [52]. If the time series is not complete enough, only the NDVI when the background vegetation has the best growth is available, the $FVC_{\text{ever}}$ estimation results will be unsatisfactory. Fig. 10 shows the key phenology periods in which evergreen forest is characterized as distinct from other vegetation. If only Landsat-8 is used as the data source, some information about the key phenology period (such as June) will be lost, leading to inaccurate $\text{NDVI}_{\text{ann-min}}$. Introducing multisource data can increase the possibility of obtaining the optimal $\text{NDVI}_{\text{ann-min}}$, which directly impacts the accuracy of the $FVC_{\text{ever}}$ estimation in pixels. The comparison of the $FVC_{\text{ever}}$ estimation results between $\text{NDVI}_{\text{CV}-\text{LS}}$ and $\text{NDVI}_{\text{CV}-\text{L}}$ model also proves this point (Section C, Chapter IV). To ensure the time series is complete enough, the dataset must be supplemented with images from multiple sources.

Images acquired from different satellites vary in spatial resolution and spectral band range due to differences in solar altitude, atmospheric conditions, and sensors. In the process of harmonizing multisource images, in addition to correction through preprocessing, selecting similar data is also critical to minimize the biases. Therefore, Landsat-8 and Sentinel-2, whose spatial resolution and spectral information were similar, were selected as the data sources of harmonizing in the study. To assess the feasibility of harmonizing multisource data undertaken here, we selected three types of ground objects—forest, farmland, and urban area—to verify if the Landsat-8 and Sentinel-2 data were consistent. Sample points representing each type of ground object were selected from Landsat-8 images acquired on January 23, 2019 and Sentinel-2 images acquired on January 22, 2019. Then, regression analysis was performed using the reflectance of the red and NIR bands and the NDVI values of corresponding pixels. The results showed that Landsat-8 and Sentinel-2 images exhibited high consistency and can be used to construct a time-series dataset (Fig. 11). Nevertheless, some inconsistencies were observed for ground objects depending on

the spectral band. Within the Red band, forest was the most consistent between the two sensors, followed by farmland and urban area. Within the NIR band, forest and farmland were both highly consistent between the two sensors, whereas urban area is somewhat inconsistent. Because compared with forest and farmland, urban area has lower near-infrared reflectance values and is more susceptible to changes in solar altitude. The trend in consistency in the NDVI regarding the three types of ground objects was similar to that observed in the NIR band. There are many types of impervious surfaces with different characteristics, including roads and buildings; thus, this gives rise to a range of NDVI values. Additionally, NDVI values are calculated using NIR and Red band reflectance values. The spectral values associated with forest and farmland increase significantly going from Red to NIR wavelengths, whereas the spectral value associated with impervious surfaces do not differ much between the two bands. As a result, spectral differences between the two sensors were amplified by the normalization process.

B. Applicability of the $\text{NDVI}_{\text{CV}-\text{LS}}$ Model

The accuracy of this model is greatly affected by spatial heterogeneity, especially for areas with complex and fragmented underlying surfaces. The $\text{NDVI}_{\text{ann-min}}$ values associated with pure background pixels containing different underlying surfaces varied; e.g., the $\text{NDVI}_{\text{ann-min}}$ was 0.2 for pixels containing buildings and 0.3 for pixels containing farmland and deciduous vegetation. In urban areas, building shadows can lead to low $\text{NDVI}_{\text{ann-min}}$ values, whereas in nonurban areas, the mixing of withered crops, grasses, and leaves into the soil can lead to higher $\text{NDVI}_{\text{ann-min}}$ values. If the same threshold value is used for extracting evergreen forest in mixed pixels with different underlying surfaces, the results will not be satisfactory. In the GF-1 images [Fig. 12(a)], residential areas alternate with evergreen forest, resulting in a large number of mixed pixels. However, if evergreen forest is extracted without considering thresholds for different land-cover types, only pixels in which evergreen forest is mixed with other types of vegetation (deciduous vegetation,
Farmland, grassland, etc.) would be extracted. In this case, pixels in which evergreen forest co-occurs with impervious surfaces (residential land, roads, etc.) would not be considered, and any urban areas with a low coverage of evergreen forest would be ignored [Fig. 12(b)]. By using NDBI to divide the study area and setting parameters by area, good extraction results were obtained [Fig. 12(c)]. With this modification, the NDVI-LS model can be used for the fine extraction of evergreen forest in urban areas.

