Improved Mapping Results of 10 m Resolution Land Cover Classification in Guangdong, China Using Multisource Remote Sensing Data With Google Earth Engine

Ying Tu, Wei Lang, Le Yu, Ying Li, Junhao Jiang, Yawen Qin, Jiemin Wu, Tingting Chen, and Bing Xu

***Abstract*—Land cover information depicting the complex interactions between human activities and surface change is critically essential for nature conservation, social management, and sustainable development. Recent advances have shown great potentials of remote sensing data in generating high-resolution land cover maps, but it remains unclear how different models, data sources, and inclusive features affect the classification results, which hinders its applications in regional studies requiring more accurate land cover data. Informing these issues, here we developed a robust framework to improve the mapping results of 10 m resolution land cover classification in Guangdong Province, China using thousands of manually collected samples, multisource remote sensing data (Sentinel-1, Sentinel-2, and Luojia-1), machine learning algorithms, and a free cloud-based platform of Google Earth Engine. Results showed that an overall accuracy of 86.12% and a Kappa coefficient of 0.84 could be achieved for land cover classification in Guangdong for 2019. We found that random forest models achieved better performance than classification and regression trees, minimum distance, and support vector machine models. We also found that features derived from Sentinel-1 data and Sentinel-2 spectral indices greatly contributed to the classification process, while the feature of Luojia-1 data was not as much important as other configurations. A comparison between our results and several existing land cover products in terms of classification accuracy and visual interpretation further validated the effectiveness and robustness of the proposed framework. Our experiments and findings not only systematically elucidate the role of classification methods and data sources in deriving more accurate and reliable land cover maps but also provide certain guidelines for future land cover mapping from regional to global scales.**

***Index Terms*—Google Earth Engine (GEE), high-resolution land cover mapping, Luojia-1, machine learning, Sentinel-2, Sentinel-1.**

Manuscript received July 3, 2020; revised July 26, 2020 and August 22, 2020; accepted August 28, 2020. Date of publication September 7, 2020; date of current version September 21, 2020. This work was supported in part by the National Key Research and Development Program of China under Grant 2016YFA0600104, in part by the National Natural Science Foundation of China under Grant 41801163, in part by the Natural Science Foundation of Guangdong Province under Grant 2018A030313221, in part by the National Key Scientific and Technological Infrastructure project "Earth System Science Numerical Simulator Facility" (EarthLab), and in part by donations from Delos Living LLC and the Cyrus Tang Foundation to Tsinghua University. *(Corresponding author: Bing Xu.)*

Ying Tu and Le Yu are with the Ministry of Education Key Laboratory for Earth System Modeling, Department of Earth System Science,Tsinghua University, Beijing 100084, China (e-mail: tu-y19@mails.tsinghua.edu.cn; leyu@tsinghua.edu.cn).

Wei Lang and Tingting Chen are with the Department of Urban and Regional Planning, School of Geography and Planning, Sun Yat-sen University, Guangzhou 510275, China, and also with the China Regional Coordinated Development and Rural Construction Institute,Sun Yat-sen University, Guangzhou 510275, China (e-mail: langw3@mail.sysu.edu.cn; chentt53@mail.sysu.edu.cn).

Ying Li, Junhao Jiang, and Yawen Qin are with the Department of Urban and Regional Planning, School of Geography and Planning, Sun Yat-sen University, Guangzhou 510275, China (e-mail: liying268@mail2.sysu.edu.cn; jiangjh26@mail2.sysu.edu.cn; qinyw5@mail2.sysu.edu.cn).

Jiemin Wu is with the Department of Remote Sensing and Geographic Information Engineering, School of Geography and Planning, Sun Yat-sen University, Guangzhou 510275, China (e-mail: wujm23@mail2.sysu.edu.cn).

Bing Xu is with the Ministry of Education Key Laboratory for Earth System Modeling, Department of Earth System Science,Tsinghua University, Beijing 100084, China, and also with the Center for Healthy Cities, Institute for China Sustainable Urbanization,Tsinghua University, Beijing 100084, China (e-mail: bingxu@tsinghua.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2020.3022210

I. Introduction

LAND cover, as a key element for earth system science, provides fundamental information for understanding the complex interactions between human activities and surface change. Land cover maps play an important role in natural resources management, including biodiversity conservation, carbon cycling, climate change, ecosystem protection, and hydrological process [1]–[6]. They are also essential to studies of public health, sustainable development, and urban planning [7]–[10]. Under this context, there has raised a growing demand for broad-scale and high-precision land cover products.

The advent and development of remote sensing technology have greatly facilitated the application of land cover mapping. During the past few decades, numerous global land cover (GLC) datasets have been developed and applied with resolution ranging from 300 m to 1 km, using coarse resolution satellite imagery such as AVHRR, MODIS, and SPOT [11]–[15]. For regional-scale studies, popular national products such as the 1 km LCC85-05 for Canada [16], the 250 m DLCDv1 for Australia [17], and the 1:50 000 scale national LULC database for India [18] have caught much spotlight by the remote sensing community. However, the relatively coarse resolution and considerably low accuracies of these land cover datasets make it difficult to provide sufficient details of the Earth's surface [19]–[21]. Consequently, they are far from satisfactory for many sophisticated applications such as the identification of cropping types, evaluation of disasters, and management of urban transportation.

It was not until the 2010s when Gong *et al*. [22] reported the first 30 m resolution GLC product of Finer Resolution Observation and Monitoring of GLC (FROM-GLC30) maps using more than 8900 scenes Landsat TM/ETM+ images. Since then, several 30 m resolution GLC products, including FROM-GLC-seg [23], FROM-GLC-agg [24], and GlobeLand30 [25], have been produced and released. Recently, the first 10 m resolution GLC map for 2017, FROM-GLC10, was developed by Gong *et al*. [26] with an overall accuracy of 72.76%. Based on the theory of "stable classification with limited sample," a seasonal sample set collected from the 30 m resolution Landsat 8 images in 2015 [27] was successfully transferred to classify the 10 m resolution Sentinel-2 images in 2017 [26]. However, this product has some potential shortcomings, especially when it comes to regional-scale research: 1) Notable misclassifications between cropland, grassland, and forest. Due to the spatially mixed vegetation structures, seasonal variations of vegetation, and the temporal inconsistency between Landsat and Sentinel data, the classification accuracy for these vegetation types is relatively low [28]; 2) Usage of limited features. The predictors are mainly depending on Sentinel-2-based features such as spectral and derived remote sensing indices, without considering other information that is highly correlated with land covers such as nighttime light (NTL) and incident microwave radiation [29], [30]; and 3) Uncertainty of models. A unified model is used to classify GLC, which may lead to biased performance for localized experiments. To sum up, for regional-scale studies that require higher data quality, it is of great importance to discuss the impact of models, data sources, and feature selections on the classification results, thus gaining experiences to produce more accurate land cover maps.

In recent years, a free cloud-based platform, Google Earth Engine (GEE), has caught much spotlight by the remote sensing community. GEE stores petabyte scales of over 40 years of remotely sensed, climate-weather, geophysical datasets, and additional ready-to-use products [31]. It also enables users to discover, analyze, and visualize geospatial big data in powerful ways without needing access to supercomputers or specialized coding expertise [32]. A series of survey studies, ranging from regional to global scales, have been carried out based on GEE, including land cover and land use classification [26], [33]–[36], crop mapping and yield estimation [37]–[41], forest mapping [42], [43], surface water detection [44], [45], etc.

Leveraging the research advances, this study will explore land cover classification in Guangdong Province, China, based on a collection of thousands of manually collected samples, using multisource remote sensing data (Sentinel-2, Sentinel-1, and Luojia-1) with the GEE platform. Our ultimate goal is to develop a robust and cost-effective framework for provincial land cover mapping, thus present an update and improvement results of 10 m resolution land cover classification in Guangdong for 2019. Specifically, we seek to answer the following scientific questions.

1) Which model achieves relatively robust and accurate performance for land cover classification?
2) What is the relative importance of inclusive features and how do different data sources contribute to the classification process?
3) How is the performance of our results compared with existing land cover products (such as the FROM-GLC10 product), and what are the similarities and differences between them?

An improved understanding of these issues is needed to guide and move forward the campaign of land cover classification from regional to global scales.

II. Study Area and Data

A. Study Area

We choose Guangdong Province (20°13′N–25°31′N, 109°39′N E–117°19′E) in China as the study area (see Fig. 1). Located on the north shore of the South China Sea, Guangdong possesses a total area of about 178405.85 km^2 with a population of 115.21 million in 2019. The landscape of Guangdong slopes from north to south: the northern part is mostly mountainous, while the south is mainly covered by plains and hills. As shown in Fig. 1, there are 21 prefecture-level cities in Guangdong. Among them, Guangzhou and Shenzhen are among the most populous and important cities in China and have now become two of the world's most populous megacities.

B. Sentinel-2 Optical Data

Sentinel-2 is a wide-swath, high-resolution, multispectral imaging mission with a global 5-day revisit frequency. As shown in Table VII, the Sentinel-2 data includes 12 spectral bands: four visible and NIR bands at 10 m, six red edge and SWIR bands at 20 m, and two atmospheric bands at 60 m spatial resolution. In this study, the Level-2A product acquired in 2019, which provided surface reflectance values, was used for further analysis [46].

C. Sentinel-1 SAR Data

The Sentinel-1 mission provides ground range detected (GRD) data from a dual-polarization C-band Synthetic Aperture Radar (SAR) instrument. SAR instruments are capable of acquiring meaningful data in all weather conditions (even clouds) during daytime and nighttime. The signal recorded in GRD data is the backscatter coefficient that measures the incident microwave radiation scattered by the radiated terrain. The scattering behavior depends on the geometry of terrain elements and their electromagnetic characteristics. The data has a spatial resolution of 10 m. We acquired the full coverage of Sentinel-1 covering the study area in 2019.

