Incorporating synthetic aperture radar and optical images to investigate the annual dynamics of anthropogenic impervious surface at large scale

Yinyi LIN, Hongsheng ZHANG*, Hui LIN, Paolo Ettore GAMBA, Xiaoping LIU

Yinyi LIN, Institute of Space and Earth Information Science, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong, yinyilin@link.cuhk.edu.hk

Hongsheng ZHANG*, Department of Geography, The University of Hong Kong, Pokfulam, Hong Kong, zhanghs@hku.hk

Hui LIN, Key Lab of Poyang Lake Wetland and Watershed Research of Ministry of Education/School of Geography and Environment, Jiangxi Normal University, Nanchang, China, huilin@cuhk.edu.hk

Paolo Ettore GAMBA, University of Pavia, Italy, paolo.gamba@unipv.it

Xiaoping LIU, School of Geography and Planning, Guangdong Key Laboratory for Urbanization and Geo-simulation, Sun Yat-sen University, Guangzhou, China, liuxp3@mail.sysu.edu.cn

Abstract

The area, distribution, and temporal dynamics of anthropogenic impervious surface (AIS) at large scale are significant for environmental, ecological and socio-economic studies. Remote sensing has become an important tool for monitoring large scale AIS, while it remains challenging for accurate extraction of AIS using optical datasets alone due to the high diversity of land covers over large scale. Previous studies indicated the complementary use of synthetic aperture radar (SAR) to improve the AIS estimation, while most of them were limited to local and small scales. The potential of SAR for large scale AIS mapping is still uncertain and underexplored. In this study, first, a machine learning framework incorporating both optical and SAR data based on Google Earth Engine platform was developed for mapping and analyzing the annual dynamics of AIS in China. Feature-level fusion for SAR and optical data across large scale was tested applicable considering the backscattering coefficients, texture measures and spectral characteristics. Improved accuracy (averaged 2% increased overall accuracy and averaged 4% increased kappa coefficient) and better delineation between the bright impervious surface and bare land was observed comparing with using optical data alone. Second, comprehensive assessment was conducted using high-resolution samples from Google Earth, census data from China Statistic Yearbook and benchmark datasets from the GlobeLand30 and GHSL, demonstrating the feasibility and reliability of the proposed method and results. Last but not the least, we analyzed the spatial and temporal patterns of AIS in China from national, regional and provincial levels.

Keywords: impervious surface; urban land cover; SAR; multisource data fusion

1. Introduction

Since the reform and opening up policy was implemented in 1978 (Branstetter and Lardy 2006), the urbanization process in P.R. China has progressed rapidly, especially after the 1980s (Cao et al. 2017). Chinese urbanization can be defined as the increased number of cities and expansion of urban space, accompanied by the growth of urban population and extension of urban infrastructure (Gu et al. 2012). Urban expansion produces the most obvious changes on the surface of the earth, i.e., the change from natural land cover to the anthropogenic impervious surfaces. Impervious surface includes the natural impervious surfaces (rock, glaciers and etc.) and anthropogenic impervious surface (asphalt, concrete, brick and etc.).

Urban/regional studies are usually limited to urban impervious surface mapping. With regard to large scale impervious surface study, natural impervious surface exists, and rural impervious surface also exits. So in this study, anthropogenic impervious surface (AIS) is used for national/global study region, which is related to manmade materials in both urban and rural areas that is impenetrable and includes parking lots, roads, and buildings. This land cover transition not only has economic effects but also environmental effects (Weng 2012). AIS extent and acreage can be an indicator of urbanization (Gao et al. 2012). In addition, AIS has different reflectance characteristics, producing the heat island effect (Yuan and Bauer 2007) and affecting the hydrologic cycle (Brabec et al. 2002; McDonald et al. 2014). Recently, the 2030 agenda Sustainable Development Goal 11 illustrated the importance of safe, resilient and sustainable cities and human settlements, and AIS expansion pattern is a crucial part of this development strategy as well. Therefore, mapping and monitoring AIS are a priority.

The experiences, datasets and prospects of human settlement mapping at a global scale have previously been comprehensively summarized (Gamba and Herold 2009; Weng et al. 2018). The existing land cover products include the MODIS Land Cover Type product (500 m resolution, yearly), GlobCover (GLC) 2009 (300 m resolution, Jan 1, 2009 - Jan 1, 2010), the China impervious fraction map (500 m resolution, 2000, 2005, 2010), GlobeLand30 (30 m resolution, 2000 & 2010). The coarse spatial (300-500 m) resolution / temporal (5-10 year) resolution is insufficient for finer and more detailed analysis.

For large-scale national-level AIS extraction, lower to medium resolution optical remote sensing data are often used, for example, 500 m resolution to 1 km resolution: Moderate Resolution Imaging Spectroradiometer (MODIS) data (Deng and Wu 2013; Friedl et al. 2002; Schneider et al. 2009, 2010) and nighttime light (NTL) data (Elvidge et al. 2007; Guo et al. 2018; Xie and Weng 2017); 30 m resolution: Landsat data (Homer et al. 2015; Xian and Homer 2010); 10 m resolution: Sentinel-2 data. Single-source data for AIS extraction is inadequate (Jiang et al. 2009; Liu et al. 2014; Zhang et al. 2014). 10-30 m resolution dataset did give us higher details of the AIS, but it also suffered from data quality problems, including mosaic problems (Roy et al. 2010), cloud contamination (Ju and Roy 2008), and strip effects (Zhang et al. 2007). Besides, spectral confusion among different land cover types obscures the classification, for example, bright impervious surface and bare soil, dark

impervious surface with paddy field. NTL data, closely related to socioeconomic factors, have been used to estimate the urban extent due to its ability to reflect the extent of city lights (Imhoff et al. 1997). However, this dataset suffers from blooming effect and ephemeral light sources (Imhoff et al. 1997), leading to overestimation of the urban region. Liu et al. proposed the normalized urban areas composite index (NUACI) for global urban land mapping using both Landsat and NTL data (Liu et al. 2015; Liu et al. 2018) via the index-based method, while Goldblatt et al. combined these two datasets (Goldblatt et al. 2018) via machine learning. Gong et al. used NTL as a mask together with the Landsat for national impervious surface mapping, so as to reduce the confusions between impervious surfaces and barren (Gong et al. 2019). Still, there are uncertainties between the similar spectral signatures. Overall, these studies showed the effectiveness of combining multiple datasets.

Recently, the space-borne Synthetic Aperture Radar (SAR), with its physical backscattering characteristics (Nghiem et al. 2009) in all-weather condition, has performed well in UIS extraction. It's worth to note that SAR data also has its limitation of shadow and topography effect. Several studies have demonstrated the potential to combine variety of information on land properties to improve the limitations of single dataset. Zhang et al. (2014) synergistically combined optical and SAR data to improve the mapping of impervious surfaces (Zhang et al. 2014). Later, different integration levels (pixel level, feature level, and decision level) (Zhang and Xu 2018) were explored using the two datasets. However, the multisource data integration process limits to regional scale, often 300-3000 km or individual plots (Joshi et al. 2016). Thanks to the Google Earth Engine (GEE) platform (Gorelick et al. 2017), planetary-scale analysis has become available, and the system provides Sentinel-1 and Landsat combination with large-area mapping capabilities. In 2019, Sun et al. used threshold-based method for impervious surface extraction using Sentinel-1 and Sentinel-2 at national scale (Sun et al. 2019). Gong et al. further applied the two datasets in arid regions and improved the accuracy by 15% (Gong et al. 2020). However, the method is sensitive to the threshold and empirical experiment. Hence, combining different sources of data improves AIS extraction accuracy but the combination method should be carefully considered. Whether transferring the regional multisource data classification method to large scale area is applicable or not still needs to confirm.

According to Earth observation archives, open access data show timeliness and diversity, which ensures high time frequency and high accuracy analysis. As in Figure 1, in recent years, SAR and multispectral joint observation became available, especially after 2014, when the Sentinel-1 sensor joined the group. Sentinel-1 images show great potential to support large region land cover mapping because of both the image quantity and quality. Altogether, optical data (Landsat series), SAR data (Sentinel-1), nighttime light data (DMSP and VIIRS), SRTM DEM provide a wealth of spectral, backscattering, radiance, and terrain information.

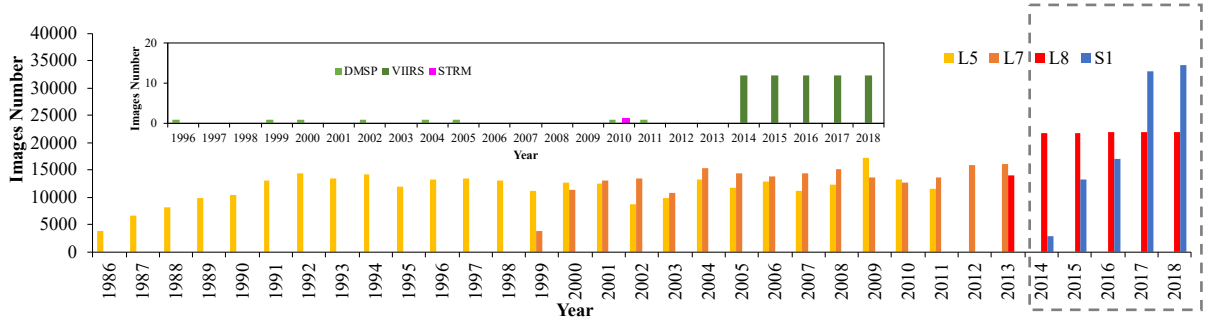


Figure 1. Data availability from 1986 to 2018 covering the study area. L stands for Landsat and S for Sentinel.

In this study, we develop a methodological framework of jointly using Sentinel-1, Landsat, nighttime light, and SRTM DEM data to estimate the large scale AIS and investigate the annual dynamics of AIS in China from 2015 to 2018 at 30 m resolution at the national level. Class reliability and the time consistency test were set up to generate the time series results. Then, the quality of the time series AIS product was assessed with the assistance of high-resolution samples on Google Earth, census data from the National Bureau of Statistics and benchmark datasets. By using the validated maps, it is possible to explore spatial and temporal changes in AIS products at the national, regional and city levels from 2015 to 2018.

2. Study area and datasets

Covering latitudes from 18° to 54° N and longitudes from 73° to 135° E, China has large north-south and sea-land differences, including complex and diverse climate systems (Herman 1966) (Figure 2). At the national level, we map the impervious surface of the whole China. Besides, seven regions are chosen for further analysis: Northeast (NE), Eastern (E), Central North (CN), Central (C), Central South (CS), Southwest (SW) and Northwest (NW). At the provincial level, 30 cities, in mainland China – are examined.

To cover the large area of 9,600,000 square kilometers, the datasets (see Table 1) include all available Tier 1 Landsat data from 2015 to 2018 (with a total of 87,685 scenes of Landsat images), 8 scenes of Global Radiance-Calibrated Nighttime Lights Version 4 (DMSP) images, 48 scenes of Visible Infrared Imaging Radiometer Suite (VIIRS) Stray Light Corrected Nighttime Day/Night Band Composites Version 1 images, 97,790 scenes of Sentinel-1 SAR GRD images and 1 scene SRTM image. The nighttime light and SRTM images were used for potential anthropogenic impervious surface (PAIS) mask extraction. Within the PAIS, the spectral features of Landsat data and backscattering coefficient features of Sentinel-1 data were used for the land cover classification.

