

Supplementary Material for

Uncovering the nature of urban land use composition using multi-source open big data with ensemble learning

Ying Tu¹, Bin Chen², Wei Lang^{3,4}, Tingting Chen^{3,4}, Miao Li¹, Tao Zhang¹, and Bing Xu^{1,*}

¹Department of Earth System Science, Ministry of Education Key Laboratory for Earth System Modeling, Institute for Global Change Studies, Tsinghua University, Beijing, 100084, China

²Division of Landscape Architecture, Faculty of Architecture, The University of Hong Kong, Hong Kong SAR, China

³Department of Urban and Regional Planning, School of Geography and Planning, Sun Yat-sen University, Guangzhou, 510275, China

⁴China Regional Coordinated Development and Rural Construction Institute, Sun Yat-sen University, Guangzhou, 510275, China

*Corresponding author: Bing Xu (bingxu@tsinghua.edu.cn)

Table of Contents

Section S1. Mapping units generation.

Section S2. Base models and parameter tuning in ensemble learning.

Section S3. Quantifying influencing factors of land use mix.

Section S4. Data preparation and processing.

Tables S1-8.

Figures S1-4.

Section S1. Mapping units generation.

For parcels, we first used the OpenStreetMap (OSM) data to generate road buffers. Initiated in 2004 as a volunteer effort, OSM is a substantial global spatial database that maps a variety of point, line, and polygon features. The positional accuracy of mapped features (± 20 m) is mainly determined by the positioning technologies (e.g., GPS) employed and references used while digitizing these features. It has been reported that road features from OSM largely surpassed the accuracy of other publicly available global datasets such as Global Roads Open Access Data Set (± 500 m) (Haklay 2010), and the high precision and wide coverage make OSM the best available seamless dataset (Barrington-Leigh and Millard-Ball 2017; Meijer et al. 2018). We downloaded all the road data from <https://www.openstreetmap.org/>. Similar to Kennedy et al. (2019), we grouped road features of OSM as either “highway”, “motorway”, “trunk”, “primary”, or “secondary” into a single layer of major roads, grouped road features coded as “tertiary”, “unclassified”, and “residential” into another layer of minor roads, and considered features coded as “tracks” as track roads. We then created buffers according to the given width of different roads, which was estimated based on a nationwide sampling experiment with 2669 samples for major roads and 2358 samples for minor roads across 322 prefecture-level cities in China (Gong et al. 2020). After that, we overlapped the road buffers with the 30-m resolution global urban boundaries (GUB) data in 2018 (Li et al. 2020) to extract initial urban parcels. Different from the commonly used administrative boundaries, the urban boundaries we used mark a physical region that consists of not only the built-up areas but also the associated natural lands in the urban center such as greenspace and water bodies (Li et al. 2020). Furthermore, to

mitigate the influence from water bodies, we excluded water pixels in these parcels using the 10-m resolution FROM-GLC10 land cover product (Gong et al. 2019). Both the GUB and FROM-GLC10 datasets were accessed from <http://data.ess.tsinghua.edu.cn/>. More detailed information about parcel generation can be referred to Gong et al. (2020).

For objects, we adopted an object-based segmentation approach, the simple non-iterative clustering (SNIC) algorithm (Achanta and Susstrunk 2017), to segment the multispectral Sentinel-2 imagery (see Section S4 for an introduction of this data) into homogeneous units. The SNIC is a bottom-up, seed-based segmentation approach that groups neighboring pixels together into clusters based on input data and parameters such as compactness, connectivity, and neighborhood size (Achanta and Susstrunk 2017). It has been widely used in various studies including land use and land cover classification (Ghorbanian et al. 2020; Tassi and Vizzari 2020; Tu et al. 2020a), cropland mapping (Brinkhoff et al. 2020; Paludo et al. 2020; Yang et al. 2021), wildfire monitoring (Crowley et al. 2019), and wetland inventory (Amani et al. 2019). The implementation of SNIC generated a set of homogeneous objects.

Specifically, road widths for buffering major roads and minor roads were set to 36 m and 8 m, respectively. Parameters for SNIC segmentation were designated as follows: compactness was set to 1 to achieve larger clusters, connectivity was set to 8, the neighborhood size was set to 8 to avoid tile boundary artifacts, and the seeds were created in a square pattern using a super-pixel seed spacing of 16 pixels. We intersected all segmented objects with the derived parcels to make sure that each object spatially corresponded to a parcel. As a summary, we generated a total of 5824 parcels and 58,311 objects in our study area.