D. Luojia-1 NTL Data

Developed by Wuhan University in China, the new generation of Luojia 1-01 remote sensing satellite was successfully launched on 2 June 2018 [47]. Compared with previous NTL data such as the Defense Meteorological Satellite Program's Operational Linescan System, Luojia-1 imagery has a finer spatial resolution (about 130 m) and a higher radiometric quantization (14 bits), and it does not suffer the problems of saturation and blooming [48]. The advantages of this new data can significantly enhance the detection capacity of artificial lightings, thus

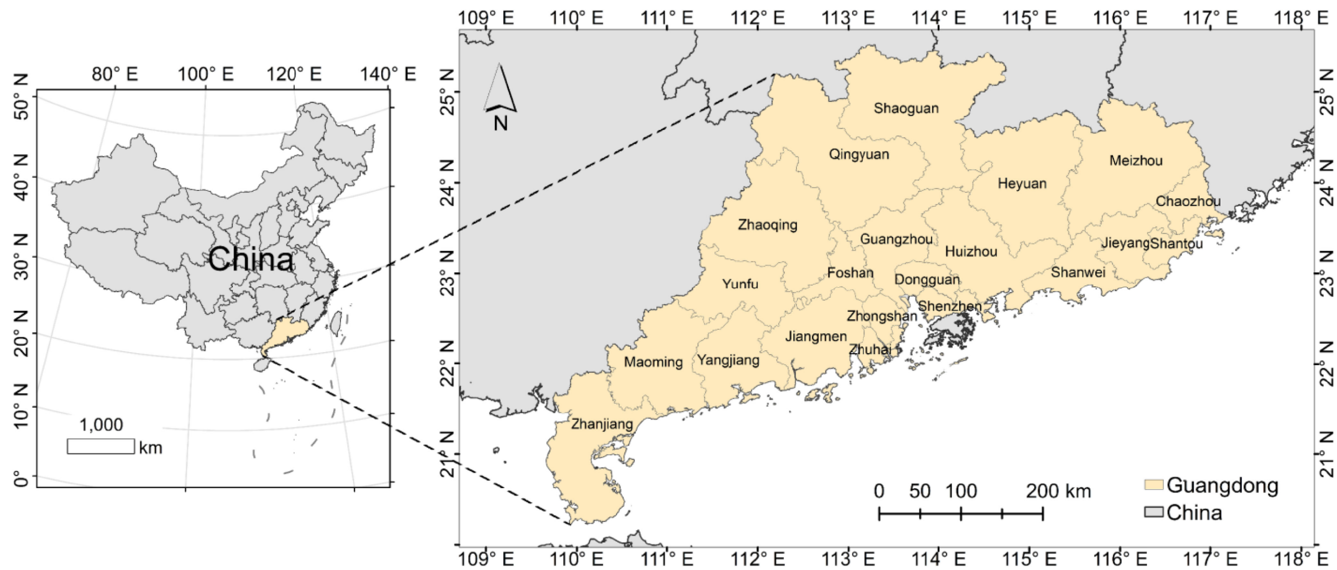

Fig. 1. Study area of Guangdong Province in China.

bringing new insights and possibilities to the research works on urban and environment [49]. We acquired the national Luojia-1 NTL imagery for 2018 from the Hubei Data and Application Center.[1] This data was produced using 275 scenes of Luojia-1 NTL images acquired between July and October that covered the land region of China [50]. Before image matching, NTL images were processed through cloud-contaminated data exclusion and stray lights removal [50]. The root mean square error of the tie points was 0.983 pixels and 195.491 m for independent checkpoints after the planar block adjustment [50]. More information about this data can be found in [50]. We clipped the national Luojia-1 NTL imagery based on the administrative boundary of Guangdong Province and later uploaded it to the GEE platform.

III. Methods

Fig. 2 presents the flowchart of the proposed framework, which consists of three main procedures: sampling, data preprocessing, and classification. All these steps can be undertaken internally and seamlessly on the GEE platform.

A. Classification System Adjustment

We adopted the classification system of FROM-GLC10, which divided land cover into ten types of cropland, forest, grassland, shrubland, tundra, wetland, water, impervious surface, bareland, and snow/ice. According to FROM-GLC10, we calculated the proportion of each land cover type and discovered that there was no tundra or snow/ice covering in Guangdong, a subtropical zone. As a result, we later excluded these two types in the adjusted classification system.

B. Sampling

We initialed a visual interpretation based on high-resolution satellite images from the Google Earth software. In total, 5000 sample points were collected for training and 1455 for the validation process (see Table I). Their spatial distributions, as shown in Fig. 9, covered most of the study area. Both training and validation samples were randomly generated by the computer at first and then manually interpreted.

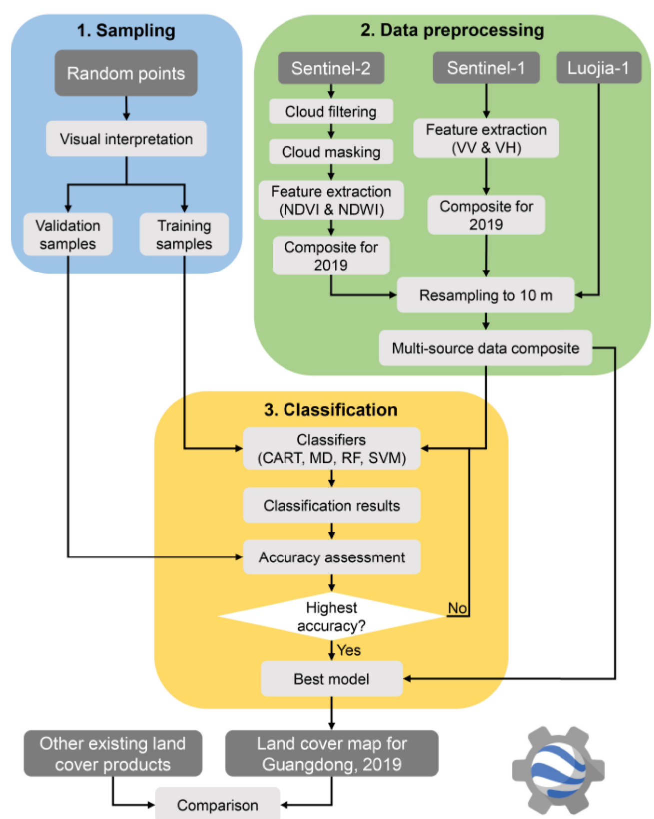

Fig. 2. Flowchart of the proposed framework.

[1][Online]. Available: http://www.hbeos.org.cn/, Accessed: Jul. 2020

TABLE I
SUMMARY OF TRAINING AND VALIDATION SAMPLES

Type	Number of training samples	Number of validation samples
Cropland	801	211
Forest	1427	238
Grassland	376	195
Shrubland	378	100
Wetland	226	100
Water	492	201
Impervious	950	220
Bareland	350	190
Total	5000	1455

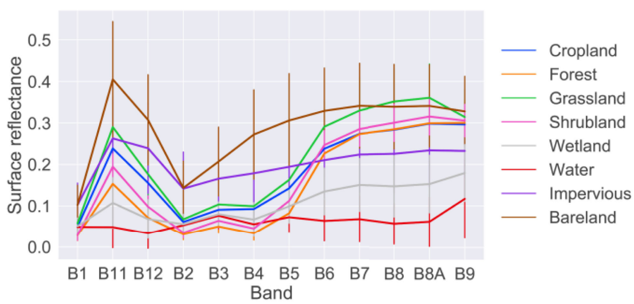

Fig. 3. Sentinel-2 spectral curves for training samples of each type. The curve of each type was obtained through calculating the "average" and "standard deviation" of the surface reflectance of all training samples (belonging to this type) for each given band in the processed cloud-free Sentinel-2 mean composite for 2019.

Fig. 3 displays the spectral curves for the training samples of each land cover type. For each type, we calculated the "average" and "standard deviation" of the surface reflectance of all training samples (belonging to this type) for each given band in the processed cloud-free Sentinel-2 mean composite for 2019 (see Section III-C) to obtain the curve. We found that the spectral characteristics had a high similarity among cropland, forest, grassland, and shrubland, whereas water and bareland were relatively distinguishable compared with other land cover types.

C. Data Preprocessing

For the Sentinel-2 optical imagery, to mitigate the limitation that arises due to cloud cover, we first filtered the whole-year archive with the percentage of cloudy pixels less than 3% using the "CLOUDY_PIXEL_PERCENTAGE" band information. Second, we did a pixel-based quality check to screen and filter out the poor-quality surface reflectance values using cloud mask and quality assessment information in the Sentinel-2 metadata. These two processes not only ensured that most of the study area had at least 5 scenes of images' coverage (see Fig. 4) but also eliminated the observations contaminated by clouds and shadows from the Sentinel-2 archive. We then calculated the Normalized difference vegetation index (NDVI = (NIR-Red)/(NIR+Red)) and Normalized difference water index (NDWI = (Green-NIR)/(Green+NIR)) values from the retained reflectance in the Green, Red, and NIR bands for each pixel. These two spectral indices were later combined with the original spectral bands of Sentinel-2 data (see Table VII) as input features into the machine learning algorithms. Finally, we calculated the average pixel values in the image collections to merge the whole-year Sentinel-2 archive and derived the cloud-free Sentinel-2 mean composite for 2019.

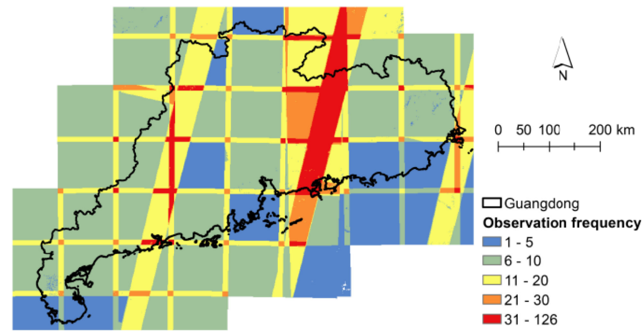

Fig. 4. Numbers of available observations in total for the Sentinel-2 optical data in Guangdong, 2019 after cloud filtering and removal. Noted that only images with cloudy pixel percentage less than 3% were used.