C. Uncertainties Associated With The Model

Because images were obtained from two different remote sensors, cumulative errors due to differences in spatial resolution and spectral band range will increase the uncertainties associated with model outcomes. Although harmonization processing was performed, some errors will still inevitably affect the classification results [45]. In addition, the shorter the time-series interval, the better the CV of NDVI reflects the actual situation [58]. Therefore, it suggests that a greater frequency of observations, such as from harmonizing data across all comparable sensors, is still needed.

VI. Conclusion

In this article, we developed a new approach, NDVI-LS, for extracting evergreen forest from harmonized multisource remote-sensing images acquired in cloudy and rainy areas. We used Landsat-8 and Sentinel-2 images with cloud cover less than 10% acquired within a year to reconstruct a continuous time-series dataset. Then, the CV of time-series NDVI was used to separate evergreen forest from other vegetation, and the NDVI\_\text{ann-min} was used as a parameter in the dimidiate pixel model to estimate FVC\_\text{ever}. Given the spatial heterogeneity and severe mixed-pixel problem, the study area was divided into areas with impervious and nonimpervious surfaces. The dimidiate pixel model parameters were simulated and analyzed for each region. The overall classification accuracy of the NDVI-LS model was 93% with the RMSE and MAE of 0.11 and 0.09, respectively. The results showed that the addition of Sentinel-2 images provided the potential to build full-time-series models. In particular, the harmonized Landsat-8 and Sentinel-2 data improved the accuracy of FVC\_\text{ever} estimation in areas with complex underlying surfaces under cloudy and rainy conditions. Compared with other evergreen forest extraction models, the NDVI\_LS model is more stable and accurate.

REFERENCES

[1] R. Ni, Y. Baiketuerhan, C. Zhang, X. Zhao, and K. Y. Gadow, “Analysing structural diversity in two temperate forests in northeastern china,” Forest Ecol. Manage., vol. 316, pp. 139–147, 2014.

[2] M. Dalponte, L. Bruzzone, and D. Gianelle, “Tree species classification in the Southern Alps based on the fusion of very high geometrical resolution multispectral/hyperspectral images and LiDAR data,” Remote Sens. Environ., vol. 123, pp. 258–270, Aug. 2012.

[3] X. Shang and P. Chaizee, “Interest of a full-waveform flown UV lidar to derive forest vertical structures and aboveground carbon,” Forests, vol. 5, pp. 1454–1480, 2014.

[4] A. K. Skidmore, N. Pettorelli, N. C. Coops, G. N. Geller, and M. Wegmann, “Agree on biodiversity metrics to track from space,” Nature, vol. 523, pp. 403–405, 2015.

[5] D. I. Forrester, “The spatial and temporal dynamics of species interactions in mixed-species forests: From pattern to process,” Forest Ecol. Manage., vol. 312, pp. 282–292, 2014.

[6] J. L. Hansen, C. G. Lorimer, C. R. Halpin, and B. J. Palik, “Ecological forestry in an uneven-aged, late-successional forest: Simulated effects of contrasting treatments on structure and yield,” Forest Ecol. Manage., vol. 270, pp. 94–107, 2012.

[7] S. Jose, R. Williams, and D. Zamora, “Belowground ecological interactions in mixed-species forest plantations,” Forest Ecol. Manage., vol. 233, pp. 231–239, 2006.

[8] A. Redondo-Brenes and F. Montagnini, “Growth, productivity, aboveground biomass, and carbon sequestration of pure and mixed native tree plantations in the Caribbean lowlands of Costa Rica,” Forest Ecol. Manage., vol. 232, pp. 168–178, 2006.

[9] J. Dong, X. Xiao, S. Sheldon, C. Biradar, and G. Xie, “Mapping tropical forests and rubber plantations in complex landscapes by integrating PAL-SAR and MODIS imagery,” ISPRS J. Photogramm. Remote Sens., vol. 74, pp. 20–33, 2012.

[10] L. D. Prior, D. Earnus, and D. M. J. S. Bowman, “Leaf attributes in the seasonally dry tropics: A comparison of four habitats in northern Australia,” Functional Ecol., vol. 17, pp. 504–515, Aug. 2003.

[11] H. Wang, H. Shi, and Y. Li, “Leaf dust capturing capacity of urban greening plant species in relation to leaf micromorphology,” in Proc. Int. Symp. Urb. Forestr. Urban Greening Protect. Xi’an, China, 2011, pp. 2198–2201.