Section S2. Base models and parameter tuning in ensemble learning.

During ensemble learning, five base models including Random Forest, Extremely Randomized Trees, Gradient Boosted Decision Trees (CatBoost), Light Gradient Boosting Machine (LightGBM), and Neural Networks were utilized for object-based urban land use classification in this study. For each base model, we tested its performance under 20 sets of parameter combinations and chose values with the highest overall accuracy as the optimal parameters. Table S4 lists the tuned optimal parameters for all base models. An introduction to each base model and its parameter settings was provided as followed.

Random Forest is a classic machine learning algorithm that consists 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 (Ho 1995, 1998). 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 algorithms (Breiman 2001). The most important parameter for Random Forest is “n_estimators”, which is the number of decision trees used during the training process. According to Oshiro et al. (2012), as the number of trees grows, it does not always mean the performance of the forest is significantly better than previous forests (fewer trees), and there is a threshold beyond which there is no significant gain. In this study, the parameter “n_estimators” in Random Forest was set between 50 to 300. The parameter “criterion” was set to “gini” and “entropy”, which represented the function to measure the quality of a split.

Extremely Randomized Trees is yielded by adding one further step of randomization to Random Forest. While similar to ordinary random forests in that they are an ensemble of individual trees, there are two main differences: first, each tree is trained using the whole learning sample (rather than a bootstrap sample), and second, the top-down splitting in the tree learner is randomized (Geurts et al. 2006). Instead of computing the locally optimal cut-point for each feature under consideration (e.g., based on the Gini impurity), a random cut-point is selected (Geurts et al. 2006). This value is selected from a uniform distribution within the feature's empirical range (in the tree's training set). The parameter settings of Extremely Randomized Trees were the same as those of Random Forest.

CatBoost is a machine learning algorithm that uses gradient boosting on decision trees (Dorogush et al. 2018). One of the differences between CatBoost and other gradient boosting trees is its advanced processing of the categorical features, which helps to maintain a high level of accuracy (Al Daoud 2019). In CatBoost, the parameter “iterations” indicates the maximum number of trees that can be built when solving machine learning problems. The parameter “learning_rate” is the speed used for reducing the gradient step, which affects the overall time of training. When the number of iterations decreases, the learning rate needs to be increased (Prokhorenkova et al. 2017). In this study, “iterations” was set from 5000 to 50000 and “learning_rate” was set from 0.01 to 0.5.

Similar to CatBoost, LightGBM is a gradient boosting framework developed based on the decision tree algorithm as well (Ke et al. 2017). It implements a highly optimized histogram-based decision tree learning algorithm, which yields great advantages on both efficiency and

memory consumption (Machado et al. 2019). One main parameter in LightGBM is the number of leaves (“num_leaves”) that controls the complexity of the tree mode. Besides, “min_data_in_leaf” is a very important parameter that prevents over-fitting in a leaf-wise tree. Setting it to a large value can avoid growing too deep a tree, but may cause under-fitting. In this study, “num_leaves” was set between 30 and 200, “min_data_in_leaf” was set between 2 and 30, and “learning_rate” was set from 0.01 to 0.5.

Figure S1 shows the architecture of Neural Networks used in this study. It shares similar design choices as the models of Cheng et al. (2016) and Howard and Gugger (2020). Our network applies a separate embedding layer to each categorical feature, where the embedding dimension is selected proportionally to the number of unique levels observed for this feature (Guo and Berkhahn 2016). For multivariate data, the individual embedding layers enable our network to separately learn about each categorical feature before its representation is blended with other variables (Erickson et al. 2020). The embeddings of categorical features are concatenated with the numerical features into a large vector which is both fed into a 3-layer feedforward network as well as directly connected to the output predictions via a linear skip-connection (Figure S1). In Neural Networks, the parameter “num_epoch” is the number of training epochs that controls the training time models and the parameter “batch_size” defines the number of samples that will be propagated through the network. Activate functions were tested with “relu”, “softrelu”, and “tanh”, respectively. More details about the structure and parameters of Neural Networks can be found in Erickson et al. (2020).

Section S3. Quantifying influencing factors of land use mix.

1. Calculating potential variables that influence land use mix

In this study, 19 specific variables from the four aspects of geography, socioeconomy, accessibility, and landscape were included to represent the potential influencing factors of urban land use mix.