Given that there were different combinations of instrument mode and polarization in the Sentinel-1 data, we chose a homogeneous GRD subset by selecting GRD scenes with a dual-polarization (i.e., VV and VH) from the instrument mode of the interferometric wide swath. We acquired the Sentinel-1 composite for 2019 by calculating the average values of VV and VH of all the obtained SAR images.

Finally, all the Sentinel-2, Sentinel-1, and Luojia-1 data were resampled to 10 m spatial resolution to get the multisource data composite, which corresponded to the fine resolution of Sentinel-2. The multisource data composite had 17 bands in total (14 for Sentinel-2, 2 for Sentinel-1, and 1 for Luojia-1).

D. Machine Learning Algorithms

To find out the best classifier that is more suitable and robust for provincial land cover classification, we here include a group of machine learning algorithms including classification and regression trees (CART), minimum distance (MD), random forest (RF), and support vector machine (SVM) to conduct model-to-model comparisons of mapping performance. These algorithms have been widely used for remote sensing classification [52]–[55]. A brief description of these inclusive models is provided below.

CART is an umbrella term used to refer to the decision tree classifier, first introduced by Breiman *et al.* [56] in 1984. CART identifies relationships between a single continuous response (dependent variable) and multiple, continuous and/or discrete, explanatory (independent) variables, through a binary recursive partitioning process, where the data are split repeatedly into increasingly homogeneous groups (nodes), using combinations of variables (rules) that best distinguish the variation of the response variable.

In MD classification, a sample (i.e., group of vectors) is classified into the class whose known or estimated distribution most closely resembles the estimated distribution of the sample to be classified [57]. The measure of resemblance is a distance measure in the space of distribution functions such as the Euclidean distance [58].

RF is a machine learning algorithm consisting of a large ensemble of regression trees. It is operated by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees [59], [60]. The majority "vote" of all the trees is used to assign a final class for each unknown. RF corrects for the overfitting problem of decision tree (CART) algorithms [61]. The relative importance of each band can be evaluated by systematically comparing the performance of the trees that use a specific band, and those that do not [62].

SVM is a supervised nonparametric learning technique aiming to determine the location of optimal decision boundaries separating different classes [63]. The nearest data points to the resulting hyperplane that are used to measure the margin are called support vectors [64], which takes the advantage of dealing with limited training sets and high-dimensional data.

In this study, the selected four algorithms were implemented using the *ee.classifier* package in GEE (*smileCart*, *minimumDistance*, *smileRandomForest,* and *libsvm* representing for CART, MD, RF, and SVM, respectively) to model the relationship between explanatory features and response land cover types. All models were tuned to derive the best-fitting parameters (see Table VIII).

E. Accuracy Assessment and Method Comparison

We used the calculated overall accuracy and Kappa coefficient [65] from the confusion matrix as major indicators to compare the classification performance of different models. We also calculated the variable importance derived from the RF model to differentiate the relative contribution of inclusive variables. Additionally, since we had included multisource datasets with different spatial resolutions into the classifiers, it would be useful to identify specific types of data sources with a higher contribution to the classification performance, thus gaining potential insights of data selections for regional land cover classification practices. Therefore, we conducted another set of classification comparison using different combinations of inclusive features.

1) Sentinel-1 (VV and VH).
2) Sentinel-2 60 m (B1 and B9).
3) Sentinel-2 20 m (B5, B6, B7, B8A, B11, and B12).
4) Sentinel-2 10 m (B2, B3, B4, and B8).
5) Sentinel-2 spectral indices (NDVI and NDWI).
6) Sentinel-2.
7) Sentinel-1&2 (Sentinel-1 and Sentinel-2).
8) All (Sentinel-1, Sentinel-2, and Luojia-1).

All the above-mentioned eight scenarios were trained and validated using the same samples as described in Section III-B. Moreover, we conducted a comparison between our results and several land cover products including the 1 km China's land-use/cover dataset (CLUD),[2] the 500 m MODIS land cover product (MCD12Q1),[3] the 30 m FROM-GLC30 product[4] [22], and the 10 m FROM-GLC10 product[5] [26] from the perspectives of classification accuracy and visual interpretation. The classification systems of CLUD and MCD12Q1 were converted to the adjusted classification system of this study (Section III-A) based on the cross-walking table in Table IX.

TABLE II
CLASSIFICATION RESULTS OF DIFFERENT MODELS IN TERMS OF OVERALL ACCURACY AND KAPPA COEFFICIENT

Model	Overall accuracy (%)	Kappa coefficient
CART	80.21	0.77
MD	81.58	0.79
RF	86.12	0.84
SVM	66.80	0.61

IV. RESULTS

A. Comparison of Different Models

Table II compares the classification results of different models and Fig. 10 shows the corresponding land cover maps for Guangdong Province in 2019 based on the selected four classification models. In general, the RF model achieved the best performance for land cover classification (overall accuracy of 86.12% and Kappa coefficient of 0.84), followed by MD and CART models with slightly lower overall accuracies of 81.58% and 80.21% (Kappa coefficients of 0.79 and 0.77), respectively. In contrast, the SVM model yielded the lowest overall accuracies of 66.80% and the lowest Kappa coefficient of 0.61. As for specific land cover types, our results revealed that all the four models could classify forest and impervious surface types rather well, with user's accuracies higher than 80% [see Fig. 5(a)]. However, as for cropland, grassland, water, and bareland types, the difference among models was huge. For example, the user's accuracy of the SVM model for grassland was only 30.77%, much lower than the other three ones [see Fig. 5(a)]. In terms of producer's accuracy, the RF model achieved a promising result (ranging from 78.65% to 96.23%) for all the eight land cover types and had the highest producer's accuracy for grassland, shrubland, wetland, impervious surface, and bareland types among the four comparative models [see Fig. 5(b)]. Given the fact that the RF model outperformed the others, we adopted this model for land cover classification in Guangdong as well as for subsequent analysis.

B. Feature Contributions

Fig. 6 shows the relative importance of different input features derived from the RF importance analysis implementation in GEE. All training samples were used for the evaluation of feature importance here. Features derived from Sentinel-1 data (VV and VH) and Sentinel-2 spectral indices (NDVI and NDWI) greatly contributed to the classification process. In contrast, the

[2][Online]. Available: http://www.resdc.cn/

[3][Online]. Available: https://lpdaac.usgs.gov/products/mcd12q1v006/

[4][Online]. Available: http://data.ess.tsinghua.edu.cn/

[5][Online]. Available: http://data.ess.tsinghua.edu.cn/

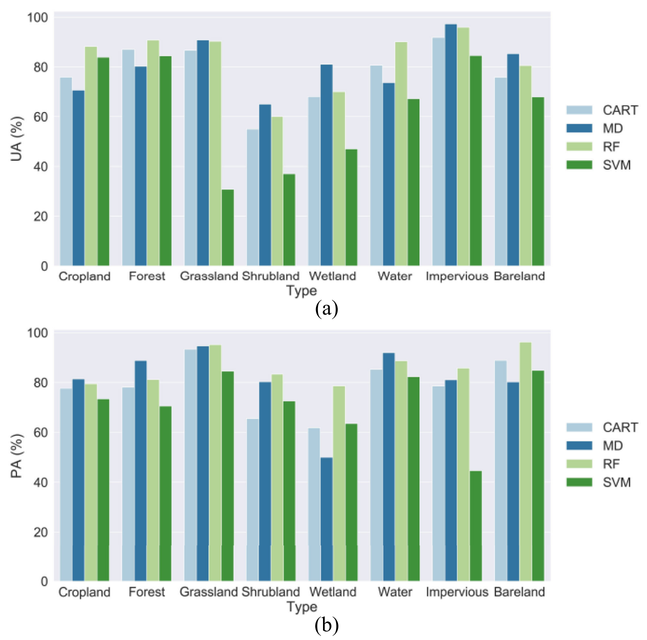

Fig. 5. Comparison of (a) user's accuracy and (b) producer's accuracy of different models for each land cover type. UA and PA denote user's accuracy and producer's accuracy, respectively.

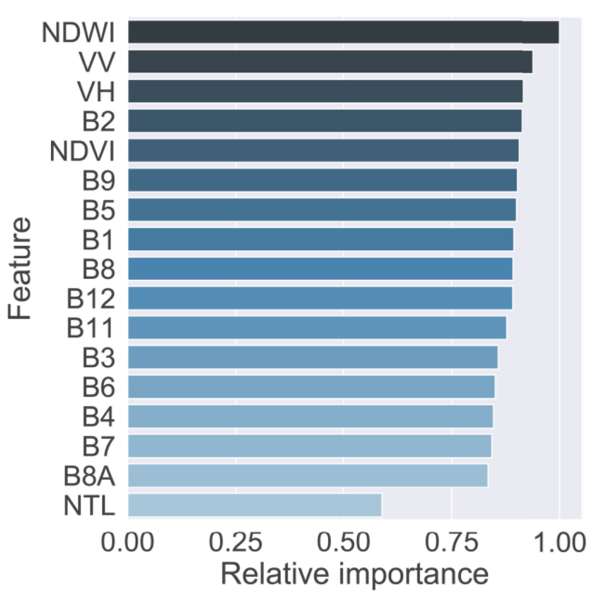

Fig. 6. Relative importance of features using all training samples in the RF model (ranging from 0 to 1). The larger the value, the more important the feature.

feature of NTL showed lower importance than other configurations. This could have been caused by the relatively low spatial resolution of Luojia-1 data (130 m), despite the fact that we had resampled all the data to 10 m for land cover classification.