[12] Z. Zhang, Y. Lv, and H. Pan, “Cooling and humidifying effect of plant communities in subtropical urban parks,” Urban Forestry Urban Greening, vol. 12, pp. 323–329, 2013.

[13] B. Chan and J. Guldum, “Impact of greening on the urban heat island: Seasonal variations and mitigation strategies,” Comput. Environ. Urban Syst., vol. 71, pp. 165–176, 2018.

[14] Y. Xie, Z. Sha, and M. Yu, “Remote sensing imagery in vegetation mapping: A review,” Narnaia, vol. 1, pp. 9–23, 2008.

[15] J. Graesser and N. Ramankutty, “Detection of cropland field parcels from Landsat imagery,” Remote Sens. Environ., vol. 201, pp. 165–180, 2017.

[16] C. Senf, D. Pfugmacher, S. van der Linden, and P. Hostert, “Mapping rubber plantations and natural forests in Xishuangbanna (Southwest China) using multi-spectral phenological metrics from MODIS time series,” Remote Sens., vol. 5, pp. 2795–2812, 2013.

[17] L. Duarte, A. C. Teodoro, A. T. Monteiro, M. Cunha, and H. Goncalves, “\text{QPhenoMetrics}: An open source software application to assess vegetation phenology metrics,” Comput. Electron. Agriculture, vol. 148, pp. 82–94, 2018.

[18] I. Moreau and P. Defourny, “The vegetation phenology detection in Amazon tropical evergreen forests using SPOT-VEGETATION 11-y time series,” in Proc. IEEE Int. Geosci. Remote Sens. Symp., 2012, pp. 40–43.

[19] W. Koo, C. Liang, L. Wei, A. Hernandez, and X. Yang, “Phenology-based method for mapping tropical evergreen forests by integrating of MODIS and Landsat imagery,” Forests, vol. 8, 2017, Art. no. 34.

[20] Y. Qin et al., “Mapping forests in monsoon Asia with ALOS PALSAR 50-m mosaic images and MODIS imagery in 2010,” Sci. Rep., vol. 6, 2016, Art. no. 20880.

[21] X. Xiao, C. Biradar, C. Czarnecki, T. Alabi, and M. Keller, “A simple algorithm for large-scale mapping of evergreen forests in tropical America, Africa and Asia,” Remote Sens., vol. 1, pp. 357–374, 2009.

[22] X. Xiao, S. Hagen, Q. Zhang, M. Keller, and B. Moore, “Detecting leaf phenology of seasonally moist tropical forests in South America with multi-temporal MODIS images,” Remote Sens. Environ., vol. 103, pp. 465–473, 2006.

[23] X. Xiao et al., “Mapping paddy rice agriculture in Southern China using multi-temporal MODIS images,” Remote Sens. Environ., vol. 95, pp. 480–492, 2005.

[24] Z. Zhang, H. Su, and Liu, “Harmonizing multi-source remote sensing images for summer corn growth monitoring,” Remote Sens., vol. 11, 2019, Art. no. 1266.

[25] B. D. Wardlow, S. L. Egbert, and J. H. Kastens, “Analysis of time-series MODIS 250 m vegetation index data for crop classification in the U.S. Central Great Plains,” Remote Sens. Environ., vol. 108, pp. 290–310, 2007.

[26] L. Zeng et al., “A hybrid approach for detecting corn and soybean phenology with time-series MODIS data,” Remote Sens. Environ., vol. 181, pp. 237–250, 2016.

[27] D. P. Roy et al., “Landsat-8: Science and product vision for terrestrial global change research,” Remote Sens. Environ., vol. 145, pp. 154–172, 2014.
L. Yan and D. P. Roy, “Large-area gap filling of landsat reflectance time series by spectral-angle-mapper based spatio-temporal similarity (SAM-STS),” Remote Sens., vol. 10, 2018, Art. no. 609.

H. Fang, B. Wu, H. Liu, and X. Huang, “Using NOAA AVHRR and Landsat TM to estimate rice area year-by-year,” Int. J. Remote Sens., vol. 19, pp. 521–525, 1998.

Q. Huang et al., “Discharge estimation in high-mountain regions with improved methods using multisource remote sensing: A case study of the Upper Brahmaputra River,” Remote Sens. Environ., vol. 219, pp. 115–134, 2018.