1.1. Geography

1.1.1. Mean of elevation (*elevation*)

The Shuttle Radar Topography Mission (SRTM) V3 product is a digital elevation model (DEM) that uses interferometric radar data to obtain near-global elevation information at 1 arcsec resolution (approximately 30 m). The dataset has undergone a void-filling process using open-source data (ASTER GDEM2, GMTED2010, and NED), as opposed to other versions that contain voids or have been void-filled with commercial sources (Farr et al. 2007). We acquired the DEM data covering the study area of Ningbo from the Google Earth Engine (GEE) platform (<https://earthengine.google.com/>) (Gorelick et al. 2017). We then calculated the average elevation value for each parcel.

1.1.2. Mean of NDVI (*ndvi*)

We extracted the average normalized difference vegetation index (NDVI) value for each parcel based on the NDVI band in the processed Sentinel-2 data (see Section S4). NDVI reflects the vegetation coverage of parcels with larger NDVI values indicate more greenspaces.

1.1.3. Fractions of clay, sand, and silt (*fra_clay, fra_sand, fra_silt*)

Soil texture, defined as the combination of mineral particles of different sizes and diameters in the soil, is one of the physical properties of the ground surface. It is regarded as an important basis for soil utilization and management, and is closely related to soil aeration, fertilizer and water retention, and farming rotation. China's soil texture data is compiled based on the 1:1 million (~1000 m) soil type map and soil profile data obtained from the second national soil survey. Soil texture is divided into three types of sand, silt, and clay according to different contents. Each type of data reflects the content of particles of different textures by percentage. We acquired the soil texture data from the Data Center for Resources and Environmental Sciences, Chinese Academy of Sciences (RESDC) (<http://www.resdc.cn>). For each parcel, we calculated fractions of clay, sand, and silt, respectively.

1.2. Socioeconomy

1.2.1. Number of business and commercial points (*business, commercial*)

We counted the total numbers of business and commercial points of parcels using the Baidu POI data (see Section S4). These two variables reflect the vitality of streets.

1.2.2. Mean of population (*pop*)

We calculated the average population for each parcel using the WorldPop data (see Section S4).

1.2.3. Mean of nighttime light (*ntl*)

We calculated the average light intensity for each parcel using the LuoJia-1 NTL data (see Section S4).

1.2.4. Mean of house price (*house_price*)

We obtained the house price data from Lianjia (<https://m.lianjia.com/>), one of the largest real estate brokerage firms in China. Specifically, we first collected an expansive set of house lists using the web-crawling Python program. Second, we cleaned the original data by removing errors and redundancy, and then geocoded physical addresses for each record with Baidu Map API. To ensure the consistency of house tenancy and structure, we further removed properties that were not commodity apartments, such as subsidized housing from work units. In the end, we included a total of 3133 house price records within the study area of Ningbo.

Since these house price records were point data with geographical coordinates, we further interpolated them into a raster surface using the inverse distance weighted (IDW) approach. Specifically, in IDW analysis, the output cell size was set to 100 m, the exponent of distance was set to 2, and the search distance was set to 3000 m. Lastly, we calculated the mean house price for each parcel according to the IDW interpolation result.

1.3. Accessibility

1.3.1. Distances to bus and subway stations (*dis_bus*, *dis_subway*)

We separated bus and subway categories from the obtained Baidu POI data respectively and then calculated the Euclidean distance of each parcel to the nearest bus (subway) station.

1.3.2. Distances to railway and roads (*dis_railway*, *dis_major_road*, *dis_minor_road*, *dis_track_road*)

Using the regrouped OSM data (as processed in Section S1), we calculated the minimum Euclidean distances of each parcel to the railway and to roads of major, minor, and track, respectively.

1.4. Landscape

1.4.1. Shape index of parcel (*shape*)

The shape index (SI) of each parcel was calculated according to the following equation:

$$SI = \frac{0.25P}{\sqrt{A}} \quad (S1)$$

where P and A denotes the perimeter and area of the parcel, respectively. SI represents the deviation between the shape of a parcel and the square of the same area. A larger SI value indicates a more complicated parcel shape.

1.4.2. Richness index of parcel (*richness*)

In Equation (S2), the richness index (RI) was defined as:

$$RI = \frac{C}{A} \quad (S2)$$

where C is the total number of objects within a given parcel and A is the area of the parcel. RI characterizes the richness of parcels with higher RI values indicate richer landscapes.