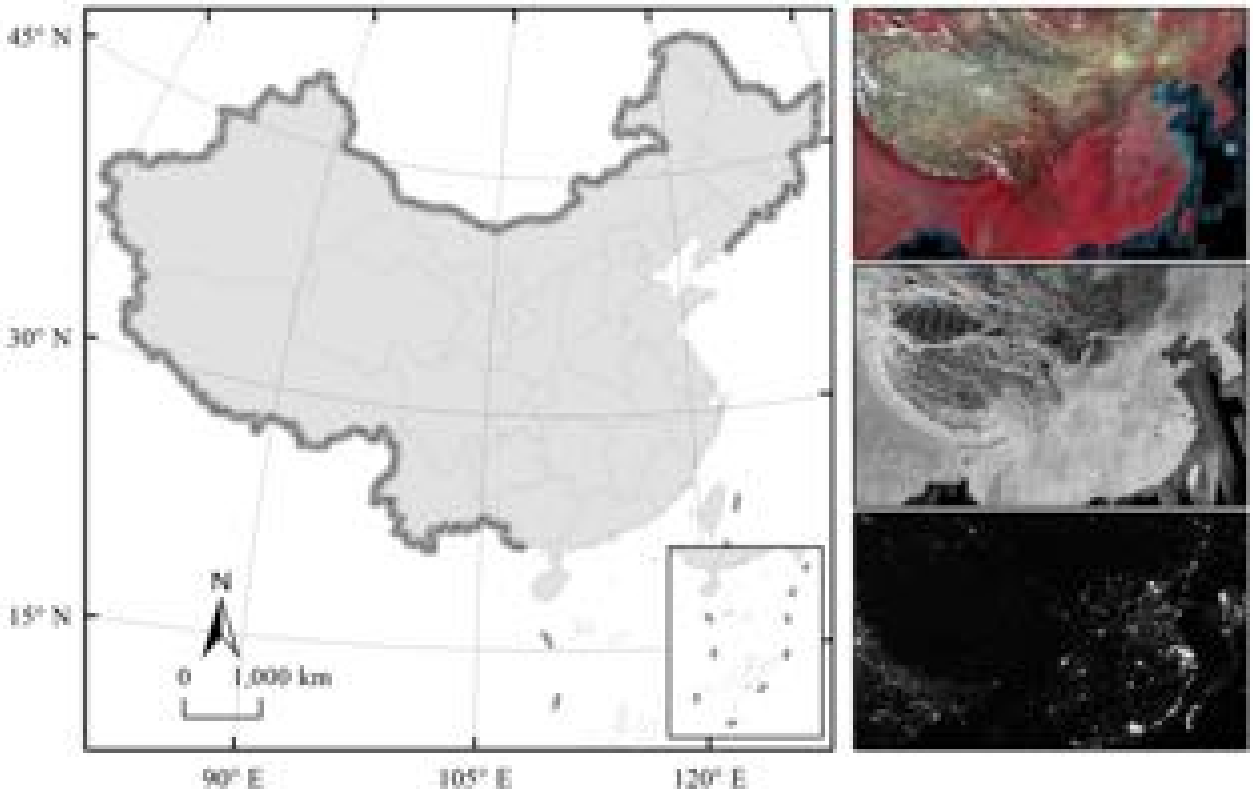


Figure 2. Research area and three datasets. The research area is China. The three datasets are in 2016. Landsat data are shown in B5, B4, B3 band composition. Sentinel-1 data are shown in VV polarization. NTL data are shown with the 'cf_cvg' band.

135

136 Table 1. Sources of the Landsat, Sentinel-1, NTL and SRTM data. Data are listed according to their
137 temporal availability.

Data Type	Product	Available Time	Scenes
Landsat	USGS Landsat 8 Collection 1 Tier 1 and Real-Time data TOA Reflectance	Jan 1, 2015 - Dec 31, 2018	87685
Sentinel-1	Sentinel-1 SAR GRD: C-band Synthetic Aperture Radar Ground Range Detected, log scaling	Jan 1, 2015 - Dec 31, 2018	97790
NTL	Global Radiance-Calibrated Nighttime Lights Version 4, Defense Meteorological Program Operational Linescan System (DMSP)	Mar 16, 1996 – Jul 31, 2011	8
	VIIRS Stray Light Corrected Nighttime Day/Night Band Composites Version 1 (VIIRS)	Jan 1, 2014 - Dec 31, 2018	60
SRTM	SRTM Digital Elevation Data Version 4 (srtm.csi.cgiar.org)	Feb 11, 2000 - Feb 22, 2000	1

3. Methodology

The overall framework for the extraction of AIS is shown in Figure 3. For one-year classification, Landsat and Sentinel-1 data were used according to the data availability. NTL and SRTM collection were used to generate the potential anthropogenic impervious surface mask. For the time series result, the time criterion of class reliability and time consistency was set up as in section 3.3. The four datasets are available in GEE, and the processing framework has been developed using the embedded code editor (https://code.earthengine.google.com). When different sources of data were combined, inputs were registered to the Mercator map coordinates (EPSG:3857), with a 30 m spatial resolution for processing and analysis.

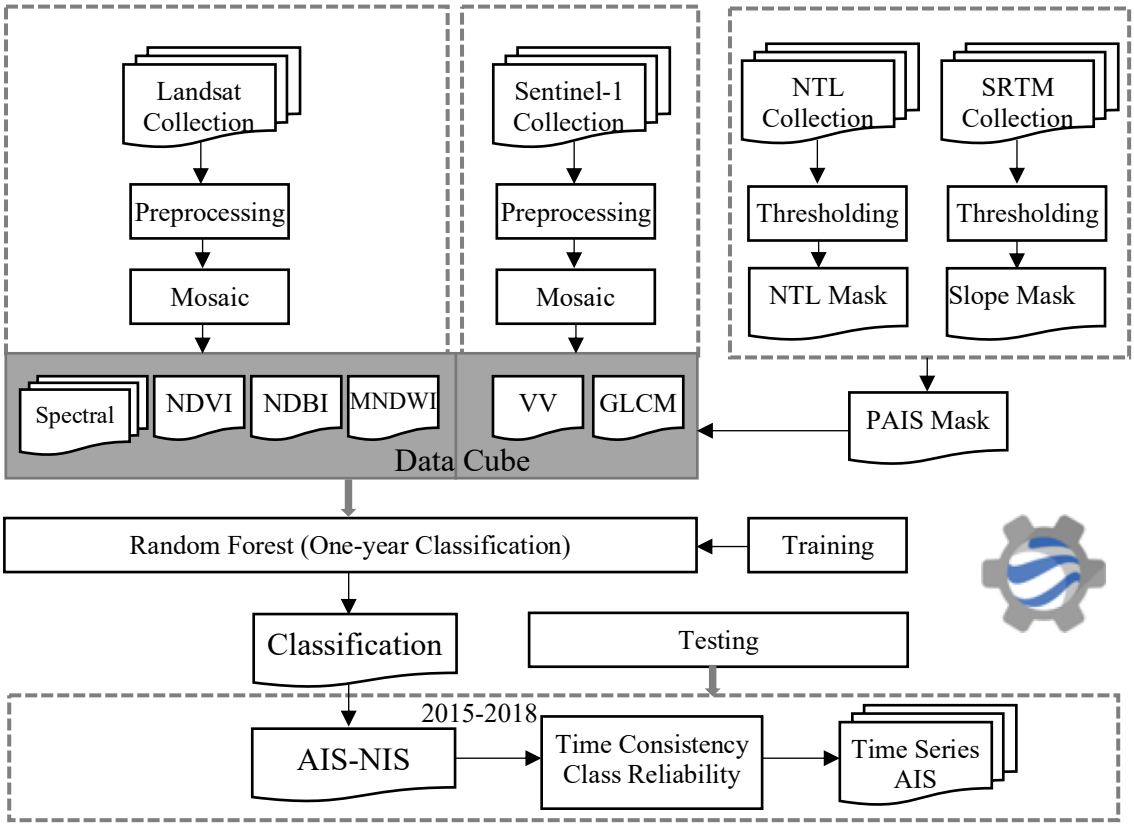


Figure 3. The proposed methodological framework in this study, PAIS: potential anthropogenic impervious surface; NIS: Non-impervious surface

3.1 Preprocessing of optical and SAR data

3.1.1 Cloud mask and mosaic for Landsat data

Landsat Collection 1 Tier 1 top of atmosphere (TOA) reflectance data were used in this research. These data have the high data quality and are ready for time series analysis. The value of the data represents the scaled, calibrated at-sensor radiance. On an annual basis (e.g., 2015-01-01 to 2015-12-31), we chose the available sensor collection observations of Landsat 8. Clouds were identified by evaluating the cloud score using the cloudiness indicators, assuming that clouds are bright in blue bands, visible bands, and infrared bands, cool in temperature, and the snow was excluded using the normalized difference snow index. A cloud score value equal to 0 stands for clear conditions, while 100 represents complete cloud coverage. Accordingly, pixels with cloud score greater than 50 were masked. Because our study area covered a large extent, the threshold of 50 ensured greater available data coverage. If all the Landsat images were contaminated by clouds over the same area throughout the year, there would be no available values for that area. Fortunately, most of the pixels had enough clear-sky images. Only a limited number of single pixels were never clear, and they were assigned a value of 0, indicating they are missing data. These 0-value pixels appeared as transparent in the images in the GEE platform and were filled via post-classification processing.

We assumed that the impervious surface change is relatively insignificant within a year, which is acceptable for national scale study. With regards to the climate zones like the moderate continental and monsoon climate zones, the spectral characteristics of vegetation and bare soil are easily affected by seasonal effects, i.e., the leaf on/off season(Deng et al. 2017). In tropical and subtropical monsoon climate zones, the seasonal effect of vegetation is relatively smaller but the cloud free observation is less due to the cloudy and rainy climate (Zhou et al. 2019). The optical dataset of mountain plateau climate zone will suffer from mountainous shadow influence. Rapid urbanization, phenological climate zones, and seasonal climate changes attributed to the short-term changes in urban area from our previous research (Zhang et al. 2018), which should be first considered in the mosaicking process. The mean value within a year interval (e.g., 2015-01-01 to 2015-12-31) for mosaicking was considered for the relatively stable inter-year impervious surface. Besides, the extreme regional phenological effects and terrain influence can be reduced. The data collection in this study

was mosaicking together using the rule of mean value at each pixel dimension. Therefore, if there were multiple images for the same area throughout the year, the mean value of each pixel was used.

Then, three indices, the normalized difference vegetation index (NDVI), normalized difference built-up index (NDBI) and modified normalized difference water index (MNDWI), which performed well in urban area extraction (Hansen et al. 2013; Liu et al. 2018; Xu 2008; Xu 2010), were calculated for the following classification.