Li, D. Long, Q. Huang, P. Han, F. Zhao, and Y. Wada, “High-temporal-resolution water level and storage change data sets for lakes on the Tibetan Plateau during 2000–2017 using multiple altimetric missions and Landsat-derived lake shoreline positions,” Earth Syst. Sci. Data, vol. 11, pp. 1603–1627, 2019.

F. Gao et al., “Toward mapping crop progress at field scales through fusion of Landsat and MODIS imagery,” Remote Sens. Environ., vol. 188, pp. 9–25, 2017.

B. Huang and H. Song, “Spatiotemporal reflectance fusion via sparse representation (Article),” IEEE Trans. Geosci. Remote Sens., vol. 50, no. 10, pp. 3707–3716, Oct. 2012.

S. Skakun, B. Franch, E. Vermote, J. C. Roger, and E. Murphy, “Winter wheat yield assessment using Landsat 8 and sentinel-2 data,” in Proc. IEEE Int. Geosci. Remote Sens. Symp., 2018, pp. 5964–5967.

P. A. P. G. Griffiths, C. A. Nendel, J. A. Pickert, and P. A. Hostert, “Towards national-scale characterization of grassland use intensity from integrated Sentinel-2 and Landsat time series,” Remote Sens. Environ., vol. 238, 2020, Art. no. 111124.

S. Falanga Bolognesi, E. Pasolli, O. R. Belfiore, C. De Michele, and G. D’Urso, “Harmonized landsat 8 and sentinel-2 time series data to detect irrigated areas: An application in Southern Italy,” Remote Sens., vol. 12, 2020, Art. no. 1275.

C. Ichoku and A. Karnieli, “A review of mixture modeling techniques for sub-pixel land cover estimation,” Remote Sens. Rev., vol. 13, pp. 161–186, 1996.

D. A. Roberts, M. Gardner, R. Church, S. Ustin, G. Scheer, and R. O. Green, “Mapping chaparral in the Santa Monica mountains using multiple endmember spectral mixture models,” Remote Sens. Environ., vol. 65, pp. 267–279, 1998.

J. J. Settle and N. A. Drake, “Linear mixing and the estimation of ground cover proportions,” Int. J. Remote Sens., vol. 14, pp. 1159–1177, 1993.

R. Tateishi, Y. Shimazaki, and P. D. Gunin, “Spectral and temporal linear mixing model for vegetation classification,” Int. J. Remote Sens., vol. 25, pp. 4203–4218, 2004.

M. Claverie et al., “The harmonized landsat and sentinel-2 surface reflectance data set,” Remote Sens. Environ., vol. 219, pp. 145–161, 2018.

V. Eric, J. Chris, C. Martin, and F. Belen, “Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product,” Remote Sens. Environ., vol. 185, pp. 46–56, 2016.

Z. Zhu, S. Wang, and C. E. Woodcock, “Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4-7, 8, and Sentinel 2 images,” Remote Sens. Environ., vol. 159, pp. 269–277, 2015.

F. Gao, J. G. Masek, and R. E. Wolfe, “Automated registration and orthorectification package for Landsat and Landsat-like data processing,” J. Appl. Remote Sens., vol. 3, 2009, Art. no. 033515.

D. P. Roy et al., “A general method to normalize Landsat reflectance data to nadir BRDF adjusted reflectance,” Remote Sens. Environ., vol. 176, pp. 255–271, 2016.

Q. Zhou et al., “Monitoring landscape dynamics in Central U.S. grasslands with harmonized landsat-3 and sentinel-2 time series data,” Remote Sens., vol. 11, 2019, Art. no. 328.

D. Long et al., “Generation of spatially complete and daily continuous surface soil moisture of high spatial resolution,” Remote Sens. Environ., vol. 233, 2019, Art. no. 111364.

A. R. S. Marcal and G. G. Wright, “The use of ‘overlapping’ NOAA-AVHRR NDVI maximum value composites for Scotland and initial comparisons with the land cover census on a Scottish Regional and District basis,” Int. J. Remote Sens., vol. 18, pp. 491–503, 2010.

C. J. Tucker, “Red and photographic infrared linear combinations for monitoring vegetation,” Remote Sens. Environ., vol. 8, pp. 127–150, 1979.

E. Weiss, S. E. Marsh, and E. S. Pitman, “Application of NOAA-AVHRR NDVI time-series data to assess changes in Saudi Arabia’s rangelands,” Int. J. Remote Sens., vol. 22, pp. 1005–1027, 2001.