2. Exploring the association between land use mix and potential variables

One problem in multivariate analysis (MVA) is the collinearity between variables, especially when the number of variables is large. Since MVA requires all the variables to be independent, it is necessary to remove the collinearity in advance. Here we adopt the principal components analysis (PCA) to linearly transform correlated variables into a smaller number of uncorrelated variables, which is achieved by projecting the original data into the reduced PCA space using the eigenvectors of the covariance matrix (Wold et al. 1987). The resulting projected data are essentially linear combinations of the original data capturing most of the variance in the data (Jolliffe and Cadima 2016). Each combination is called a principal component (PC). Normally, the

first PC (PC_1) explains the most variance in the original data with each subsequent component explaining less. PCA is particularly useful in reducing the dimensionality of data and removing the collinearity of variables (Ringnér 2008), and had been utilized in many MVA studies that measure abstract concepts such as urban deprivation or neighborhood inequality (Havard et al. 2008; Lalloué et al. 2013; Langlois and Kitchen 2001; Messer et al. 2006).

In this study, the scikit-learn Python library (Pedregosa et al. 2011) was used to realize PCA. Before transforming, each variable (column) in the original feature matrix X was normalized to 0-1 according to the following equation:

$$X_{norm_{i,j}} = \frac{X_{i,j} - X_{min_j}}{X_{max_j} - X_{min_j}}, (j = 1, 2, \dots, 19) \quad (S3)$$

where $X_{norm_{i,j}}$ is the value in the i -th row and j -th column of the matrix after normalization; $X_{i,j}$ is the i -th row and j -th column value of the original matrix; X_{min_j} is the minimum value of the j -th column and X_{max_j} is the maximum value of the j -th column. During PCA, the number of components was set to 10. To better understand the physical meaning of each PC, we also calculated the importance of each variable reflected by the magnitude of the corresponding values in eigenvectors. The larger they are these absolute values, the more a specific variable contributes to that PC.

Lastly, a multiple linear regression model was performed to analyze the association between land use mix (complexity index (CI)) and the retained PCs according to Equation (S4):

$$CI_i = \beta_0 + \beta_1 PC_{i,1} + \beta_2 PC_{i,2} + \dots + \beta_k PC_{i,k} + \varepsilon_i, (k = 1, 2, \dots, m) \quad (S4)$$

where CI_i is the observed CI value of the i -th parcel, $PC_{i,k}$ is k -th PC value of the i -th parcel, and m is the number of PCs (equals 10 in this case). β_k represents the k -th parameter estimated

using the ordinary least squares method and ε_i is the i -th independent identically distributed normal error. A positive β_k indicates a positive correlation between CI and the k -th PC, and vice versa. All regression analyses were carried out using the statsmodels library (Seabold and Perktold 2010) in Python and three levels of significance were tested at $p<0.001$, $p<0.01$, and $p<0.05$, respectively. Noted all the analysis above was performed for the Level I category.

Section S4. Data preparation and processing.

The Sentinel-1 mission provides ground range detected (GRD) data from a dual-polarization C-band SAR instrument. One advantage of SAR instruments is the capacity in 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. In this study, the Sentinel-1 mean composite for Ningbo, 2018 was processed and acquired from the GEE platform (<https://earthengine.google.com/>) (Gorelick et al. 2017). Specifically, two dual-polarization bands of VV and VH were extracted at a spatial resolution of 10 m.

Sentinel-2 is a wide-swath, high-resolution, multispectral imaging mission with a global 5-day revisit frequency. The data includes 12 spectral bands altogether with spatial resolutions ranging from 10 m to 60 m. In this study, only the four 10 m bands of Blue, Green, Red, and Near Infrared (NIR) in the Level-2A surface reflectance product (Louis et al. 2016) for 2018 were used for further analysis. Specifically, we first filtered the whole-year archive with the percentage of cloudy pixels less than 30% 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 band (QA60) information. Third, we extracted values of normalized difference vegetation index ($NDVI = (NIR-Red)/(NIR+Red)$) and normalized difference water index ($NDWI = (Green-NIR)/(Green+NIR)$) using the retained reflectance in the Green, Red, and NIR bands for each pixel. Finally, we calculated the average pixel values in the image collections and derived the cloud-free Sentinel-2 mean composite for 2018. Like Sentinel-1, we processed and downloaded all the Sentinel-2 data in the GEE platform.