3.1.2 Polarimetric backscattering coefficients generation for SAR data

The dataset covers the entire study area since 2015 so we can use the available observations from 2015 (e.g., 2015-01-01 to 2015-12-31) on annual basis for our research. The collection in this study was the ground range detected scenes with a resolution of 10 m. The preprocessing step (implemented as the Sentinel Toolbox by GEE) is ready for further study, including thermal noise removal, radiometric calibration, terrain correction and decibels conversion. Vertical-vertical (VV) with both ascending and descending modes were merged. Once the analysis-ready collection was obtained, temporal mean mosaicking was applied to each pixel within the collection as equation (1). The temporal mean value considers the time series effect, which can further reduce the speckle noise (Quin et al. 2014).

$$\sigma_{VV} = \frac{1}{n}\sum_{i=1}^{n} \sigma_{VV}^{i} \quad (1)$$

Where σ_{VV}^{i} refers to the *i* th VV backscattering coefficient observation at each pixel, *n* refers to the number of available observations within a year.

Then grey-level co-occurrence matrix (GLCM) (Conners et al. 1984; Haralick et al. 1973) was computed from VV backscattering coefficients (σ_{VV}) for later use. The results were automatically registered with Landsat data with 30 m resolution for processing.

3.1.3 Potential anthropogenic impervious surface mask

We proposed that the area where AIS areas are located, called potential anthropogenic impervious surface (PAIS), which is obtained by considering any area with a sufficient level of nighttime (Shi et al. 2014), and relatively flat slope (Sun et al. 2019). Two masks, including the NTL mask and slope mask, were thus generated for the PAIS extraction. The

VIIRS Stray Light Corrected Nighttime Day/Night Band Composites Version 1 data was used for the NTL mask generation. Indeed, because of the similarities of bare soil and BIS, spectral characteristics are not sufficient to discriminate between the two land cover types. However, the NTL data are sensitive to light; AIS reflects light because of its social attributes while other land cover doesn't. Since the PAIS should cover the urban area as much as possible, we used the whole time series nighttime light data for the PAIS generation. For DMSP average DNB radiance 'avg_vis', we first calculated the sum value at pixel dimension of the collection (from 1996-03-16 to 2011-07-31), then for each pixel, we used the urban and nonurban threshold, 1 empirically to set the DMSP mask. For VIIRS average DNB radiance 'avg_rad', we did the same process as DMSP mask to generate the VIIRS mask. Finally, we merged the two masks to compute the final NTL mask as in equation (2).

$$M_{NTL} = M_{DMSP} \cup M_{VIIRS} \quad (2)$$

With regard to the SRTM dataset, we first generated the slope from the DEM data. Then, we defined the flat and non-flat area. The non-flat area refers to the rock area instead of urban fabric, both of which has high possibility of high backscattering. The threshold was empirically set 15° as in (Ban et al. 2015) to calculate the slope mask. For the data cube mentioned in section 3.1.1/3.1.2, we used the NTL mask (M_{NTL}) and slope mask (M_{slope}) successively to extract the PAIS as equation (3).

$$PAIS = M_{NTL} \cap M_{slope} \quad (3)$$

3.2 One-year classification

For each year interval (e.g., 2015-01-01 to 2015-12-31), we generated one classification of this specific year called one-year classification. The one-year classification was conducted within the PAIS. Spectral and vegetation indices of Landsat, as well as backscattering feature and textural features of Sentinel-1 will be generated for classification. For Landsat images, both the spectral reflectance and the spectral indices were calculated. Three indices, NDVI, which highlights the vegetation component, NDBI, which represents the AIS component, and MNDWI, which represents the water component, were put into the random forest classifier (Breiman 2001) for the whole region classification. The feature extraction from SAR images

and the classification with feature level fusion between SAR and optical data were described in detail in the two sections below.

3.2.1 Sentinel-1 VV backscattering coefficients and textural feature extraction

Grey-level co-occurrence matrix (GLCM) was computed from the SAR dual-polarimetric σ_{VV}. Then, the GLCM-based texture measures were registered to the Landsat data and resampled using the nearest neighbor sampling method from 10 m to 30 m resolution to construct the multisource data stack. After that, the SAR data were fused with Landsat images at the feature level using the proposed strategy in our previous research (Zhang and Xu 2018) with the random forest (RF) for one-year classification of AIS. Overall, we computed the following SAR textural features (Conners et al. 1984; Haralick et al. 1973) as in Table 2: angular second moment (*ASM*), contrast (*CON*), correlation (*CORR*), variance (*VAR*), inverse difference moment (*IDM*), sum average (*SAVG*), sum variance (*SVAR*), sum entropy (*SENT*), entropy (*ENT*), difference variance (*DVAR*), difference entropy (*DENT*), information measures of correlation (*IMCORR1 & 2*), maximal correlation coefficient (*MAXCORR*), dissimilarity (*DISS*), inertia (*INE*), cluster shade (*SHADE*), cluster prominence (*PROM*).

Table 2. GLCM-based SAR texture measures.

Texture	Equation	Texture	Equation
ASM	$f_{ASM} = \sum_i \sum_j \{p(i,j)\}^2$	*CON*	$f_{CON} = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{\substack{j=1 \\ \lvert i-j \rvert = n}}^{N_g} p(i,j) \right\}$
CORR	$f_{CORR} = \frac{\sum_i \sum_j p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}$	*VAR*	$f_{VAR} = \sum_i \sum_j (i-\mu)^2 p(i,j)$
IDM	$f_{IDM} = \sum_i \sum_j \frac{1}{1+(i-j)^2} p(i,j)$	*SVAG*	$f_{SVAG} = \sum_{i=2}^{2N_g} i p_{x+y}(i)$
SVAR	$f_{SVAR} = \sum_{i=2}^{2N_g} (i - f_{SENT})^2 p_x +_y (i)$	*SENT*	$f_{SENT} = -\sum_{i=2}^{2N_g} p_x +_y (i) log\{p_{x+y}(i)\}$
ENT	$f_{ENT} = -\sum_i \sum_j p(i,j) log(p(i,j))$	*DVAR*	$f_{DVAR} = variance\ of\ p_{x-y}$

DENT	$f_{DENT} = -\sum_{i=0}^{N_{g-1}} p_{x-y}(i) log\{p_{x-y}(i)\}$	*MAXCORR*	$f_{MAXCORR}$ $= (Second\ largest\ eigenvalue\ of\ Q)$ $Q(i,j) = \sum_k \frac{p(i,k)p(j,k)}{p_x(i)p_y(k)}$
IMCORR1 & 2	$f_{IMCORR1} = \frac{HXY - HXY1}{max\{HX,HY\}}$ $f_{IMCORR2} = (1 - exp[-2.0(HXY2 - HXY)])^{1/2}$ $HXY = -\sum_i \sum_j p(i,j) log(p(i,j))$ $HXY1 = -\sum_i \sum_j p(i,j) log\{p_x(i)p_y(j)\}$ $HXY2 = -\sum_i \sum_j p_x(i)p_y(j) log\{p_x(i)p_y(j)\}$		
DISS	$f_{DISS} = \sum_{k=0}^{N_g-1} k p_{x-y}(k)$	*INE*	$f_{INE} = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (i-j)^2 p(i,j)$
SHADE	$f_{SHADE} = \sum_{i=1}^{N_g-1} \sum_{j=1}^{N_g-1} ((i-\mu) + (j-\mu))^3 p(i,j)$	*PROM*	$f_{PROM} = \sum_{i=1}^{N_g-1} \sum_{j=1}^{N_g-1} ((i-\mu) + (j-\mu))^4 p(i,j)$

Where $p(i,j)$ denotes the (i,j) th entry in a grey-tone spatial-dependence matrix, N_g represents the distinct gray levels of the images.

A tree-based feature selection method, the extremely randomized tree (Extra-tree) (Geurts et al. 2006), was applied for the SAR feature selection. VV backscattering coefficients and different GLCM feature importance of 4-year training samples were computed for comparison. The importance was measured by the particular information gain. For a sample S and a split s, the importance measure is given by:

$$Importance(s,S) = \frac{2I_c^s(S)}{H_s(S)+H_c(S)} \quad (3)$$

Where $H_c(S)$ is the log entropy of the classification in S, $H_s(S)$ is the split entropy, $I_c^s(S)$ is the mutual information of the split outcome and the classification.

With regard to the GLCM parameters, 4 directions (-1, -1), (0, -1), (1, -1), (-1, 0) with lag parameter of 1 were used in this study. The directional bands of each matrix were averaged. Then, two major issues including the size of image block (Marceau et al. 1990) and random forest tree numbers will be further analyzed in the following random forest classifier (Puissant et al. 2005; Zhang et al. 2014).

279 *3.2.2 Classification with feature level fusion of SAR and optical data*

280 In the PAIS, we carefully analysed the land cover types and determined the following land
281 cover composition including impervious surface (AIS), vegetation (VEG), water (WAT), and
282 others (including the composition of bare land, farm, and other land cover types). Historical
283 high-resolution images from Google Earth were used to assist the random sampling with
284 visual interpretation in the PAIS. In the spectral analysis, bright impervious surface (BIS) is
285 easily confused with desert soil, while dark impervious surface (DIS) is easily confused with
286 dry farmland. BIS, which is related to the newly built buildings, has high reflectance values.
287 In addition, DIS and BIS have different spectral characteristics, and it is important to choose
288 these two types of AIS for training. With regard to SAR image, it doesn't have the DIS and
289 BIS difference, so the DIS and BIS samples are all labeled as AIS. China is an agricultural
290 country, so in this study, farmland is chosen. Since the seasonally alternating farming exits,
291 dry farmland and wet farmland will be chosen as class others. It's worth to note that in the
292 training process, we used 10 percent of the two-class (AIS and NIS) samples for classification.
293 The random forest classifier (Breiman 2001) is chosen because of its fast computing ability
294 and flexibility with multidimensional data.

295

296 3.3 Time series classification using temporal filtering

297 Before a temporal filter was applied to the time series classification results, simple post-
298 classification processing was applied to each one-year classification map. A majority filtering
299 by counting the 0 and 1 values was applied within a 3-by-3 pixel window to remove the
300 NULL pixel values in the one-year classification results. Then, to improve the extracted AIS
301 estimate based on temporal information, a time consistency check was applied because the
302 radiometric conditions and sensors could have been different during the time period over the
303 same area (Li et al. 2015). In this study, two time criteria – class reliability and time
304 consistency – were proposed for the time series consistency check, in order to derive and
305 improve the final AIS estimate. The class reliability check assumes that the adjacent years

have similar classification results. For each pixel, this test considers the previous and the following year classification results and applies the majority rule as its final label. The probability distribution of each pixel according to the time series was calculated with sliding windows. The class reliability function is shown in Equation (4).

$$C_{(i,j)}^{t} = \begin{cases} \sum_{t-k}^{t+k} C_{(i,j)} \geq (k+1), & 1 \\ else, & 0 \end{cases} \quad (4)$$

where C is the binary AIS map (1 for AIS, 0 for pervious surfaces, NIS), t represents the present year, (i,j) represents the pixel location, and k is the size of one side of the sliding window. In this study, we tested different k values and calculated the accuracy. Since the classification result tends to be more similar closer to the target year, k was empirically set to 1. Thus, the size of the sliding window is 3, which means the results within 3 years were considered each time. The term t starts from 2017 and ranges from 2017 to 2016 since later years have multisource data that ensure higher accuracy than earlier years. For the time consistency test, we assumed that urbanization is irreversible from 2015 to 2018 and only allow NIS to transform into AIS. We used the accumulated AIS probability to determine the current land cover type. The rule is that, for each pixel the year t from 2018 to 2016, if the pixel in the previous year and the following year is AIS, then the pixel in the present year is an AIS. The time consistency check was conducted with Equation (5).

$$C_{(i,j)}^{t} = \begin{cases} \sum_{t}^{end} C_{(i,j)} \geq (end-t) \text{ and } \sum_{t-1}^{end} C_{(i,j)} \geq (end-t+1), & 1 \\ else, & 0 \end{cases} \quad (5)$$

where *end* is the final year, in our case, is 2018. For a common year, not the first year, the consistency check requires the data from both the current year and previous year. For the first year, since the previous year is missing in the study period, we have to set some special criteria for this case. As in the beginning year 2015, we set the rule that if two consecutive years (i.e., 2015 and 2016) are AIS, the pixel of 2015 is AIS.