We further compared the classification performance with different combinations of features (see Table III). For the multispectral data of Sentinel-2, when the spatial resolution of inclusive features grew from 60 to 10 m, the classification

TABLE III
CLASSIFICATION RESULTS OF DIFFERENT SCENARIOS IN TERMS OF OVERALL ACCURACY AND KAPPA COEFFICIENT

Scenario	Overall accuracy (%)	Kappa coefficient
Sentinel-1	46.60	0.38
Sentinel-2 60 m	57.87	0.51
Sentinel-2 20 m	76.91	0.73
Sentinel-2 10 m	78.49	0.75
Sentinel-2 spectral indices	64.81	0.59
Sentinel-2	82.13	0.79
Sentinel-1&2	83.09	0.80
All	86.12	0.84

accuracy increased accordingly (57.89% for 60 m, 76.91% for 20 m, and 78.49% for 10 m). For multisource remote sensing data, when features from Sentinel-1 and Sentinel-2 data were both considered into the RF model, a higher overall accuracy of 83.09% and a higher Kappa coefficient of 0.80 were obtained, as compared to scenarios when only Sentinel-1 or Sentinel-2 data was used. The highest overall accuracy (86.12%) and Kappa coefficient (0.84) occurred when all the features from the three data sources were utilized. Moreover, to assess the contribution of features to different land covers, we compared classification accuracy for each land cover type under different scenarios (see Table X). Results showed that Sentinel-1 data could distinguish water rather well, with a user's accuracy of 72.64% and a producer's accuracy of 70.19%. But it had a difficulty in classifying other land cover types, especially for grasslands, shrublands, and wetlands (see Table X). In contrast, when only using the Sentinel-2 data, the classification accuracy for each land cover type was relatively ideal, with a user's accuracy ranging from 63.00% to 95.45% and a producer's accuracy ranging from 66.54% to 94.87%. The contribution of Luojia-1 data was mainly reflected in improving the classification accuracy of cropland and grassland types. Compared with the scenario where only Sentinel-1 and Sentinel-2 data was used, after adding the feature of Luojia-1 data (i.e., Scenario: All), the user's accuracy of cropland and grassland types increased by 3.32% and 14.87%, respectively, and their producer's accuracy increased by 8.17% and 7.64%, respectively. These findings demonstrated that although the importance of features might differ among data sources with different spatial resolutions, they all contributed to the improvement of classification performance.

C. Land Cover Map for Guangdong, 2019

Fig. 7 presents the classified land cover map for Guangdong Province in 2019 using all the features derived from the three data sources and the RF model. Forests covered most of the study area, while the impervious surfaces mainly distributed on the Pearl River Delta. The Leizhou Peninsula was dominated by croplands. Statistically, within the 178405.85 km^2 land cover area of Guangdong Province in 2019, croplands accounted for 19.58%, forests accounted for 56.84%, grasslands accounted for 1.05%, shrublands accounted for 6.41%, wetlands accounted for 2.40%, waters accounted for 3.95%, impervious surfaces accounted for 8.85%, and barelands accounted for 0.92% (see Table IV). The confusion matrix based on the RF model, as

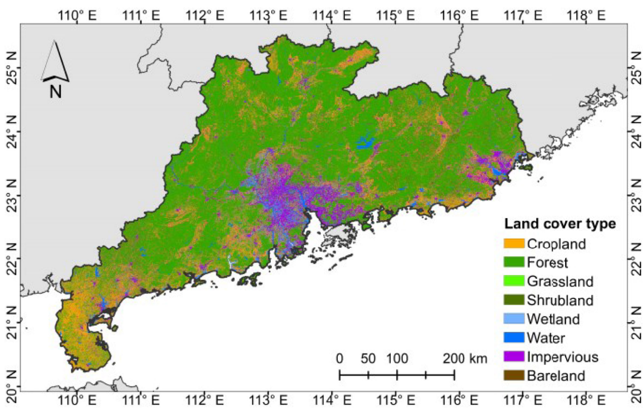

Fig. 7. Land cover map for Guangdong Province in 2019.

TABLE IV
STATISTICS OF LAND COVER COMPOSITION FOR GUANGDONG, 2019

Type	Area (km^2)	Proportion (%)
Cropland	34929.04	19.58
Forest	101408.90	56.84
Grassland	1867.05	1.05
Shrubland	11433.75	6.41
Wetland	4282.63	2.40
Water	7049.09	3.95
Impervious	15785.63	8.85
Bareland	1649.75	0.92

shown in Table V, denoted that the proposed framework could classify land cover types cropland, forest, grassland, water, impervious, and bareland rather well, with a user's accuracy ranging from 80.53% to 95.91% and a producer's accuracy ranging from 74.49% to 96.23%. In contrast, shrubland and wetland types had a relatively low classification with a user's accuracy of 60.00% and 70.00% and a producer's accuracy of 83.33% and 78.65%, respectively (see Table V).

D. Comparison With Other Land Cover Products

A quantitative assessment was carried out among three land cover products of the 30 m FROM-GLC30 in 2017, the 10 m FROM-GLC10 in 2017, and our 10 m results in 2019 (we did not include CLUD and MCD12Q1 here due to their coarse spatial resolutions). Using the same validation samples, we obtained an overall accuracy of 52.92% (Kappa coefficient: 0.45) for the FROM-GLC30 product and an overall accuracy of 71.34% (Kappa coefficient: 0.67) for the FROM-GLC10 product, both lower than our derived land cover results (overall accuracy of 86.12% and Kappa coefficient of 0.84). As shown in Table VI, except for special cases such as cropland and forest types, our results had a significant improvement over FROM-GLC30 in terms of user's accuracy (ranging from 24.09% to 74.21%) and producer's accuracy (ranging from 11.53% to 57.50%). As for FROM-GLC10, our results had a 0.42%–63.16% increase in user's accuracy for the eight land cover types (see Table VI). Among them, the accuracy of the analogy between forest, grassland, shrubland, and impervious surface was not much

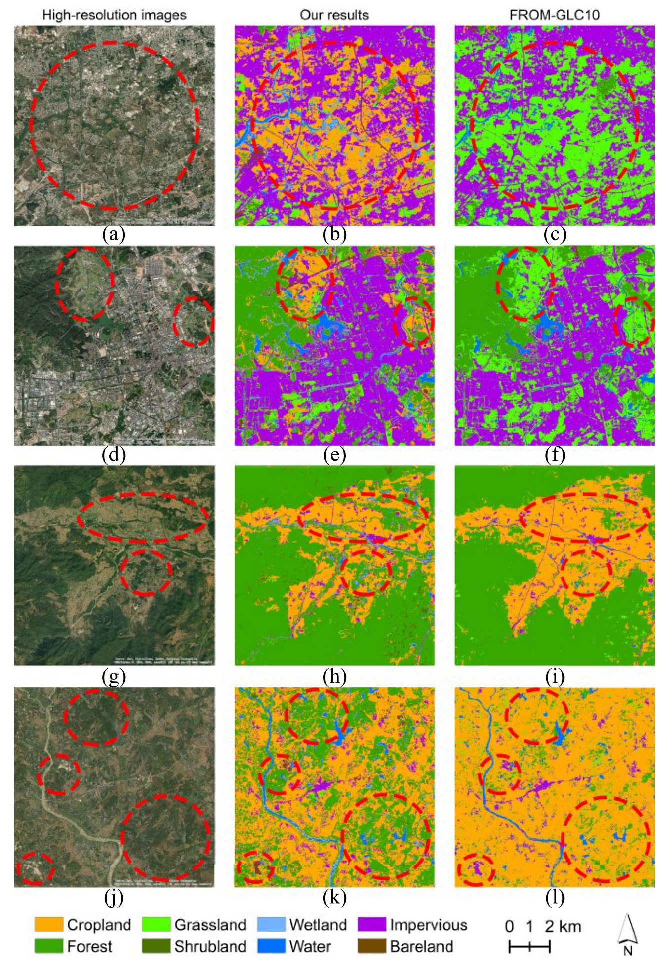

Fig. 8. High-resolution images, land cover maps of our results, and land cover maps of FROM-GLC10 in four zoomed areas: (a)–(c) Jieyang city (116°33′E, 23°34′N); (d)–(f) Dongguan city (114°04′E, 22°49′N); (g)–(i) Shaoguan city (113°30′E, 24°32′N); (j)–(l) Zhanjiang city (110°03′E, 21°37′N). Red ovals highlight some significant difference between our results and FROM-GLC10.

different, while our results significantly improved the accuracy of cropland, wetland, water, and bareland types (more than 10%).

Fig. 11 compares land cover maps among five products (the 1 km CLUD in 2018, the 500 m MCD12Q1 in 2018, the 30 m FROM-GLC30 in 2017, the 10 m FROM-GLC10 in 2017, and our 10 m results in 2019) in the Pearl River Delta area, one of the most densely urbanized regions in South China. All the different land cover systems were unified and converted to the classification system used in this study. In general, land cover maps of these five products had a good spatial agreement: the Golden Delta area located at the mouth of the Pearl River was dominated by impervious surfaces, while forests and croplands covered most of the remaining area (except for MCD12Q1). In terms of spatial resolution, three 10–30 m products (FROM-GLC30, FROM-GLC10, and our results) presented more spatial details, as compared to those of the 1 km CLUD and the 500 m MCD12Q1 products.