Developed by Wuhan University in China, the new generation of Luojia 1-01 remote sensing satellite was successfully launched on 2 June 2018 (Zhang et al. 2019). Compared with previous NTL data such as the Defense Meteorological Satellite Program's Operational Linescan System, Luojia-1 has many advantages such as finer spatial resolution (~ 130 m) and higher radiometric quantization (14 bits), which significantly enhance its capacity in detecting artificial lightings at night (Li et al. 2019; Tu et al. 2020b). We acquired the Luojia-1 NTL imagery covering the study area for 2018 from the Hubei Data and Application Center (<http://www.hbeos.org.cn/>).

WorldPop is a high-resolution, contemporary, and accurate dataset that provides estimated population distributions at a spatial resolution of ~ 100 m. The data was generated by matching and disaggregating population census to grid cells based on machine learning approaches (Tatem 2017). We accessed the spatial demographics of Ningbo in 2018 from the WorldPop website (<https://www.worldpop.org/>).

The Baidu POI data for 2018 was acquired via the Baidu Map Application Programming Interface (API) (<http://lbsyun.baidu.com/>), one of the most popular web mapping services in China. Each POI record consisted of a series of information including name, location coordinates,

and urban function. We filtered the original POIs by excluding those irrelevant points and reclassified them into 9 general categories including residential, business, commercial, industrial, transportation, administrative, educational, medical, and sport and culture. After cleaning and selecting, a total of 707,297 POI records were obtained within the study area.

Tables S1-8.

Table S1. Texture metrics used in this study.

Texture metric	Formula	Description
Variance	$\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (i - \mu)^2 g(i, j)$	A measure of heterogeneity. Variance increases when the gray level values differ from their mean.
Correlation	$\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \frac{ijg(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y}$	A measure of linear dependency of gray levels on those of neighboring pixels. Correlation will be high if an image contains a considerable amount of linear structure.
Contrast	$\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (i - j)^2 g^2(i, j)$	A measure of exponential difference between the highest and lowest values of neighboring pixels. A high value of contrast indicates the presence of edges, noise, or wrinkled textures in the image.
Dissimilarity	$\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} g(i, j) i - j $	A measure of linear weighted difference between values. It is similar to contrast.
Entropy	$\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} -g(i, j) \ln(g(i, j))$	A measure of the degree of disorder among pixels in the image. Images with a larger number of gray levels have larger entropy.
Angular second moment	$\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} g(i, j)^2$	Measures the uniformity (or orderliness) of the gray level distribution of the image; images with a smaller number of gray levels have larger angular second moment.

Note. All the metrics are calculated based on the gray level co-occurrence matrix (GLCM) (Haralick et al. 1973). In all equations, N is the total number of gray levels in the image; $g(i, j)$ is the (i, j) -th entry of the normalized GLCM, that is, $g(i, j) = p(i, j) / \sum_{ij} p(i, j)$, where $p(i, j)$ is the (i, j) -th entry of the computed GLCM; and μ_x, μ_y and σ_x, σ_y denote the mean and standard deviations of the row and column sums of the GLCM, respectively, that is, $\mu = \sum_{ij} ig(i, j)$ and $\sigma = \sum_{ij} (i - \mu)^2 g(i, j)$.

Table S2. Summary of features used in the stage-1 mapping of EULUC-seg.

Data source	Features	Variables	Count
Sentinel-1	Mean and the standard deviation of VV and VH bands	<i>VV_mean, VH_mean,</i>	2
	Standard deviation of VV and VH bands	<i>VV_std, VH_std</i>	2
	Textures (variance, correlation, contrast, dissimilarity, entropy, and angular second moment) of VV and VH bands	<i>VV_var, VV_corr, VV_contrast, VV_diss, VV_ent, VV_asm, VH_var, VH_corr, VH_contrast, VH_diss, VH_ent, VH_asm</i>	12
Sentinel-2	Mean of blue, green, red, near-infrared bands, NDVI, and NDWI bands	<i>B2_mean, B3_mean, B4_mean, B8_mean, NDVI_mean, NDWI_mean</i>	6
	Standard deviation of blue, green, red, near-infrared bands, NDVI, and NDWI bands	<i>B2_std, B3_std, B4_std, B8_std, NDVI_std, NDWI_std</i>	6
	Textures (variance, correlation, contrast, dissimilarity, entropy, and angular second moment) of blue, green, red, and near-infrared bands	<i>B2_var, B2_corr, B2_contrast, B2_diss, B2_ent, B2_asm, B3_var, B3_corr, B3_contrast, B3_diss, B3_ent, B3_asm, B4_var, B4_corr, B4_contrast, B4_diss, B4_ent, B4_asm, B8_var, B8_corr, B8_contrast, B8_diss, B8_ent, B8_asm</i>	24
Luoja-1	Mean of DN values	<i>Luoja1_mean</i>	1
	Sum of DN values	<i>Luoja1_sum</i>	1
WorldPop	Mean of population	<i>Pop_mean</i>	1
	Sum of population	<i>Pop_sum</i>	1
/	Area of each object	<i>Area</i>	1
Baidu POI	Total number of all POIs	<i>pAll</i>	1
	Total number of each type of POIs	<i>p101, p201, p202, p301, p401, p501, p502, p503, p504</i>	9
	Proportion of each type of POIs	<i>r101, r201, r202, r301, r401, r501, r502, r503, r504</i>	9
Total			76