3.4 Validation with ground reference data and public products

We used three data sources for validation, including high-resolution samples derived from Google Earth images, census data from the National Bureau of Statistics, and benchmark datasets. According to the (Stehman and Foody 2019), cluster sampling was used in this study.

The primary sampling unit was plots of prefecture-level city and the secondary sampling unit was the simple random samples within each plot. **First**, two-class reference samples were manually selected in study period according to the historical Google Earth high-resolution images. A total of 3249 stable random plots of prefecture-level city are chosen. There will be 5 random samples within each plot. Altogether, approximately 16245 samples (8115 AIS samples and 8130 NIS samples) were selected separately in 2018. AIS samples include the different types of buildings, roads and parking lots while the NIS samples include the vegetation (VEG), water (WAT), and cropland, bare land and other land cover types (others). We used the commonly used sample size formula as in (Stehman and Foody 2019):

$$n = \frac{z^2 p(1-p)}{d^2} \quad (6)$$

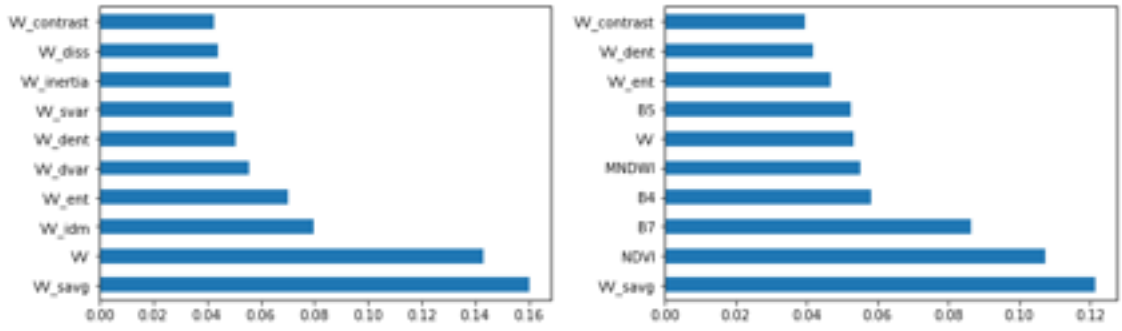
where $z = 1.96$ stands for a 95% confidence interval and d denotes the half-width of the confidence interval. p is the anticipated user's accuracy. In this study, we set the $\mathrm{d} = 0.02$ for the two-class classification and anticipated $p = 0.9$. So the anticipated AIS/NIS class number would be 864. In our cases,10 percent of the samples were used for training while the latter 90 percent were used for testing. There were a total of 1644 training samples (811 AIS training samples and 833 NIS training samples) and 14601 testing samples (7304 AIS testing samples and 7297 NIS testing samples). Finally, the overall accuracy (OA), user's accuracy (UA), producer's accuracy (PA) and Kappa coefficient (Stehman 1997) were calculated via the GEE platform. **Second**, census data from the China Statistic Yearbook were obtained from the National Bureau of Statistics. R-squared values were calculated between the extracted AIS data and the statistical urban area at the provincial level over 3 years (from 2015 to 2017). **Third**, two benchmark datasets, GlobeLand30 (www.globallandcover.com) (Jun et al. 2014) and GHSL (https://ghsl.jrc.ec.europa.eu/datasets.php) , were employed for validation. Two-class samples (AIS and NIS) were used to evaluate the two benchmark datasets. GlobeLand30 year 2010 and GHSL built-up grid year 2014 was used. The results were compared with the accuracy results of the corresponding AIS classification result in this study.

4. Results

4.1 SAR feature extraction

After the Extra-tree feature selection process, the importance of SAR features was slightly different from 2015 to 2018. To better evaluate the feature importance, all samples from the

whole period were compared in Figure 4. *SAVG* and VV feature of SAR data were in high feature importance. In Figure 4 (b), the top 10 important features were *SAVG, NDVI, B7, B4, MNDWI, VV, B5, ENT, DENT* and *CON* and the top 10 SAR features were *SAVG, VV, IDM, ENT, DVAR, DENT, SVAR, INE, DISS, CON*. To keep the result from overfitting, we didn't include nearly every possible feature, especially for the case in large scale mapping, which requires huge computational costs. The feature selection process can improve the prediction performance, provide faster and more cost-effective predictors, and better understanding of the generated data (Guyon et al. 2002). At the feature level, the proposed method included optical spectrum, vegetation indices, SAR VV backscattering coefficient σ_{VV} and *SVAG* textural information.



(a) (b)

Figure 4. Extra-tree feature selection results using all samples from 2015 to 2018, (a) SAR feature only, (b) multisource feature.

Then, the single source classification result was used as the baseline, and SAR backscattering coefficient as well as the different GLCM texture was added to evaluate the OA and Kappa. From Figure 5, we can conclude that not all GLCM texture feature worked positive to the impervious surface extraction. The proposed selected SAR features, together with the optical features, performed the best result (final column group), with OA larger than 0.97 and Kappa

larger than 0.95. Further parameter optimization was conducted to determine the effects of GLCM block size (neighborhood size) and RF parameter (RF tree number) (Figure 6). The fluctuation resulted from the samples variation among different years. With neighborhood size of GLCM becoming larger, the accuracy fluctuated and decreased gradually. With regard to RF tree number, the accuracy curve first improved then became stable. To simplify the classification complexity and computation time, we empirically set the GLCM neighborhood size of 1 and RF tree number of 50 in our experiment.

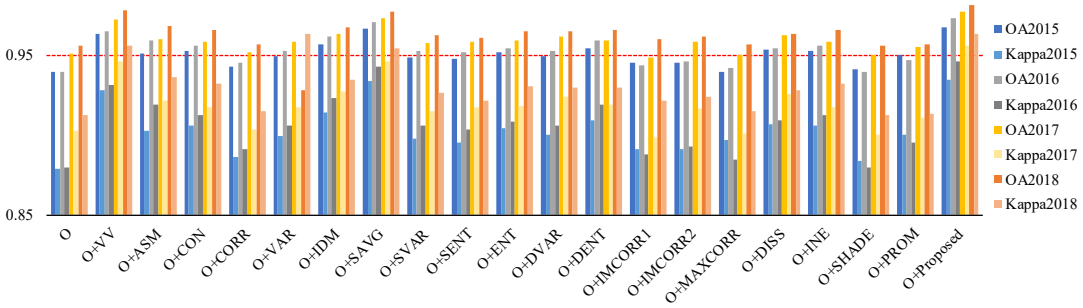
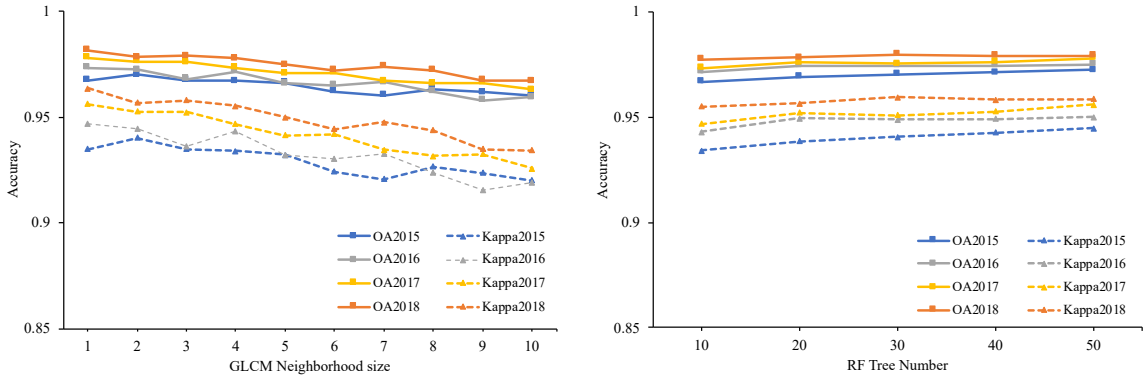


Figure 5. VV backscattering coefficients and texture feature analysis, O denotes optical data, including the spectral bands and indices, which work as the baseline.



(a) (b)

Figure 6. Parameter sensitivity analysis. (a) GLCM neighboring effect, 1-10: window size. (b) RF Tree number. 10-50: RF tree numbers.

4.2 Accuracy assessment of the AIS layer using various reference data

4.2.1 Accuracy assessment using samples from high-resolution Google Earth images

The OA, UA, PA and Kappa of the one-year classification result from 2015 to 2018 are demonstrated in Figure 7. Generally, an increase in OA, UA, PA and Kappa in multisource (MS) comparing to optical (O) or SAR (S) only is observed, which indicates the effectiveness of using multiple data sets of Landsat and Sentinel-1. We found that the proposed method provided averaged 2% increased OA, 3% increased UA, 1% increased PA and 4% increased Kappa comparing to using optical data only, and corresponding 7%, 6%, 8% and 14% increased accuracy comparing to using SAR data only. Paired t-test was conducted to test the significance between the data groups (multisource, optical and SAR data), and accuracy groups (OA, UA, PA and Kappa) within 4 years results (2015-2018). All groups demonstrated the p value less than 0.001, which confirmed that the improvement was statistically significant. Actually, the improvements located in the confusion area of AIS and others class. The physical backscattering characteristics of SAR data help to discriminate the AIS from bare land. The OA and Kappa of the national AIS data across the 2015-2018 was generally higher than 95%, indicating the reliability and stability of the national AIS product in different years.