J. Qi et al., “Spatial and temporal dynamics of vegetation in the San Pedro River basin area,” Agricultural Forest Meteorol., vol. 105, pp. 55–68, 2000.

A. K. Ambika, B. Wardlow, and V. Mishra, “Remotely sensed high resolution, irrigated area mapping in India for 2000 to 2015,” Sci. Data, vol. 3, 2016, Art. no. 160118.

S. K. Maxwell and K. M. Sylvestre, “Identification of ‘ever-cropped’ land (1984–2010) using Landsat annual maximum NDVI image composites: Southwestern Kansas case study,” Remote Sens. Environ., vol. 121, pp. 186–195, 2012.

Y. Zha, J. Gao, and S. Ni, “Use of normalized difference built-up index in automatically mapping urban areas from TM imagery,” Int. J. Remote Sens., vol. 24, pp. 583–594, 2010.

L. Mentascchi, G. Besio, F. Cassola, and A. Mazzino, “Problems in RMSE-based wave model validations,” Ocean Model., vol. 72, pp. 53–58, 2013.

C. Cortes and V. Vapnik, “Support-vector networks,” Mach. Learn., vol. 20, pp. 273–297, 1995.

A. R. John, Remote Sensing Digital Image Analysis. Berlin, Germany: Springer, 1994.

J. D. Paola and R. A. Schowengerdt, “A detailed comparison of back-propagation neural network and maximum-likelihood classifiers for urban land use classification,” IEEE Trans. Geosci. Remote Sens., vol. 33, no. 33, pp. 981–996, Jul. 1995.

W. Liu, J. Dong, K. Xiang, S. Wang, W. Han, and W. Yuan, “A sub-pixel method for estimating planting fraction of paddy rice in Northeast China,” Remote Sens. Environ., vol. 205, pp. 305–314, 2018.

Taixia Wu received the M.S. degree in cartography and GIS from the College of Urban and Environmental Sciences, Northeast Normal University, Changchun, China, in 2006, and the Ph.D. degree in remote sensing from Peking University, Beijing, China, in 2010.

He is currently a Professor with the School of Earth Sciences and Engineering, Hohai University, Nanjing, China. His research interests include hyperspectral remote sensing and water remote sensing applications.

Yuting Zhao is currently working toward the M.S. degree in photogrammetry and remote sensing at the School of Earth Sciences and Engineering, Hohai University, Nanjing, China.

Her research interests include forest resources monitoring and extracting using multiple remote sensing datasets.
Hongjun Su (Senior Member, IEEE) received the B.S. degree in geography information system from the China University of Mining and Technology, Xuzhou, China, in 2006, and the Ph.D. degree in cartography and geography information system from the Key Laboratory of Virtual Geographic Environment (Ministry of Education), Nanjing Normal University, Nanjing, China, in 2011.

He was a Visiting Student with the Department of Electrical and Computer Engineering, Mississippi State University, Starkville, MS, USA, from 2009 to 2010. He was also a Postdoctoral Fellow with the Department of Geographical Information Science, Nanjing University, Nanjing, from 2012 to 2014. He was also a Research Scholar with the Department of Geography, University of Wisconsin–Madison, Madison, WI, USA, from 2019 to 2020. He is currently a Full Professor with the School of Earth Sciences and Engineering, Hohai University, Nanjing. He has authored more than 50 journal papers. His research interests include hyperspectral remote sensing dimensionality reduction, classification, and spectral unmixing.

Prof. Su is an Associate Editor of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, a Guest Editor of the Remote Sensing. He was the recipient of the 2016 Best Reviewer Award from the IEEE Geoscience and Remote Sensing Society. He is an Active Reviewer for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, and the International Journal of Remote Sensing.

Yingying Yang is currently working toward the Ph.D. degree in ecological remote sensing at the School of Earth Sciences and Engineering, Hohai University, Nanjing, China.

Her research interests include remote sensing of environment and forest monitoring.

Dongzhen Jia received the B.Eng. degree in surveying and mapping engineering, the M.S. and Ph.D. degrees in geodesy from Hohai University, Nanjing, China, in 2003, 2006, and 2018, respectively. From December 2007 to December 2008, he was a Visiting Researcher with Nanyang Technological University, Singapore. He is currently a Lecturer with the School of Earth Science and Engineering, Hohai University. His research interests include deformation monitoring and 3-D laser scanning applications.