A detailed visual comparison further validated the effectiveness and robustness of our proposed framework. Fig. 8 displays high-resolution images, land cover maps of our results, and land

TABLE V
CONFUSION MATRIX OF LAND COVER CLASSIFICATION RESULTS DERIVED FROM THE RF MODEL. UA AND PA DENOTE USER'S ACCURACY AND PRODUCER'S ACCURACY, RESPECTIVELY

Type	Cropland	Forest	Grassland	Shrubland	Wetland	Impervious	Water	Bareland	UA (%)	PA (%)
Cropland	186	10	4	1	3	1	4	2	88.15	79.49
Forest	8	216	1	11	2	0	0	0	90.76	81.20
Grassland	13	4	176	0	0	0	1	1	90.26	95.14
Shrubland	5	31	3	60	0	0	1	0	60.00	83.33
Wetland	4	3	0	0	70	19	4	0	70.00	78.65
Water	4	0	0	0	13	181	3	0	90.05	88.73
Impervious	5	0	0	0	1	0	211	3	95.91	85.77
Bareland	9	2	1	0	0	3	22	153	80.53	96.23

TABLE VI
COMPARISON OF CLASSIFICATION ACCURACY FOR EACH LAND COVER TYPE UNDER DIFFERENT SCENARIOS. UA AND PA DENOTE USER'S ACCURACY AND PRODUCER'S ACCURACY, RESPECTIVELY

Type	UA (%)		
	FROM-GLC30	FROM-GLC10	Our results
Cropland	91.47	77.73	88.15
Forest	92.02	90.34	90.76
Grassland	17.95	84.10	90.26
Shrubland	20.00	59.00	60.00
Wetland	1.00	57.00	70.00
Water	65.67	78.11	90.05
Impervious	71.82	85.91	95.91
Bareland	6.32	17.37	80.53
Type	PA (%)		
	FROM-GLC30	FROM-GLC10	Our results
Cropland	37.84	60.74	79.49
Forest	59.84	70.49	81.20
Grassland	37.63	66.94	95.14
Shrubland	39.22	89.39	83.33
Wetland	33.33	100.00	78.65
Water	77.19	86.26	88.73
Impervious	63.45	66.08	85.77
Bareland	100.00	75.00	96.23

cover maps of FROM-GLC10 in four zoomed areas. Overall, land cover maps of our results and FROM-GLC10 both had a good agreement with the ground truth. But there were some subtle differences between these two products, especially between several easily confused land cover types such as forest, cropland, and grassland (highlighted as red ovals in Fig. 8). For instance, located in Rongcheng district of Jieyang city, the zoomed area in Fig. 8(a) was mainly covered by impervious surfaces and croplands. However, most of the croplands were wrongly classified as grasslands in FROM-GLC10 [see Fig. 8(c)]. Similar misclassifications were also discovered in Dongguan city [see Fig. 8(d)], where some small pieces of croplands were incorrectly identified as grasslands in FROM-GLC10 [see Fig. 8(f)]. Fig. 8(g) and (j) were two rural areas located in the city of Shaoguan and Zhanjiang, respectively, whose dominant land cover types were forests and croplands. As seen in FROM-GLC10, many forests were mistakenly recognized as croplands [see Fig. 8(i) and (l)]. Besides, a piece of barelands was correctly detected in our results [see Fig. 8(k)], but failed to be identified in the FROM-GLC10 product [see Fig. 8(l)].

V. DISCUSSION

A. *Strengths and Future Implications*

Taking Guangdong as a starting point, this study investigates the potential of accurately high-resolution land cover mapping at a large scale with the utilization of machine learning algorithms, multisource remote sensing data, and the GEE platform. Compared with other existing land cover products, a higher classification accuracy was obtained in our results using the same validation samples (see Table VI) and the derived land cover map was observed to better aligned with the ground truth (see Figs. 8 and 11).

Machine learning algorithms have been widely adopted for remote sensing classification [52]–[55]. In this study, using the same training and validation samples, we tested multimodel performance in land cover classification. Our results revealed that RF models achieved the best performance in both computational expanse and classification accuracy. Compared with other classifiers such as SVM, the selection of RF could yield a net increase in the overall accuracy of 19.32% and in the Kappa coefficient of 0.23 (see Table II). Since the protocol of RF is to assign the final class based on the majority "vote" of all built trees, such kind of strategy is especially suitable for processing high dimensional features. The efficiency and robustness of RF models in land cover mapping have also been observed and discussed in our previous experiments [66], [67].

To investigate the impact of data sources on classification performance, we calculated the relative importance of inclusive features from the RF model and compared eight classification scenarios under different features combinations. We discovered that the strategy of integrating multisource data sources, which provided complementary information to separate between different classes, contributed to classification results. For example, when features from Sentinel-1 and Sentinel-2 data were both considered into the RF model, a higher overall accuracy of 83.09% was obtained, as compared to scenarios when only Sentinel-1 or Sentinel-2 data was used (see Table III). We also discovered that the spatial resolution of data sources played an important role in land cover classification. As shown in Fig. 6, the most contributory features were derived from Sentinel-1 data (VV and VH) and Sentinel-2 spectral indices (NDVI and NDWI). Both of them had a spatial resolution of 10 m. On the contrary, the NTL feature derived from the 130 m Luojia-1 data showed the lowest importance among all (see Fig. 6). For the multispectral data of Sentinel-2, the classification results of using 10 m bands were better than that of using 20 m or 60 m bands (see Table III). These experiences suggested that leveraging multiple data sources with a higher spatial resolution was helpful in generating more accurate land cover maps.

The utilization of the GEE platform has greatly improved work efficiency. GEE consists of a multipetabyte analysis-ready data catalog colocated with a high-performance, intrinsically parallel computation service [31]. Unlike the traditional way that downloads massive data to the local for subsequent analysis, GEE enables users to directly process and analyze enormous multitemporal remote sensing data in Google's cloud and thus helps save much time. Apart from that, it provides a Git repository for storing, sharing, and script versioning of users' codes that leads to more user collaboration [31]. Following the proposed framework, multitemporal land cover mapping, ranging from regional to global scales, could be effectively done on GEE in the future once training and validation samples are ready.

B. Limitations and Uncertainties

Since we used the whole-year Sentinel-2 archive to derive the cloud-free Sentinel-2 mean composite for 2019, the seasonal effect of land cover types was neglected. This might have led to a misclassification among vegetation types (such as forest, grassland, and shrubland). Based on the differences in phenological characteristics, Zhu *et al.* [68] discovered that multitemporal optical images (i.e., images from different seasons) were helpful in better distinguishing vegetation types. Future work could consider seasonal compositing strategy [69] instead of the annual composite.

The time difference among datasets might cause uncertainty in the classification and comparison results. On the one hand, the target classification year of this study was 2019 while the Luojia-1 data was obtained in 2018. We did not select the Luojia-1 data for 2019 because the data coverage for that year is incomplete. On the other hand, even though we had tried our best to choose existing products corresponding to the target year 2019, there was still a one-year's or two-years' interval between our results and the four comparative land cover products (CLUD and MCD12Q1 for 2018, FROM-GLC30, and FROM-GLC10 for 2017). Nevertheless, given that land cover would not change dramatically over the years, we assumed the effect of this difference to be insignificant.

Despite its advantages, GEE is restricted by some limitations, which can be classified into three categories according to Tamiminia *et al.* [32]: 1) Computation. In our case, GEE would run into memory issues when processing is performed on a huge number of datasets. 2) Dataset. The Sentinel series is the only dataset on GEE that meets the scope of this research (with a spatial resolution of 10 m and the full coverage of Guangdong), and it is only available from 2015 to the present. 3) Algorithms. Even though GEE has provided a few classic machine learning algorithms (such as RF), deep learning algorithms are not yet supported directly by GEE. All those limitations of GEE would hinder its capacity for large-scale, high-resolution, historical, and accurate land cover mapping at present.

VI. Conclusion

Leveraging thousands of manually collected samples, multisource remote sensing data, machine learning algorithms, and the GEE platform, this study aims to develop a robust and cost-effective framework for accurately high-resolution land cover mapping at a large scale. Following this framework, we conducted land cover classification in Guangdong Province, China, using various open-source geospatial data layers (Sentinel-1 SAR, Sentinel-2 optical, and Luojia-1 NTL data) and a complete sample set (5000 for training and 1455 for validation). Results showed that RF models achieved the best performance for land cover classification (overall accuracy of 86.12% and Kappa coefficient of 0.84), as compared to other models of CART, MD, and SVM. We also found that features derived from Sentinel-1 data and Sentinel-2 spectral indices had a great contribution to the classification process while the feature of Luojia-1 data showed the lowest importance among all configurations. Using the same validation samples, we obtained a higher classification accuracy over other existing land cover products, and the derived map was discovered to be more correspond with the ground truth. This study systematically elucidates the role of classification methods and data sources in generating more accurate and reliable land cover maps. Our product could serve as critical variables for future applications such as biodiversity conservation, climate change, and urban planning. The derived 10 m resolution land cover map for Guangdong in 2019 can be downloaded.[6]

Appendix

TABLE VII
Band Information for the Sentinel-2 Data

Name	Resolution	Wavelength	Spectral description
B1	60 m	443.9nm (S2A) / 442.3nm (S2B)	Aerosols
B2	10 m	496.6nm (S2A) / 492.1nm (S2B)	Blue
B3	10 m	560nm (S2A) / 559nm (S2B)	Green
B4	10 m	664.5nm (S2A) / 665nm (S2B)	Red
B5	20 m	703.9nm (S2A) / 703.8nm (S2B)	Red Edge 1
B6	20 m	740.2nm (S2A) / 739.1nm (S2B)	Red Edge 2
B7	20 m	782.5nm (S2A) / 779.7nm (S2B)	Red Edge 3
B8	10 m	835.1nm (S2A) / 833nm (S2B)	NIR
B8A	20 m	864.8nm (S2A) / 864nm (S2B)	Red Edge 4
B9	60 m	945nm (S2A) / 943.2nm (S2B)	Water vapor
B11	20 m	1613.7nm (S2A) / 1610.4nm (S2B)	SWIR 1
B12	20 m	2202.4nm (S2A) / 2185.7nm (S2B)	SWIR 2

[6][Online]. Available: https://drive.google.com/drive/folders/1jMmM6VXK8yNpAQ0M4TZ0c3DSwxtGa46l?usp=sharing

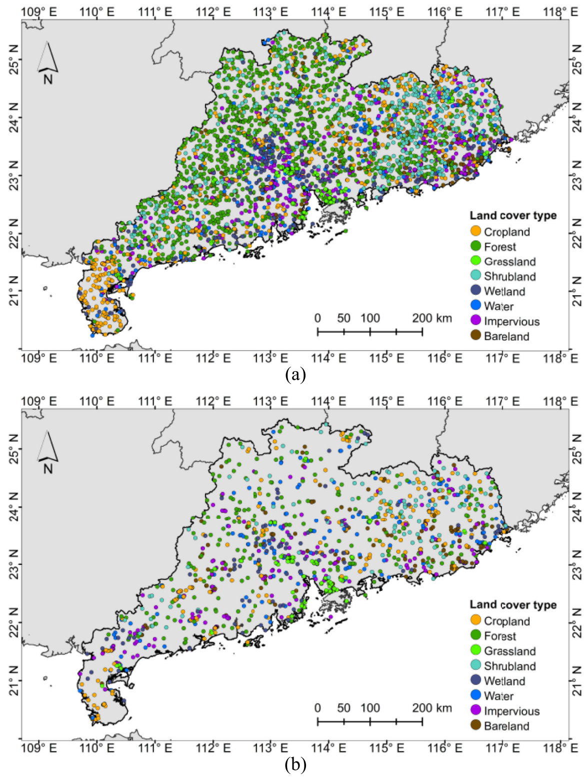

Fig. 9. Spatial distributions of (a) training samples and (b) validation samples for the study area of Guangdong in 2019.