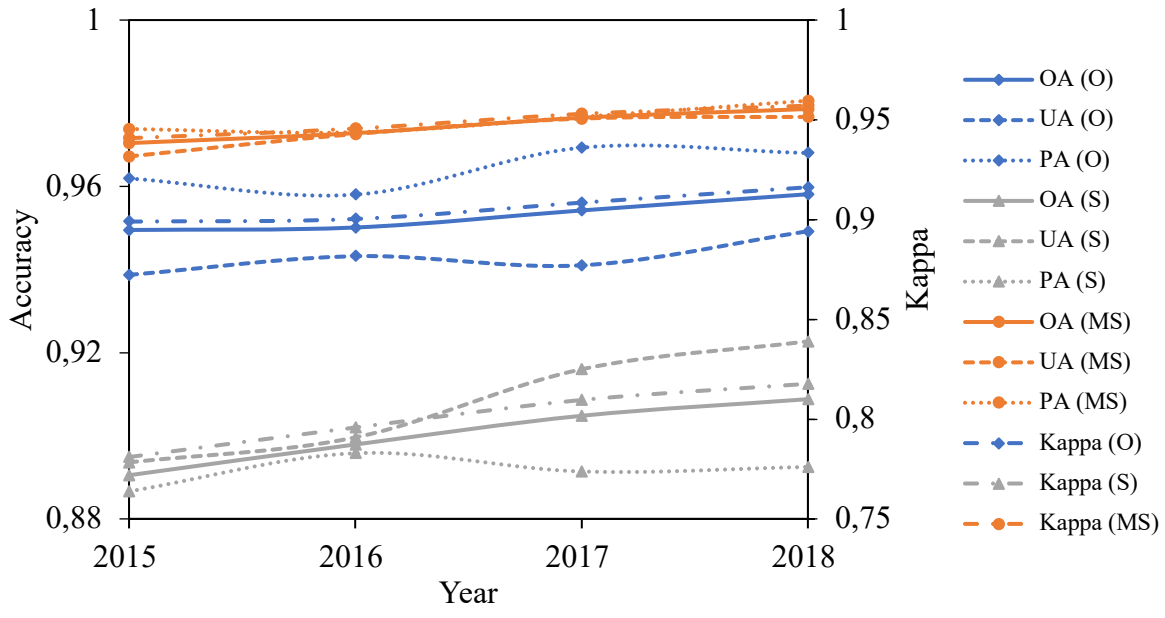


Figure 7. Overall accuracy (OA), user's accuracy (UA), producer's accuracy (PA) and Kappa coefficient for the impervious surface extraction (MS denotes multisource method, O denotes optical data only, S denotes SAR data only).

4.2.2 Accuracy assessment using census data from the National Bureau of Statistics

Census data from the National Bureau of Statistics were used in the assessment. Figure 8 shows the correlation of AIS area and built-up district (BD) (China Statistical Yearbook). Note that the definitions of the two terms are a little different: one refers to all the impermeable materials, while the other refers to the development area. Thus, their extents are not exactly the same. Considering the whole statistics from spatial and temporal scale, the correlation is significant with R square of 0.60. It should be mentioned that the BD was referred to the city urban built-up area, so its area was less than the AIS, which contained the rural area. Therefore, the census data showed the rationality of the AIS estimation.

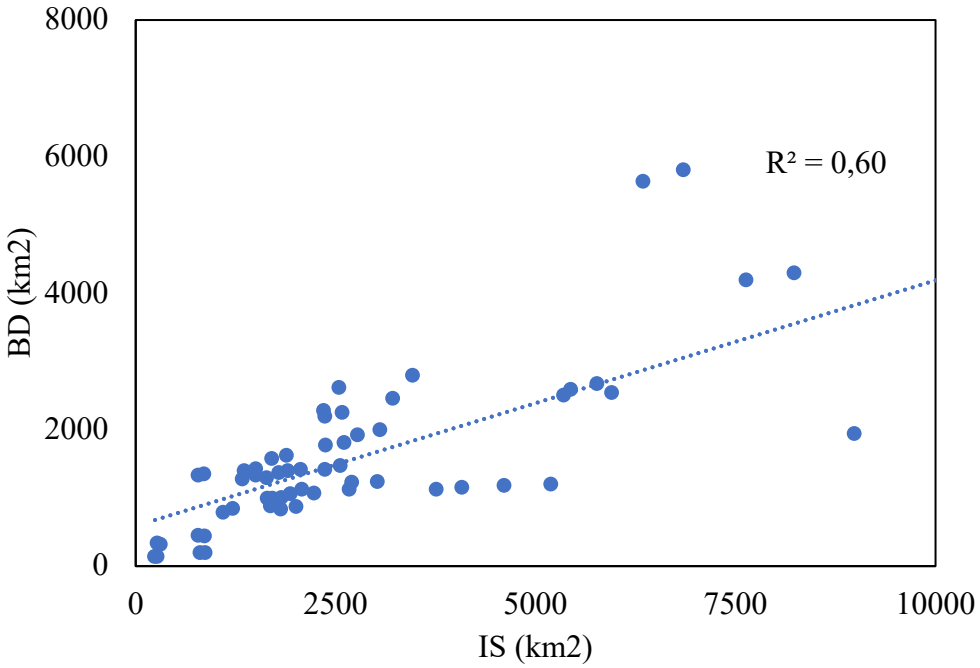


Figure 8. Comparison between the estimated impervious surface (AIS) area and built-up district (BD) area from China Statistical Yearbook in different provinces from 2015 to 2017.

4.2.3 Accuracy assessment using benchmark datasets

The GHSL and GlobeLand30 were used as benchmark datasets for our comparison. Two-class validation samples (as specified in 3.4) were used as ground truth to test the accuracy of different products. The results are shown in Table 3 with GHSL in 2014, GlobeLand30 in 2010, and AIS in 2015. Improved OA and kappa coefficient values were observed comparing

443 to GHSL and GlobeLand30. Six impervious surface morphology were chosen for detailed
444 analysis (Figure 9), including urban area, rural area, mountainous area 1, mountainous area 2,
445 delta and dessert from top to bottom. As we can see, GHSL kept good performance in urban
446 area, while in rural area, it underestimated the buildings in village. For GlobeLand30, since it
447 used the object-based method for classification, the impervious surface periphery was
448 smoother but it overestimated the impervious surface area when there was vegetation within
449 the settlement. Also, the impervious pattern was missed when the whole region was
450 determined as one category. We can see that AIS can not only classify the village settlements
451 but also keep good impervious surface pattern. Besides, in mountainous area, delta as well as
452 the dessert, the mapping result was also reliable. It's worth to note that SAR data contributed
453 to the impervious surface mapping since the strong sensitivity of backscattering
454 characteristics to man-made structures. Double bounce characteristic was shown in settlement,
455 which reflected as high VV intensity in the SAR images. To some extent, SAR data helps to
456 discriminate the bare land from impervious surface.
457

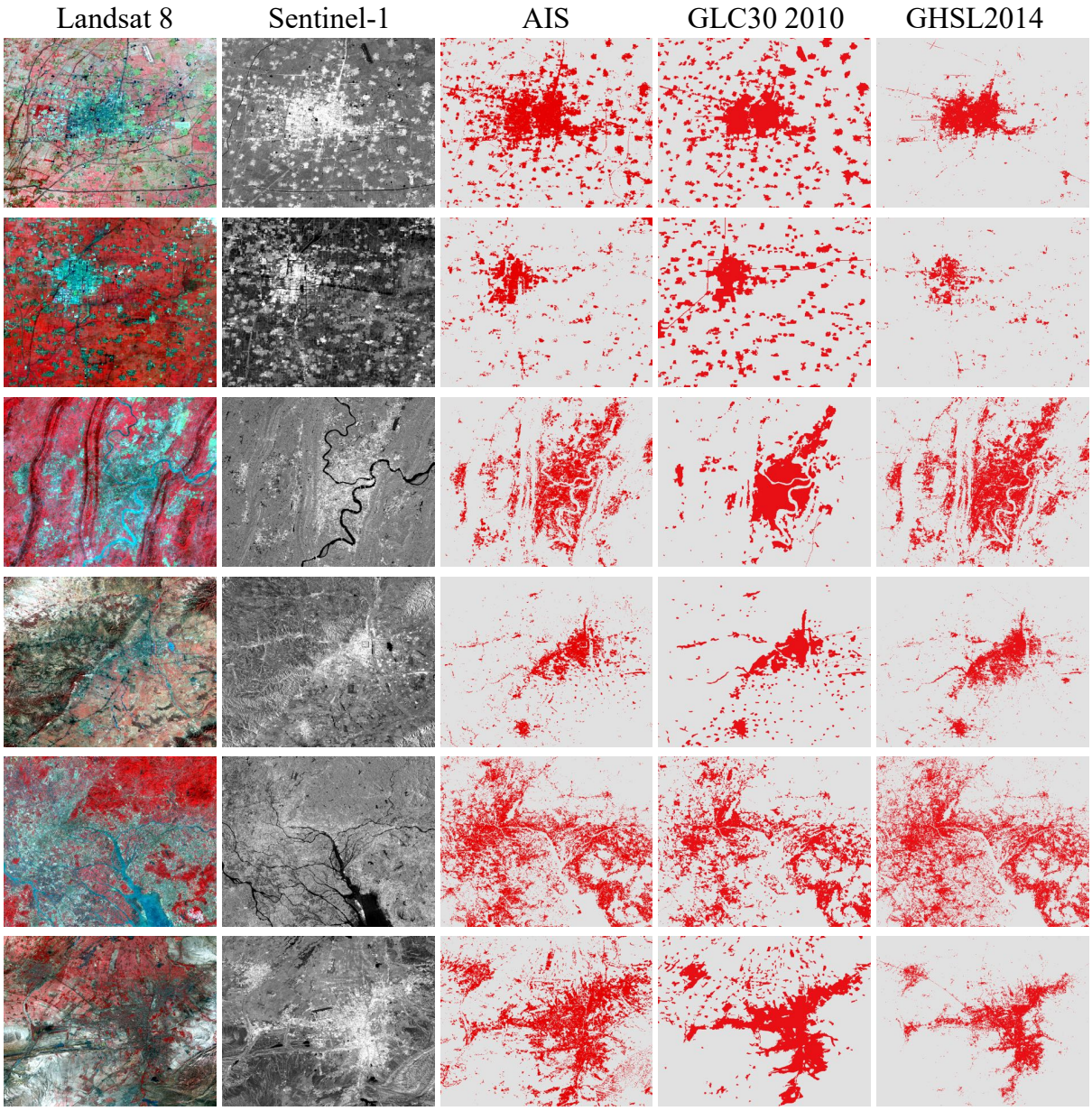


Figure 9. The comparison of different datasets, i.e., Landsat 8 (2015) in B5, B4, B3 composition, Sentinel-1 (2015) VV backscattering coefficient, AIS in this study, GlobeLand30, and GHSL. Impervious surface morphology: urban area, rural area, mountainous area 1, mountainous area 2, delta and dessert from top to bottom.

458

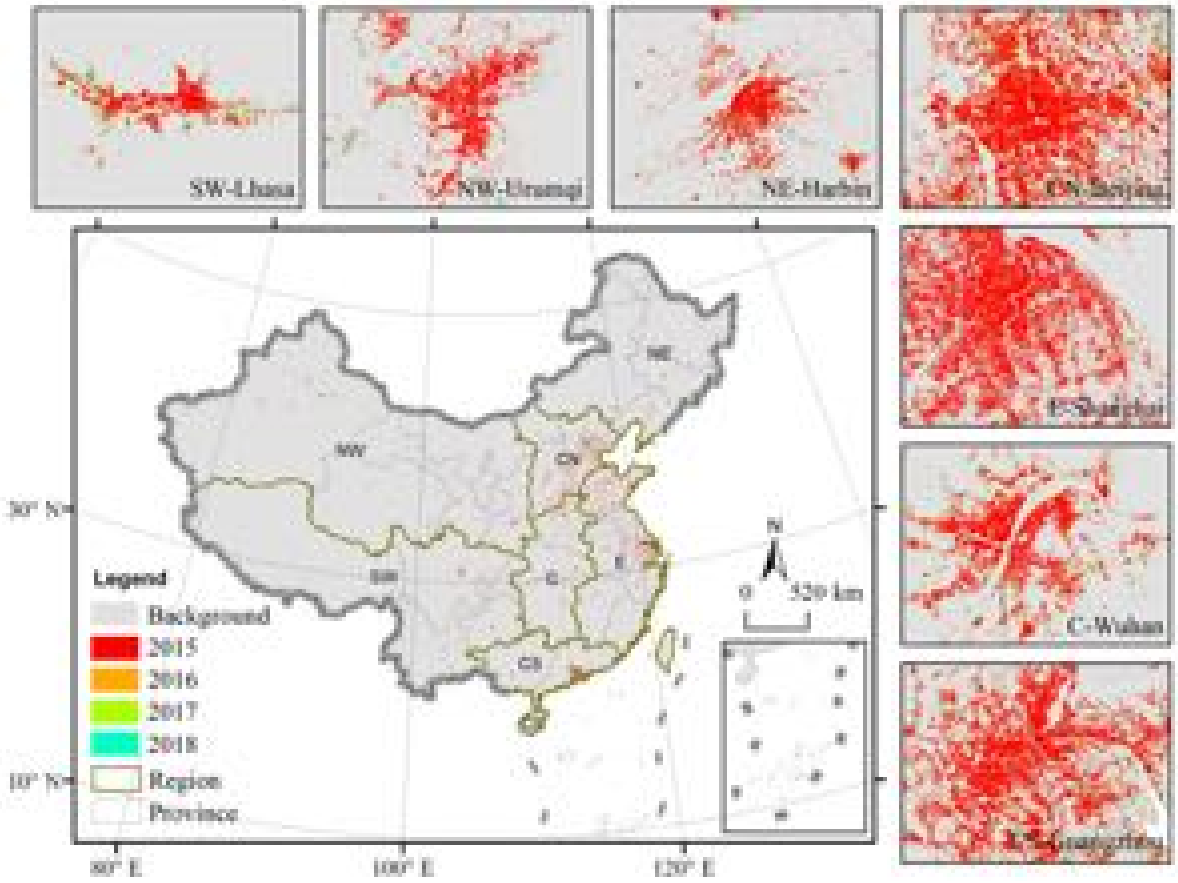
Table 3. Accuracy assessment between different datasets. The ground truth samples are two-class validation samples with 7316 impervious surface samples and 7293 non-impervious surface samples.

	GHSL	GlobeLand30	AIS 2015
OA (%)	89.79	86.55	97.51
UA (%)	93.72	85.85	97.48
PA (%)	85.33	87.68	97.71
Kappa	0.7959	0.7309	0.9501

4.3 Spatial and temporal characteristics of the impervious surfaces of China

4.3.1 The spatial and temporal dynamics of AIS at different levels

The AIS distribution fraction map over China from 2015 to 2018 is shown in Figure 10. Most AISs are distributed in the eastern and southern parts of China. Statistically, over the whole of China, the AIS area was 94588 square km in 2015 and increased to 131143 square km in 2018. To demonstrate the capability of explaining the annual AIS dynamics, seven representative cities of seven regions were selected for comparison of the relative area changes: Harbin in Northeast (NE), Shanghai in Eastern (E), Beijing in Central North (CN), Wuhan in Central (C), Guangzhou in Central South (CS), Lhasa in Southwest (SW) and Urumqi in Northwest (NW). Central North, Eastern, and Central South regions have dense and compact impervious surface pattern while Northwest, Northern and Central regions have relatively coarse and scattering impervious surface pattern. This pattern further demonstrates that the urbanization scale and speed in the eastern coastal area has been larger and faster than that in the central part of China. Southwest located on the Qinghai-Tibet Plateau and accounted for one-eighth of China's area. The AIS area there was relatively lower than that in other provinces because of its high altitude. Overall, the urbanization speed was the largest in eastern China, followed by central China, and the smallest in western China.



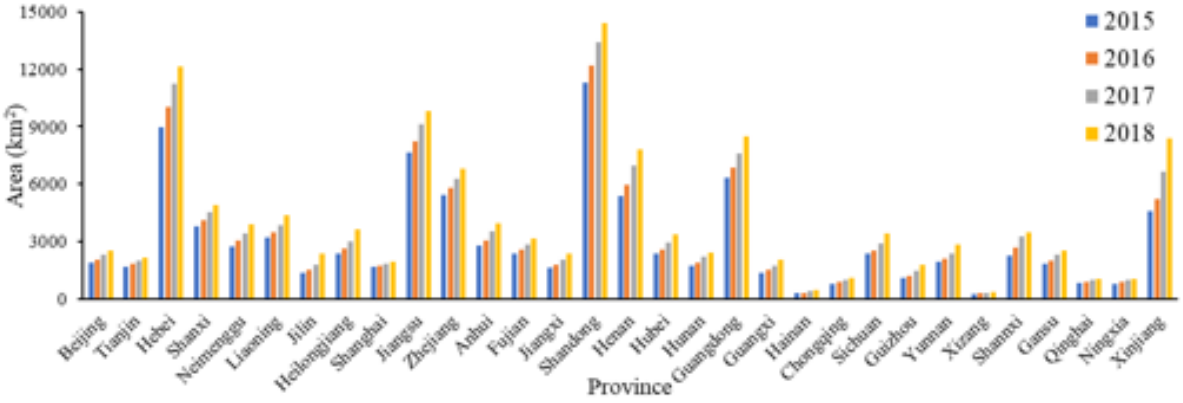
479 *Figure 10.* The AIS extent in China over 4 years. SW: Southwest, NW: Northwest, NE: Northeast, CN:
480 Central North, E: Eastern, C: Central, CS: Central South. The seven subplots demonstrated the
481 representative city of the region.

482

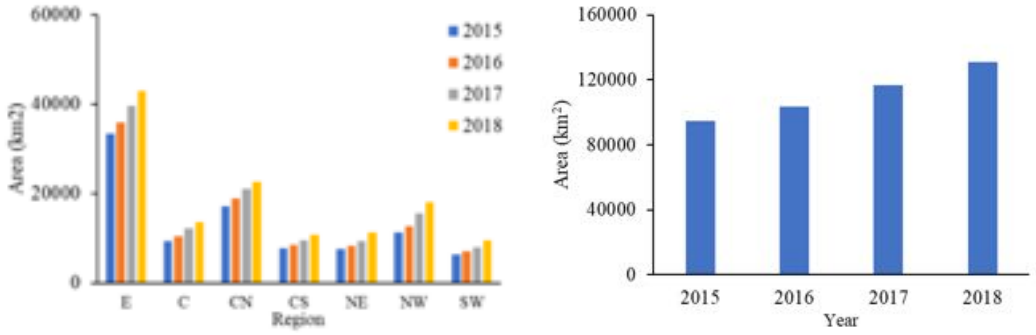
483 *4.3.2 Quantifying the spatial and temporal changes of AIS at different scales*

484 The area change was compared at different levels in Figure 11. The national growth was
485 steady from 2015 to 2018. At the regional level, the AIS area in Eastern region was the much
486 larger than the other regions, followed by Central North region. At the provincial level, the
487 AIS area was the largest in Shandong, followed by Hebei, Jiangsu, which corresponded to the
488 Central North and Eastern regional result respectively.

489



(a)



(b) (c)

Figure 11. The AIS area at different scales from 2015 to 2018: (a) provincial scale, (b) regional scale, (c) national scale.

5. Discussion

The results presented in the previous section are based on some assumptions and make explicit a few important aspects, which we would like to summarize in the following paragraphs.

5.1 Land cover sampling strategy

In this work, impervious surfaces are extracted by mapping AIS and NIS, that is, the binary class mapping. This choice is due to the fact that with such a wide North-South and East-West extent, China presents very diverse land use types. The selection of the most appropriate legend to represent these types is a huge task. Things are simpler if land cover, as opposed to land use, classes is considered. For instance, “farm lands” is a very important land use class but, due to regional and phenological effects, may correspond to very different land cover

classes. Therefore (and we acknowledge this is a simplification) farms with inundated fields, farms with bare fields, and farms with growing weeds to NIS class. As we mentioned before, NIS includes the VEG, WAT and other class, the diversity of NIS is much larger than AIS. Not only the AIS and NIS diversity, but also their proportion will affect the final classification result.

The accuracy assessment is closely related the sampling design. Cluster sampling was implemented considering plots of prefecture-level city and simple random samples within each plot. The samples were chosen by visual inspection of Google Earth images. In our cases, 10 percent of samples were used for training while the latter 90 percent were used for testing. We further tested the sample proportion influences on our study as in Figure 12. When the proportion of training samples increased, OA, UA, PA and Kappa improved correspondingly. But the 50-50 training-testing proportion only improved 1% compared with the 10-50 one. On one hand, the accuracy performance indicated the representative and reliability of the samples. On the other hand, it tested out that 10 percent of samples were in line with the sampling size in (Stehman and Foody 2019), which was considered sufficient for large scale mapping.

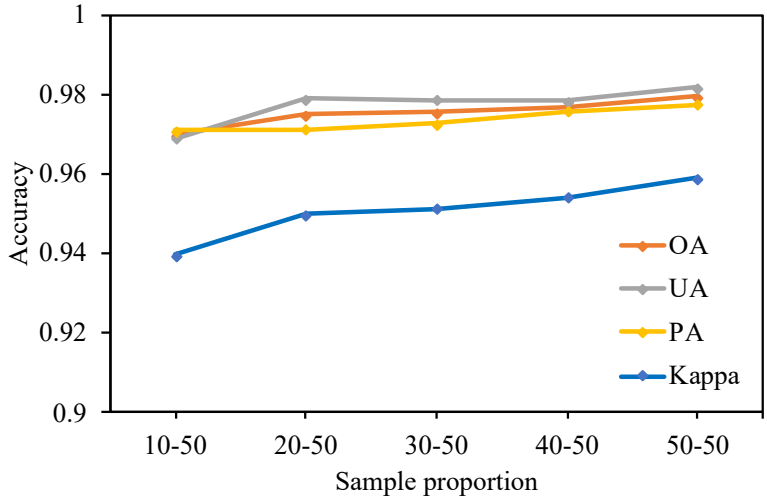


Figure 12. Accuracy according to sample proportion. OA: overall accuracy, UA: user's accuracy, PA: producer's accuracy. Sample proportion: e.g., 10-50, 10% training samples and 50% testing samples.

5.2 Multi-source data contributions

This research stresses that single source data may have limitations for specific land cover classification while multisource data ensures more accurate result. This work, focusing on the wide geographical area covered by P.R. China, and spanning from 2015 to 2018, is an example of solution to this problem. However, this works show that different data provide different type of information and highlight different issues.