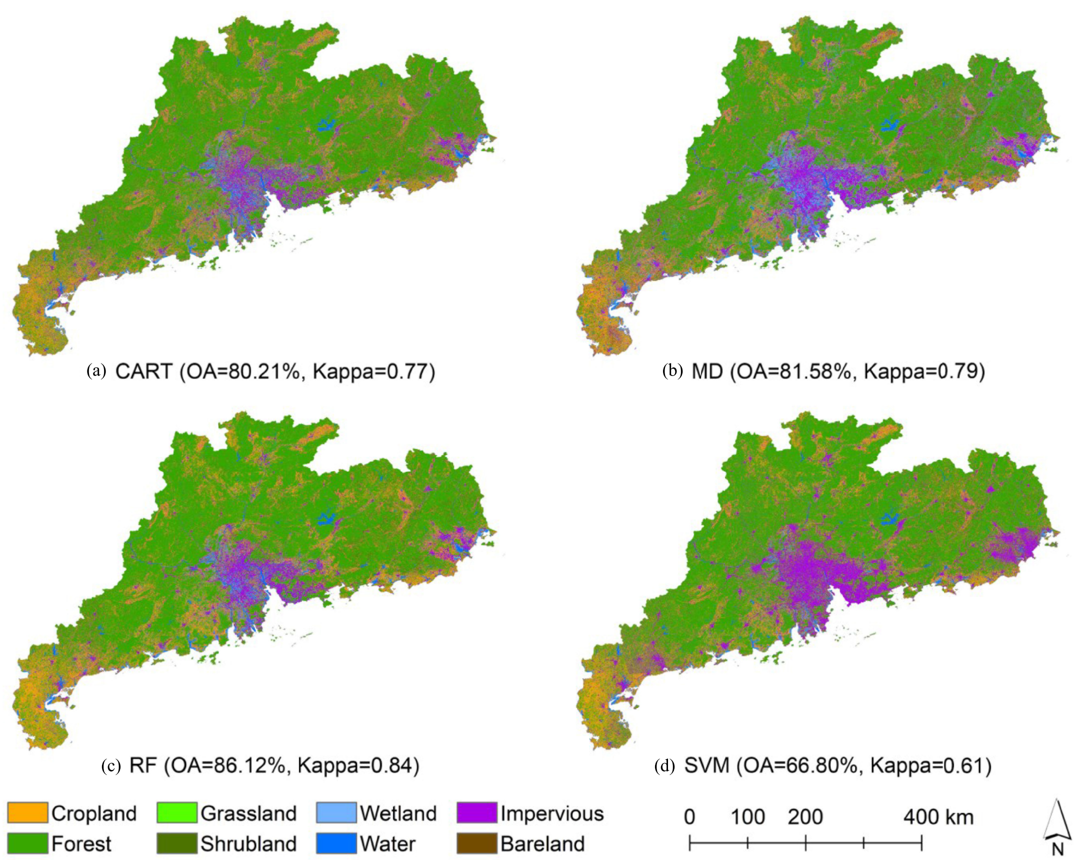

Fig. 10 Land cover maps for Guangdong Province in 2019 using different classification models. (a) CART. (b) MD. (c) RF. (d) SVM.

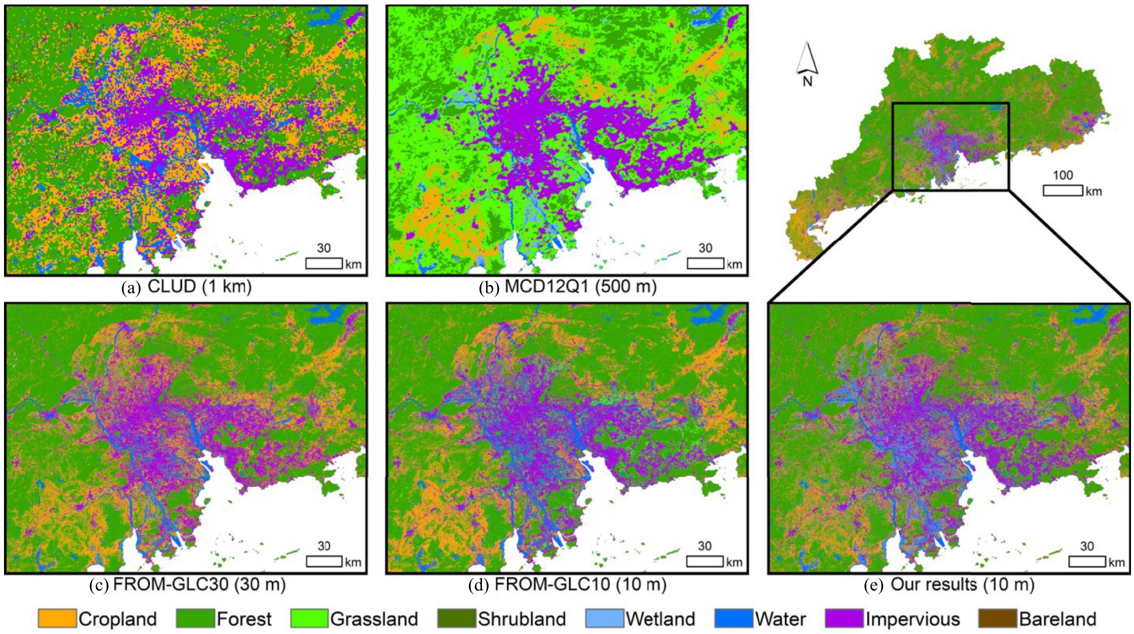

Fig. 11. Comparison of different land cover products in the Pearl River Delta area (21°52′N–23°53′N, 112°07′N E–114°49′E). (a) CLUD in 2018 (1 km), provided by Chinese Academy of Sciences. (b) MCD12Q1 in 2018 (500 m), provided by USGS. (b) FROM-GLC30 in 2017 (30 m), provided by Tsinghua University. (d) FROM-GLC10 in 2017 (10 m), provided by Tsinghua University. (d) Our results in 2019 (10 m).

TABLE VIII
IMPLEMENTATION PACKAGES AND OPTIMAL PARAMETERS FOR THE SELECTED FOUR MODELS

Model	Package in GEE	Optimal parameters
CART	*ee.classifier.smileCart()*	/
MD	*ee.classifier.minimumDistance()*	metric: 'mahalanobis'
RF	*ee.classifier.smileRandomForest()*	numberOfTrees: 120
SVM	*ee.classifier.libsvm()*	kernelType: ' RBF ', gamma: 0.01, cost: 1024

TABLE IX
CROSS-WALKING TABLE FROM CLUD AND MCD12Q1 TO THIS STUDY

Target type	Conversion type	
This study	CLUD	MCD12Q1
Cropland	Paddy	Croplands
	Dry land	Croplands mosaics
Forest	Forest	Evergreen needleleaf
	Sparse woods	Evergreen broadleaf
	Other woods	Deciduous needleleaf
		Deciduous broadleaf
		Mixed forest
Grassland	Dense grass	Woody savannas
	Moderate grass	Savannas
	Sparse grass	Grasslands
Shrubland	Shrub	Closed shrublands
		Open shrublands
Wetland	Tidalflat	Permanent wetlands
	Bottomland	
	Swampland	
Water	River and canal	Water bodies
	Lake	
	Reservoir and pond	
Impervious	Urban	Urban and built up
	Rural settlement	
	Industry-traffic land	
Bareland	Bare soil	Bare soil and rocks

TABLE X
COMPARISON OF CLASSIFICATION ACCURACY FOR EACH LAND COVER TYPE UNDER DIFFERENT SCENARIOS. UA AND PA DENOTE USER'S ACCURACY AND PRODUCER'S ACCURACY, RESPECTIVELY

Type	UA (%)			
	Sentinel-1	Sentinel-2	Sentinel-1&2	All
Cropland	58.77	82.94	84.83	88.15
Forest	68.49	90.34	90.76	90.76
Grassland	20.00	69.23	75.38	90.26
Shrubland	10.00	63.00	59.00	60.00
Wetland	4.00	69.00	69.00	70.00
Water	72.64	89.55	88.56	90.05
Impervious	57.27	95.45	95.91	95.91
Bareland	34.74	77.89	78.95	80.53
Type	PA (%)			
	Sentinel-1	Sentinel-2	Sentinel-1&2	All
Cropland	44.29	66.54	71.31	79.49
Forest	37.05	80.52	81.20	81.20
Grassland	35.45	88.82	87.50	95.14
Shrubland	20.41	84.00	84.29	83.33
Wetland	10.81	80.23	76.67	78.65
Water	70.19	87.80	88.12	88.73
Impervious	57.53	83.67	84.06	85.77
Bareland	58.93	94.87	95.54	96.23

ACKNOWLEDGMENT

The derived 10 m resolution land cover map for Guangdong in 2019 can be downloaded at https://drive.google.com/drive/folders/1jMmM6VXK8yNpAQ0M4TZ0c3DSwxtGa46l?usp=sharing. The authors would like to thank the two anonymous reviewers for their insightful comments and valuable suggestions, which have greatly improved this article.