For instance, in our study NTL images are more effective as a mask than as a classification feature. Specifically, due to their well-known blooming effect, the extent of "urban areas" in NTL is larger than the actual impervious surface extents. As an advantage, bare lands are not included in NTL areas, and these greatly reduced misclassifications between bare lands and bright impervious surfaces. It must be noted, however, that even if the classification results are improved by masking step, but this mask may miss places where the light is weak.

Instead, when it comes to the SAR data, the use of a feature selection scheme is mandatory. It has been tested out that feature selection improved both the accuracy and efficiency for large scale impervious surface mapping, which is cost-effective on the GEE cloud-based platform. Different Sentinel-1 features have to be included (and compared) in this study, namely VV, as well as the GLCM texture. We found that a direct use of SAR feature may not improve the clustering results. The spatial textures computed starting from SAR texture, deliver instead better classifications, if combined with optical data. Actually, when the physical features from SAR are added, the rough surface and flat surface should be considered. With regard to the rough surface, we should notice that there are similarities between buildings area and mountainous area with similar double bounce scattering characteristics. The slope mask is essential to eliminate the noise from mountainous area. Surface scattering is returned from the relatively flat area, for example, the bare land. For the flat dessert and the rough building area, SAR helps to discriminate the bright impervious surface from dessert, which are similar in spectral characteristics but different in SAR backscattering feature. But there are some flat impervious surfaces that are hard to distinguish, for example, the airport. The track of the airport is flat, showing low intensities in SAR images, which is easily to misclassify as bare land. As a result, it is the limitation for the classification using SAR datasets. Besides, since we combined the optical and SAR feature at feature level in the classification method, the

result is sensitive to the training data. In this study, choosing the representative training samples is of paramount importance. As for the shadow, layover, and speckle noise of SAR dataset, we used the time series mean value of VV backscattering coefficients to reduce the effect. Still, there are some noises from the data source, but it's acceptable for national level impervious surface mapping.

In this study, multisource and single source results were compared quantitatively to demonstrate the improvement of multisource data. The proposed method provided averaged 2% increased OA, 3% increased UA, 1% increased PA, 4% increased Kappa comparing to using optical data only, and corresponding 7%, 6%, 8%, 14% increased accuracy comparing to using SAR data only. Paired t-test was conducted to test the significance with p value less than 0.001, which confirmed that the improvement was statistically significant. Apparently, the improvement of the proposed method was more significant in SAR data than in optical data. Spectral confusions of optical data and backscattering confusions of SAR data in different land cover types exit, which is the limitation of using single dataset for impervious surfaces mapping. In order to quantify the SAR contribution, we further do a multiclass classification using the proposed method. As in the confusion matrix of Table 4, taking the AIS and other (farm and bare land) class as an example, 426 and 446 AIS were misclassified as other class in optical and SAR data respectively, and the confusion reduced to 232 in multisource data group. In optical data, major confusion was between AIS and other class while in SAR, it turned out to be AIS, VEG and other class. WAT was less likely to confuse in all three data groups. It demonstrated that proposed multisource method reduced the confusions between AIS, vegetation, bare land and farmland for large scale impervious surface mapping. Previous studies showed that there were a range of 1%-8% OA improvements (1-2% in (Zhang et al. 2014), 3% in (Shao et al. 2016), 4% in (Zhou et al. 2018) and 8% in (Guo et al. 2014)) comparing to using optical data alone. The variations came from the sample quality, feature selection, fusion strategy and research area. Though the improvement of the accuracy in optical data set in our study was 2% on average, it explained improvement of large scale land cover classification with the combined SAR features. The improvements of SAR data was closely related to the sampling strategy, methodology, accuracy assessment process.

591 Table 4. Confusion matrix of multisource (MS), optical (O) and SAR (S) result.

	Ground Truth				
Class	**MS**	AIS	VEG	WAT	OTHER
	AIS	7107	11	5	120
	VEG	38	1254	11	313
	WAT	32	29	1383	111
	OTHER	232	224	27	3180
Class	**O**	AIS	VEG	WAT	OTHER
	AIS	7028	12	18	180
	VEG	17	1198	22	379
	WAT	98	55	1314	88
	OTHER	426	238	25	2974
Class	**S**	AIS	VEG	WAT	OTHER
	AIS	6756	117	47	323
	VEG	696	269	17	634
	WAT	80	7	1348	120
	OTHER	446	276	157	2784

592

593 It was the first attempt to combine optical and SAR for large scale impervious surface
594 mapping, in which Landsat-8 and Sentinel-1 data were used in this study. The 30 m
595 impervious surface result was comparable and consistent to previous Landsat 30 m
596 classification results from 1986 to 2018, including GlobeLand30, GHSL, Liu et al.'s result
597 (Liu et al. 2018) and Gong et al.'s result (Gong et al. 2020). The emerging Sentinel-2 was
598 available after June 23, 2015, and full coverage of China within a year was from 2016, which
599 was a great supplement for large scale impervious surface mapping. Some issues relate to
600 impervious surface mapping should be considered when using Sentinel-2, like the various
601 resolutions (10 m, 20 m, 60 m), the shadows (Labib and Harris 2018). The potential for
602 Sentinel-1 and Sentinel-2 combination should be further explored in future study.
603
604 5.3 Temporal filtering and potential underestimation in early years
605 One of the key points in our research is the need to apply some sort of temporal filters. This
606 step is essential for the generation of reliable long time series of impervious surface maps
607 from one-year classification (Li et al. 2015; Zhang et al. 2017). Temporal filters should

consider the effect of different sensors and, for seasonal analysis, the phenological effects. Additionally, in this work it is assumed that the urbanization process is irreversible, which may be not the case in reality. In addition, the change of impervious surface is related to the abrupt change. We focused on the intra-year land cover difference and simplified the inter-year change, which also brought uncertainties in this study. The timing and magnitude of impervious surface change should be further determined.

6. Conclusions

In this study, we generated the annual high-resolution AIS at the national scale using multisource data, including Landsat series, Sentinel-1, NTL and SRTM DEM data. In the case of China, complicated land cover exists, and the use of optical data for classification can easily result in the misclassification of AIS. The PAIS from NTL and SRTM DEM data reduced part of the bare land and terrain influence. With the assistance of SAR data, the confusion between bare land and AIS can be reduced comparing to using single optical data only. Also, the proposed temporal filtering helps to ensure the reliability of the results, by eliminating a large part of the inconsistencies due to for the phenological effects and the peculiarities of the different sensors reasons used in this work. The proposed SAR and optical feature-level based fusion framework serves as a pilot study for large scale AIS mapping, and is applicable to extend to multisource data applications.

The reliability of the developed AIS has been validated using different data sources, namely high-resolution samples from Google Earth time series images, census data and other, independently acquired, benchmark datasets. The 30m AIS product from 2015 to 2018 can serve as the basis for future spatial analysis, thermal environment evaluation and population estimation. As a first attempt at displaying the potential applications of this method, we analyzed the changing patterns of China's AIS and compared its temporal patterns at the national, regional and provincial levels. This study provides a general glimpse to the complexity and diversity of the changes to the impervious surface area extent at different geographical scales.

References

Ban, Y.F., Jacob, A., & Gamba, P. (2015). Spaceborne SAR data for global urban mapping at 30 m resolution using a robust urban extractor. *Isprs Journal of Photogrammetry and Remote Sensing, 103*, 28-37

Brabec, E., Schulte, S., & Richards, P.L. (2002). Impervious surfaces and water quality: A review of current literature and its implications for watershed planning. *Journal of Planning Literature, 16*, 499-514

Branstetter, L., & Lardy, N. (2006). China's embrace of globalization. *National Bureau of Economic Research*

Breiman, L. (2001). Random forests. *Machine Learning, 45*, 5-32

Cao, S.S., Hu, D.Y., Zhao, W.J., Mo, Y., & Chen, S.S. (2017). Monitoring spatial patterns and changes of ecology, production, and living land in Chinese urban agglomerations: 35 Years after reform and opening up, where, how and why? *Sustainability, 9*

Conners, R.W., Trivedi, M.M., & Harlow, C.A. (1984). Segmentation of a High-Resolution Urban Scene Using Texture Operators. *Computer Vision Graphics and Image Processing, 25*, 273-310

Deng, C.B., Li, C.J., Zhu, Z., Lin, W.Y., & Xi, L. (2017). Subpixel urban impervious surface mapping: the impact of input Landsat images. *Isprs Journal of Photogrammetry and Remote Sensing, 133*, 89-103

Deng, C.B., & Wu, C.S. (2013). The use of single-date MODIS imagery for estimating large-scale urban impervious surface fraction with spectral mixture analysis and machine learning techniques. *Isprs Journal of Photogrammetry and Remote Sensing, 86*, 100-110

Elvidge, C.D., Tuttle, B.T., Sutton, P.S., Baugh, K.E., Howard, A.T., Milesi, C., Bhaduri, B.L., & Nemani, R. (2007). Global distribution and density of constructed impervious surfaces. *Sensors, 7*, 1962-1979

Friedl, M.A., McIver, D.K., Hodges, J.C.F., Zhang, X.Y., Muchoney, D., Strahler, A.H., Woodcock, C.E., Gopal, S., Schneider, A., Cooper, A., Baccini, A., Gao, F., & Schaaf, C. (2002). Global land cover mapping from MODIS: algorithms and early results. *Remote Sensing of Environment, 83*, 287-302

Gamba, P., & Herold, M. (2009). *Global mapping of human settlement: Experiences, datasets, and prospects*. CRC Press

Gao, F., De Colstoun, E.B., Ma, R.H., Weng, Q.H., Masek, J.G., Chen, J., Pan, Y.Z., & Song, C.H. (2012). Mapping impervious surface expansion using medium-resolution satellite image time series: A case study in the Yangtze River Delta, China. *International Journal of Remote Sensing, 33*, 7609-7628

Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely randomized trees. *Machine Learning, 63*, 3-42

Goldblatt, R., Stuhlmacher, M.F., Tellman, B., Clinton, N., Hanson, G., Georgescu, M., Wang, C.Y., Serrano-Candela, F., Khandelwal, A.K., Cheng, W.H., & Balling, R.C. (2018). Using Landsat and nighttime lights for supervised pixel-based image classification of urban land cover. *Remote Sensing of Environment, 205*, 253-275

Gong, P., Li, X.C., Wang, J., Bai, Y.Q., Cheng, B., Hu, T.Y., Liu, X.P., Xu, B., Yang, J., Zhang, W., & Zhou, Y.Y. (2020). Annual maps of global artificial impervious area (GAIA) between 1985 and 2018. *Remote Sensing of Environment, 236*

Gong, P., Li, X.C., & Zhang, W. (2019). 40-Year (1978-2017) human settlement changes in China reflected by impervious surfaces from satellite remote sensing. *Science Bulletin, 64*, 756-763

685 Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017).
686 Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of*
687 *Environment, 202*, 18-27
688 Gu, C., Wu, L., & Cook, I. (2012). Progress in research on Chinese urbanization. *Frontiers of*
689 *Architectural Research, 1*, 101-149
690 Guo, H.D., Yang, H.N., Sun, Z.C., Li, X.W., & Wang, C.Z. (2014). Synergistic use of optical
691 and polSAR imagery for urban impervious surface estimation. *Photogrammetric Engineering*
692 *and Remote Sensing, 80*, 91-102
693 Guo, W., Li, G.Y., Ni, W.J., Zhang, Y.H., & Lu, D.S. (2018). Exploring improvement of
694 impervious surface estimation at national scale through integration of nighttime light and
695 Proba-V data. *Giscience & Remote Sensing, 55*, 699-717
696 Guyon, I., Weston, J., Barnhill, S., & Vapnik, V. (2002). Gene selection for cancer
697 classification using support vector machines. *Machine Learning, 46*, 389-422
698 Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A.,
699 Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L.,
700 Justice, C.O., & Townshend, J.R. (2013). High-resolution global maps of 21st-century forest
701 cover change. *Science, 342*, 850-853
702 Haralick, R.M., Shanmugam, K., & Dinstein, I. (1973). Textural Features for Image
703 Classification. *Ieee Transactions on Systems Man and Cybernetics, Smc3*, 610-621
704 Herman, T. (1966). A geography of China. *Geographical Journal, 132*, 101-&
705 Homer, C., Dewitz, J., Yang, L.M., Jin, S., Danielson, P., Xian, G., Coulston, J., Herold, N.,
706 Wickham, J., & Megown, K. (2015). Completion of the 2011 National Land Cover Database
707 for the Conterminous United States - Representing a Decade of Land Cover Change
708 Information. *Photogrammetric Engineering and Remote Sensing, 81*, 345-354
709 Imhoff, M.L., Lawrence, W.T., Stutzer, D.C., & Elvidge, C.D. (1997). A technique for using
710 composite DMSP/OLS "city lights" satellite data to map urban area. *Remote Sensing of*
711 *Environment, 61*, 361-370
712 Jiang, L.M., Liao, M.S., Lin, H., & Yang, L.M. (2009). Synergistic use of optical and InSAR
713 data for urban impervious surface mapping: A case study in Hong Kong. *International*
714 *Journal of Remote Sensing, 30*, 2781-2796
715 Joshi, N., Baumann, M., Ehammer, A., Fensholt, R., Grogan, K., Hostert, P., Jepsen, M.R.,
716 Kuemmerle, T., Meyfroidt, P., Mitchard, E.T.A., Reiche, J., Ryan, C.M., & Waske, B. (2016).
717 A Review of the Application of Optical and Radar Remote Sensing Data Fusion to Land Use
718 Mapping and Monitoring. *Remote Sensing, 8*
719 Ju, J.C., & Roy, D.P. (2008). The availability of cloud-free Landsat ETM plus data over the
720 conterminous United States and globally. *Remote Sensing of Environment, 112*, 1196-1211
721 Jun, C., Ban, Y.F., & Li, S.N. (2014). Open access to Earth land-cover map. *Nature, 514*,
722 434-434
723 Labib, S.M., & Harris, A. (2018). The potentials of Sentinel-2 and LandSat-8 data in green
724 infrastructure extraction, using object based image analysis (OBIA) method. *European*
725 *Journal of Remote Sensing, 51*, 231-240
726 Li, X.C., Gong, P., & Liang, L. (2015). A 30-year (1984-2013) record of annual urban
727 dynamics of Beijing City derived from Landsat data. *Remote Sensing of Environment, 166*,
728 78-90
729 Liu, C., Shao, Z., Chen, M., & Luo, H. (2014). MNDISI: A multi-source composition index
730 for impervious surface area estimation at the individual city scale. *Remote Sensing Letters, 5*,
731 204-204

Liu, X.P., Hu, G.H., Ai, B., Li, X., & Shi, Q. (2015). A normalized urban areas composite index (NUACI) based on combination of DMSP-OLS and MODIS for mapping impervious surface area. *Remote Sensing, 7*, 17168-17189

Liu, X.P., Hu, G.H., Chen, Y.M., Li, X., Xu, X.C., Li, S.Y., Pei, F.S., & Wang, S.J. (2018). High-resolution multi-temporal mapping of global urban land using Landsat images based on the Google Earth Engine Platform. *Remote Sensing of Environment, 209*, 227-239

Marceau, D.J., Howarth, P.J., Dubois, J.M.M., & Gratton, D.J. (1990). Evaluation of the Grey-Level Co-Occurrence Matrix Method for Land-Cover Classification Using Spot Imagery. *Ieee Transactions on Geoscience and Remote Sensing, 28*, 513-519

McDonald, R.I., Weber, K., Padowski, J., Florke, M., Schneider, C., Green, P.A., Gleeson, T., Eckman, S., Lehner, B., Balk, D., Boucher, T., Grill, G., & Montgomery, M. (2014). Water on an urban planet: Urbanization and the reach of urban water infrastructure. *Global Environmental Change-Human and Policy Dimensions, 27*, 96-105

Nghiem, S.V., Balk, D., Rodriguez, E., Neumann, G., Sorichetta, A., Small, C., & Elvidge, C.D. (2009). Observations of urban and suburban environments with global satellite scatterometer data. *Isprs Journal of Photogrammetry and Remote Sensing, 64*, 367-380

Puissant, A., Hirsch, J., & Weber, C. (2005). The utility of texture analysis to improve per-pixel classification for high to very high spatial resolution imagery. *International Journal of Remote Sensing, 26*, 733-745

Quin, G., Pinel-Puyssegur, B., Nicolas, J.M., & Loreaux, P. (2014). MIMOSA: An Automatic Change Detection Method for SAR Time Series. *Ieee Transactions on Geoscience and Remote Sensing, 52*, 5349-5363

Roy, D.P., Ju, J.C., Kline, K., Scaramuzza, P.L., Kovalskyy, V., Hansen, M., Loveland, T.R., Vermote, E., & Zhang, C.S. (2010). Web-enabled Landsat Data (WELD): Landsat ETM plus composited mosaics of the conterminous United States. *Remote Sensing of Environment, 114*, 35-49

Schneider, A., Friedl, M.A., & Potere, D. (2009). A new map of global urban extent from MODIS satellite data. *Environmental Research Letters, 4*

Schneider, A., Friedl, M.A., & Potere, D. (2010). Mapping global urban areas using MODIS 500-m data: New methods and datasets based on 'urban ecoregions'. *Remote Sensing of Environment, 114*, 1733-1746

Shao, Z.F., Fu, H.Y., Fu, P., & Yin, L. (2016). Mapping urban impervious surface by fusing optical and SAR data at the decision level. *Remote Sensing, 8*

Shi, K.F., Huang, C., Yu, B.L., Yin, B., Huang, Y.X., & Wu, J.P. (2014). Evaluation of NPP-VIIRS night-time light composite data for extracting built-up urban areas. *Remote Sensing Letters, 5*, 358-366

Stehman, S.V. (1997). Selecting and interpreting measures of thematic classification accuracy. *Remote Sensing of Environment, 62*, 77-89

Stehman, S.V., & Foody, G.M. (2019). Key issues in rigorous accuracy assessment of land cover products. *Remote Sensing of Environment, 231*

Sun, Z.C., Xu, R., Du, W.J., Wang, L., & Lu, D.S. (2019). High-Resolution Urban Land Mapping in China from Sentinel 1A/2 Imagery Based on Google Earth Engine. *Remote Sensing, 11*

Weng, Q., Quattrochi, D., & Gamba, P.E. (2018). *Urban remote sensing*. CRC press

Weng, Q.H. (2012). Remote sensing of impervious surfaces in the urban areas: Requirements, methods, and trends. *Remote Sensing of Environment, 117*, 34-49

778 Xian, G., & Homer, C. (2010). Updating the 2001 national land cover database impervious
779 surface products to 2006 using Landsat imagery change detection methods. *Remote Sensing of*
780 *Environment, 114*, 1676-1686
781 Xie, Y.H., & Weng, Q.H. (2017). Spatiotemporally enhancing time-series DMSP/OLS
782 nighttime light imagery for assessing large-scale urban dynamics. *Isprs Journal of*
783 *Photogrammetry and Remote Sensing, 128*, 1-15
784 Xu, H. (2008). A new index for delineating built-up land features in satellite imagery.
785 *International Journal of Remote Sensing, 29*, 4269-4276
786 Xu, H.Q. (2010). Analysis of impervious surface and its impact on urban heat environment
787 using the normalized difference impervious surface index (NDISI). *Photogrammetric*
788 *Engineering and Remote Sensing, 76*, 557-565
789 Yuan, F., & Bauer, M.E. (2007). Comparison of impervious surface area and normalized
790 difference vegetation index as indicators of surface urban heat island effects in Landsat
791 imagery. *Remote Sensing of Environment, 106*, 375-386
792 Zhang, C., Li, W., & Travis, D. (2007). Gaps-fill of SLC-off Landsat ETM plus satellite
793 image using a geostatistical approach. *International Journal of Remote Sensing, 28*, 5103-
794 5122
795 Zhang, H.S., Wang, T., Zhang, Y.H., Dai, Y.R., Jia, J.J., Yu, C., Li, G., Lin, Y.Y., Lin, H., &
796 Cao, Y. (2018). Quantifying Short-Term Urban Land Cover Change with Time Series
797 Landsat Data: A Comparison of Four Different Cities. *Sensors, 18*
798 Zhang, H.S., & Xu, R. (2018). Exploring the optimal integration levels between SAR and
799 optical data for better urban land cover mapping in the Pearl River Delta. *International*
800 *Journal of Applied Earth Observation and Geoinformation, 64*, 87-95
801 Zhang, L., Weng, Q.H., & Shao, Z.F. (2017). An evaluation of monthly impervious surface
802 dynamics by fusing Landsat and MODIS time series in the Pearl River Delta, China, from
803 2000 to 2015. *Remote Sensing of Environment, 201*, 99-114
804 Zhang, Y.Z., Zhang, H.S., & Lin, H. (2014). Improving the impervious surface estimation
805 with combined use of optical and SAR remote sensing images. *Remote Sensing of*
806 *Environment, 141*, 155-167
807 Zhou, T., Zhao, M.F., Sun, C.L., & Pan, J.J. (2018). Exploring the Impact of Seasonality on
808 Urban Land-Cover Mapping Using Multi-Season Sentinel-1A and GF-1 WFV Images in a
809 Subtropical Monsoon-Climate Region. *Isprs International Journal of Geo-Information, 7*
810 Zhou, Y., Dong, J.W., Liu, J.Y., Metternicht, G., Shen, W., You, N.S., Zhao, G.S., & Xiao,
811 X.M. (2019). Are There Sufficient Landsat Observations for Retrospective and Continuous
812 Monitoring of Land Cover Changes in China? *Remote Sensing, 11*

List of Figure Captions

Figure 11. The AIS area at different scales from 2015 to 2018: (a) provincial scale, (b) regional scale, (c) national scale.

Figure 12. Accuracy according to sample proportion. OA: overall accuracy, UA: user's accuracy, PA: producer's accuracy. Sample proportion: e.g., 10-50, 10% training samples and 50% testing samples.