REFERENCES

[1] G. M. Buchanan, A. Nelson, P. Mayaux, A. Hartley, and P. F. Donald, "Delivering a global, terrestrial, biodiversity observation system through remote sensing," *Conserv. Biol.*, vol. 23, no. 2, pp. 499–502, 2009.

[2] M. Jung, K. Henkel, M. Herold, and G. Churkina, "Exploiting synergies of global land cover products for carbon cycle modeling," *Remote Sens. Eniviron.*, vol. 101, no. 4, pp. 534–553, 2006.

[3] J. Yang *et al.*, "The role of satellite remote sensing in climate change studies," *Nature Climate Change*, vol. 3, no. 10, pp. 875–883, 2013.

[4] J. A. Foley *et al.*, "Global consequences of land use," *Science*, vol. 309, no. 5734, pp. 570–574, 2005.

[5] S. M. Sterling, A. Ducharne, and J. Polcher, "The impact of global land-cover change on the terrestrial water cycle," *Nature Climate Change*, vol. 3, no. 4, pp. 385–390, 2013.

[6] N. B. Grimm *et al.*, "Global change and the ecology of cities," *Science*, vol. 319, no. 5864, pp. 756–760, 2008.

[7] B. Xu, P. Gong, G. Biging, S. Liang, E. Seto, and R. C. Spear, "Snail density prediction for schistosomiasis control using IKONOS and ASTER images," *Photogramm. Eng. Remote Sens.*, vol. 70, no. 11, pp. 1285–1294, 2004.

[8] L. Liang *et al.*, "Combining spatial-temporal and phylogenetic analysis approaches for improved understanding on global H5N1 transmission," *PLoS One*, vol. 5, no. 10, 2010.

[9] S. Pauleit and F. Duhme, "Assessing the environmental performance of land cover types for urban planning," *Landscape Urban Plann.*, vol. 52, no. 1, pp. 1–20, 2000.

[10] C. Lo and D. A. Quattrochi, "Land-use and land-cover change, urban heat island phenomenon, and health implications," *Photogramm. Eng. Remote Sens.*, vol. 69, no. 9, pp. 1053–1063, 2003.

[11] M. C. Hansen, R. S. DeFries, J. R. Townshend, and R. Sohlberg, "Global land cover classification at 1 km spatial resolution using a classification tree approach," *Int. J. Remote Sens.*, vol. 21, no. 6–7, pp. 1331–1364, 2000.

[12] T. R. Loveland *et al.*, "Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data," *Int. J. Remote Sens.*, vol. 21, no. 6–7, pp. 1303–1330, 2000.

[13] M. A. Friedl *et al.*, "Global land cover mapping from MODIS: Algorithms and early results," *Remote Sens. Environ.*, vol. 83, no. 1–2, pp. 287–302, 2002.

[14] E. Bartholome and A. S. Belward, "GLC2000: A new approach to global land cover mapping from Earth observation data," *Int. J. Remote Sens.*, vol. 26, no. 9, pp. 1959–1977, 2005.

[15] M. A. Friedl *et al.*, "MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets," *Remote Sens. Environ.*, vol. 114, no. 1, pp. 168–182, 2010.

[16] R. Latifovic and D. Pouliot, "Multitemporal land cover mapping for Canada: Methodology and products," *Can. J. Remote Sens.*, vol. 31, no. 5, pp. 347–363, 2005.

[17] L. Lymburner *et al.*, "The national dynamic land cover dataset–technical report," Nat. Earth Observ. Group, Geosci. Australia, Symonston, ACT, 2011.

[18] L. Nrsc, "Land Use/Land Cover database on 1: 50,000 scale," Nat. Res. Census Project, LUCMD, LRUMG, RSAA, Nat. Remote Sens. Centre, ISRO, Hyderabad. Natural Res. Census-Land Use Land Cover, Ver, vol. 1, pp. 1–11, 2014.

[19] M. Herold, P. Mayaux, C. Woodcock, A. Baccini, and C. C. Schmullius, "Some challenges in global land cover mapping: An assessment of agreement and accuracy in existing 1 km datasets," *Remote Sens. Environ.*, vol. 112, no. 5, pp. 2538–2556, 2008.

[20] P. H. Verburg, K. Neumann, and L. Nol, "Challenges in using land use and land cover data for global change studies," *Global Change Biol.*, vol. 17, no. 2, pp. 974–989, 2011.

[21] P. Gong, "Assessment of GLC map accuracies using flux net location data," *Progr. Natural Sci.*, vol. 19, no. 7, pp. 754–759, 2009.

[22] P. Gong *et al.*, "Finer resolution observation and monitoring of global land cover: First mapping results with Landsat TM and ETM+ data," *Int. J. Remote Sens.*, vol. 34, no. 7, pp. 2607–2654, 2013.

[23] L. Yu, J. Wang, and P. Gong, "Improving 30 m global land-cover map FROM-GLC with time series MODIS and auxiliary data sets: a segmentation-based approach," *Int. J. Remote Sens.*, vol. 34, no. 16, pp. 5851–5867, 2013.

[24] L. Yu, J. Wang, X. Li, C. C. Li, Y. Y. Zhao, and P. Gong, "A multi-resolution global land cover dataset through multisource data aggregation," *Sci. China Earth Sci.*, vol. 57, no. 10, pp. 2317–2329, 2014.

[25] J. Chen *et al.*, "Global land cover mapping at 30 m resolution: A POK-based operational approach," *ISPRS J. Photogramm. Remote Sens.*, vol. 103, pp. 7–27, 2015.

[26] P. Gong *et al.*, "Stable classification with limited sample: Transferring a 30 m resolution sample set collected in 2015 to mapping 10 m resolution global land cover in 2017," *Sci. Bull.*, vol. 64, no. 6, pp. 370–373, 2019.

[27] C. Li *et al.*, "The first all-season sample set for mapping global land cover with Landsat-8 data," *Sci. Bull.*, vol. 62, no. 7, pp. 508–515, 2017.

[28] S. Dong, B. Gao, Y. Pan, R. Li, and Z. Chen, "Assessing the suitability of FROM-GLC10 data for understanding agricultural ecosystems in China: Beijing as a case study," *Remote Sens. Lett.*, vol. 11, no. 1, pp. 11–18, 2020.

[29] Y. Tu, B. Chen, T. Zhang, and B. Xu, "Regional mapping of essential urban land use categories in China: A segmentation-based approach," *Remote Sens.*, vol. 12, no. 7, 2020, Art. no. 1058.

[30] Y. Tu, H. Zhou, W. Lang, T. Chen, X. Li, and B. Xu, "A novel cross-sensor calibration method to generate a consistent nighttime lights time series dataset," *Int. J. Remote Sens.*, vol. 41, pp. 5482–5502, 2020.

[31] N. Gorelick, M. Hancher, M. Dixon, S. Ilyuschenko, D. Thau, and R. Moore, "Google Earth Engine: Planetary-scale geospatial analysis for everyone," *Remote Sens. Environ.*, vol. 202, pp. 18–27, 2017.

[32] H. Tamiminia, B. Salehi, M. Mahdianpari, L. Quackenbush, S. Adeli, and B. Brisco, "Google Earth Engine for geo-big data applications: A meta-analysis and systematic review," *ISPRS J. Photogramm. Remote Sens.*, vol. 164, pp. 152–170, 2020.

[33] H. Huang *et al.*, "Mapping major land cover dynamics in Beijing using all Landsat images in Google Earth Engine," *Remote Sens. Environ.*, vol. 202, pp. 166–176, 2017.

[34] Q. Li, C. Qiu, L. Ma, M. Schmitt, and X. X. Zhu, "Mapping the land cover of Africa at 10 m resolution from multi-source remote sensing data with Google Earth Engine," *Remote Sens.*, vol. 12, no. 4, 2020, Art. no. 602.

[35] P. Gong *et al.*, "Mapping essential urban land use categories in China (EULUC-China): Preliminary results for 2018," *Sci. Bull.*, vol. 65, no. 3, pp. 182–187, 2019.

[36] L. Carrasco, A. W. O'Neil, R. D. Morton, and C. S. Rowland, "Evaluating combinations of temporally aggregated Sentinel-1, Sentinel-2 and Landsat 8 for land cover mapping with Google Earth Engine," *Remote Sens.*, vol. 11, no. 3, 2019, Art. no. 288.

[37] B. Chen, Y. Jin, and P. Brown, "Automatic mapping of planting year for tree crops with Landsat satellite time series stacks," *ISPRS J. Photogramm. Remote Sens.*, vol. 151, pp. 176–188, 2019.

[38] A. J. Oliphant *et al.*, "Mapping cropland extent of Southeast and Northeast Asia using multi-year time-series Landsat 30 m data using a random forest classifier on the Google Earth Engine Cloud," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 81, pp. 110–124, 2019.

[39] Y. Xie, T. J. Lark, J. F. Brown, and H. K. Gibbs, "Mapping irrigated cropland extent across the conterminous United States at 30 m resolution using a semi-automatic training approach on Google Earth Engine," *ISPRS J. Photogramm. Remote Sens.*, vol. 155, pp. 136–149, 2019.

[40] J. Xiong *et al.*, "Automated cropland mapping of continental Africa using Google Earth Engine cloud computing," *ISPRS J. Photogramm. Remote Sens.*, vol. 126, pp. 225–244, 2017.

[41] Y. Xu *et al.*, "Tracking annual cropland changes from 1984 to 2016 using time-series Landsat images with a change-detection and post-classification approach: Experiments from three sites in Africa," *Remote Sens. Environ.*, vol. 218, pp. 13–31, 2018.

[42] B. Chen *et al.*, "A mangrove forest map of China in 2015: Analysis of time series Landsat 7/8 and Sentinel-1A imagery in Google Earth Engine cloud computing platform," *ISPRS J. Photogramm. Remote Sens.*, vol. 131, pp. 104–120, 2017.

[43] Y. H. Tsai, D. Stow, H. L. Chen, R. Lewison, L. An, and L. Shi, "Mapping vegetation and land use types in Fanjingshan national nature reserve using google earth engine," *Remote Sens.*, vol. 10, no. 6, 2018, Art. no. 927.

[44] C. Wang, M. Jia, N. Chen, and W. Wang, "Long-term surface water dynamics analysis based on Landsat imagery and the Google Earth Engine platform: A case study in the middle Yangtze River Basin," *Remote Sens.*, vol. 10, no. 10, 2018, Art. no. 1635.

[45] Y. Wang, J. Ma, X. Xiao, X. Wang, S. Dai, and B. Zhao, "Long-term dynamic of Poyang Lake surface water: A mapping work based on the Google Earth Engine cloud platform," *Remote Sens.*, vol. 11, no. 3, 2019, Art. no. 313.

[46] J. Louis *et al.*, "Sentinel-2 Sen2Cor: L2A processor for users," 2016, pp. 1–8.

[47] G. Zhang, J. Wang, Y. Jiang, P. Zhou, Y. Zhou, and Y. Xu, "On-orbit geometric calibration and validation of Luojia 1-01 night-light satellite," *Remote Sens.*, vol. 11, no. 3, 2019, Art. no. 264.

[48] X. Li, X. Li, D. Li, X. He, and M. Jendryle, "A preliminary investigation of Luojia-1 night-time light imagery," *Remote Sens. Lett.*, vol. 10, no. 6, pp. 526–535, 2019.

[49] J. Ou, X. Liu, P. Liu, and X. Liu, "Evaluation of Luojia 1-01 nighttime light imagery for impervious surface detection: A comparison with NPP-VIIRS nighttime light data," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 81, pp. 1–12, 2019.

[50] X. Li *et al.*, "Planar Block Adjustment for China's Land Regions with LuoJia1-01 Nighttime Light Imagery," *Remote Sens.*, vol. 11, no. 18, 2019, Art. no. 2097.

[51] W. Li, R. Dong, H. Fu, J. Wang, L. Yu, and P. Gong, "Integrating Google Earth imagery with Landsat data to improve 30 m resolution land cover mapping," *Remote Sens. Environ.*, vol. 237, 2020, Art. no. 111563.

[52] A. E. Maxwell, T. A. Warner, and F. Fang, "Implementation of machine-learning classification in remote sensing: An applied review," *Int. J. Remote Sens.*, vol. 39, no. 9, pp. 2784–2817, 2018.

[53] D. Lu and Q. Weng, "A survey of image classification methods and techniques for improving classification performance," *Int. J. Remote Sens.*, vol. 28, no. 5, pp. 823–870, 2007.

[54] R. DeFries and J. C.-W. Chan, "Multiple criteria for evaluating machine learning algorithms for land cover classification from satellite data," *Remote Sens. Environ.*, vol. 74, no. 3, pp. 503–515, 2000.

[55] J. Rogan, J. Franklin, D. Stow, J. Miller, C. Woodcock, and D. Roberts, "Mapping land-cover modifications over large areas: A comparison of machine learning algorithms," *Remote Sens. Environ.*, vol. 112, no. 5, pp. 2272–2283, 2008.

[56] L. Breiman, J. Friedman, R. Olshen, and C. Stone, *Classification and Regression Trees (Cole Statistics/Probability Series)*. Monterey, CA, USA: Wadsworth & Brooks, 1984.

[57] A. Wacker and D. Landgrebe, "Minimum distance classification in remote sensing," LARS Tech. Rep., pp. 25, 1972.

[58] J. Wolfowitz, "The minimum distance method," *Ann. Math. Statist.*, vol. 28, pp. 75–88, 1957.

[59] T. K. Ho, "Random decision forests," in *Proc. 3rd Int. Conf. Document Anal. Recogn.*, 1995, pp. 278–282.

[60] T. K. Ho, "The random subspace method for constructing decision forests," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 8, pp. 832–844, Aug. 1998.

[61] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[62] M. Pal, "Random forest classifier for remote sensing classification," *Int. J. Remote Sens.*, vol. 26, no. 1, pp. 217–222, 2005.

[63] I. Steinwart and A. Christmann, *Support Vector Machines*. Berlin, Germany: Springer, 2008.

[64] M. Pal and P. Mather, "Support vector machines for classification in remote sensing," *Int. J. Remote Sens.*, vol. 26, no. 5, pp. 1007–1011, 2005.

[65] R. G. Congalton, "A review of assessing the accuracy of classifications of remotely sensed data," *Remote Sens. Environ.*, vol. 37, no. 1, pp. 35–46, 1991.

[66] L. Yu *et al.*, "Meta-discoveries from a synthesis of satellite-based land-cover mapping research," *Int. J. Remote Sens.*, vol. 35, no. 13, pp. 4573–4588, 2014.

[67] P. Gong *et al.*, "A new research paradigm for global land cover mapping," *Ann. GIS*, vol. 22, no. 2, pp. 87–102, 2016.

[68] Z. Zhu, C. E. Woodcock, J. Rogan, and J. Kellndorfer, "Assessment of spectral, polarimetric, temporal, and spatial dimensions for urban and peri-urban land cover classification using Landsat and SAR data," *Remote Sens. Environ.*, vol. 117, pp. 72–82, 2012.

[69] G. Azzari and D. Lobell, "Landsat-based classification in the cloud: An opportunity for a paradigm shift in land cover monitoring," *Remote Sens. Environ.*, vol. 202, pp. 64–74, 2017.

Ying Tu received the B.S. degree in geographical information science from Sun Yat-sen University, Guangzhou, China, in 2019. She is currently working toward the Ph.D. degree in ecology at the Ministry of Education Key Laboratory for Earth System Modeling, Department of Earth System Science, Tsinghua University, Beijing, China.

Her research interests include land use/land cover change detection, nighttime lights remote sensing data processing, and environmental sustainability monitoring using earth observation data.

Wei Lang received the B. Eng. degree in urban planning from Harbin Institute of Technology, Harbin, China, in 2007, the M.S. degree in urban planning and policy design from Politecnico di Milano, Milan, Italy, in 2009, and the Ph.D. degree in urban planning and spatial analysis from the Hong Kong Polytechnic University, Hong Kong, in 2016.

From 2013 to 2014, he was a Visiting Scholar with the Department of City and Regional Planning, the University of California, Berkeley, CA, USA. He is currently working as an Associate Professor of urban and rural planning with the Department of Urban and Regional Planning, Sun Yat-sen University, Guangzhou, China. The areas of his academic interest include spatial structure and urban morphology, regional development and planning, spatial analysis, and community participatory planning.

Le Yu received the B.S. degree in geographic information systems and the Ph.D. degree in remote sensing both from the School of Earth Sciences, Zhejiang University, Hangzhou, China, in 2005 and 2010, respectively.

He is currently an Associate Professor with the Department of Earth System Science, Tsinghua University, Beijing, China. His research interests include global land cover/land use mapping, regional and global land change monitoring, and modeling for sustainable development.

Ying Li received the B.S. degree from Wuhan University of Technology, Wuhan, China, in 2018. She is currently working toward the master's degree at Sun Yat-sen University, Guangzhou, China.

Her research interests include economic geography and spatial data mining.

Junhao Jiang received the B.S. degree in geographical information science from South China Normal University, Guangzhou, China, in 2019. He is currently working toward the master's degree at Sun Yat-sen University, Guangzhou, China.

His major research interest focuses on urbanization and application of deep learning in rural development.

Yawen Qin received the B.S. degree in geography science from Wuhan University, Wuhan, China, in 2017. She is currently working toward the Ph.D. degree in human geography at Sun Yat-sen University, Guangzhou, China.

Her current research interests include urban and rural planning and land use mapping.

Jiemin Wu received the B.S. degree in remote sensing and geographic information system from Sun Yat-Sen University, Guangzhou, China, in 2018. She is currently working toward the master's degree at the Center of Integrated Geographic Information Analysis, Sun Yat-Sen University, Guangzhou, China.

Her major research interests include deep learning in image processing and spatial data mining.

Tingting Chen received the B. Eng. degree in urban planning from Sun Yat-sen University, Guangzhou, China, in 2008, the M.S. degree in urban planning from the University of Hong Kong, Hong Kong, in 2010, and the Ph.D. degree in urban planning and housing studies from the Hong Kong Polytechnic University, Hong Kong, in 2016.

From 2013 to 2014, she was a Visiting Scholar with the Department of City and Regional Planning, the University of California, Los Angeles, CA, USA. She is currently working as an Associate Professor of urban and rural planning with the Department of Urban and Regional Planning, Sun Yat-sen University, Guangzhou, China. The areas of her academic interests include community planning, community vitality, regional planning and governance, and housing studies.

Bing Xu received the M.S. and Ph.D. degrees in environmental science from the University of California, Berkeley, CA, USA, in 2001 and 2003, respectively.

From 2003 to 2004, she was an Assistant Professor of geography with the Texas State University, San Marcos, TX, USA, and from 2004 to 2008, with the University of Utah, Salt Lake City, UT, USA. She is currently a Professor with the Department of Earth System Science, Tsinghua University, Beijing, China. She is the author of six book chapters and 120 journal articles, including *PNAS*, *Annals of the AAG*, *Emerging Infectious Diseases*, *Remote Sensing of Environment*, *Nature Climate Change*, *ISPRS Journal of Photogrammetry and Remote Sensing*, *Environment International, Scientific Reports*, *PLoS NTD*, etc. Her research interests include global environmental change monitoring and health, particularly the understanding of evolution and transmission mechanism of avian influenza at the molecular and cross-continental scales using phylogenetic analysis and spatio-temporal modeling tools.

Prof. Xu served as the President of Chinese Professionals with Geographic Information Science from 2006 to 2007, and was on the Board of Directors from 2011 to 2013. She served as the Editor of the *Annals of Geographic Information Sciences*, in 2009. She was the recipient of a number of research awards, including William A. Fischer Memorial Scholarship Award from the American Society for Photogrammetry and Remote Sensing, Early Career Award from Remote Sensing Specialty Group, Association of American Geographers, and Superior Research Award from the University of Utah.