Table S3. The two-level essential urban land use categories (EULUC) classification system.

Level I	Level II	Descriptions
01 Residential	0101 Residential	Residential areas with a protective gate and detailed architectural standards, including houses, apartment buildings, etc.
	0102 Village	Urban villages not regulated by any form of centralized urban planning, including rural settlements, traditional dwellings, etc.
02 Commercial	0201 Business	Buildings where people work, including office buildings, and commercial office places for finance, internet technology, e-commerce, media, etc.
	0202 Commercial	Houses and buildings for commercial retails, restaurants, lodging, and entertainment.
03 Industrial	0301 Industrial	Land and buildings used for manufacturing, warehouse, mining, etc.
04 Transportation	0401 Transportation	Transportation facilities including airport, motor, bus, train stations and ancillary facilities.
05 Public	0501 Administrative	Lands used for government, military, and public service agencies.
	0502 Educational	Lands used for education and research, including schools, universities, institutes and their ancillary facilities.
	0503 Medical	Lands used for hospitals, disease prevention, and emergency services.
	0504 Sport and cultural	Lands used for public sports and training, cultural services, including gym centers, libraries, museums, exhibition centers, etc.
	0505 Park and greenspace	Parks and greenspace lands used for entertainment and environmental conservation.
	0506 Undeveloped	Raw lands without utilities, structure, or pre-defined building site and internal roads.

Table S4. Tuned optimal parameters for each base model in ensemble learning.

Model	Optimal parameters
Random Forest	n_estimators: 300; criterion: entropy
Extremely Randomized Trees	n_estimators: 300; criterion: entropy
CatBoost	iterations: 10000; learning rate: 0.1
LightGBM	num_leaves: 60; min_data_in_leaf: 21; learning rate: 0.06
Neural Networks	num_epoch: 10; learning_rate: 0.005; activation: relu; batch_size: 256; dropout_probability: 0.035

Table S5. Confusion matrix for the Level I category of EULUC-seg. UA and PA denote user's accuracy and producer's accuracy, respectively. Definition for each land use category can be seen in Table S3.

	01	02	03	04	05	UA (%)	PA (%)
01	43	0	0	0	1	97.73	93.48
02	2	18	1	0	4	72.00	85.71
03	0	0	8	0	2	80.00	88.89
04	0	0	0	5	2	71.43	100.00
05	1	3	0	0	55	93.22	85.94
OA = 85.52%, Kappa coefficient = 0.79							

Table S6. Confusion matrix for the Level II category of EULUC-seg. UA and PA denote user's accuracy and producer's accuracy, respectively. Definition for each land use category can be seen in Table S3.

	0101	0102	0201	0202	0301	0401	0501	0502	0503	0504	0505	0506	UA (%)	PA (%)
0101	24	0	0	0	0	0	1	0	0	0	0	0	96.00	88.89
0102	0	18	0	0	1	0	0	0	0	0	0	0	94.74	90.00
0201	1	0	11	0	0	0	0	1	0	0	0	0	84.62	64.71
0202	0	0	2	5	2	0	0	0	0	2	1	0	41.67	100.00
0301	0	0	0	0	9	0	0	1	0	0	0	0	90.00	60.00
0401	0	1	0	0	0	5	0	1	0	0	0	0	71.43	100.00
0501	1	0	1	0	0	0	4	0	0	0	0	0	66.67	66.67
0502	0	0	2	0	1	0	1	5	0	0	3	0	41.67	50.00
0503	1	0	1	0	0	0	0	1	3	0	0	0	50.00	100.00
0504	0	1	0	0	0	0	0	1	0	3	1	0	50.00	60.00
0505	0	0	0	0	1	0	0	0	0	0	22	0	95.65	78.57
0506	0	0	0	0	1	0	0	0	0	0	1	4	66.67	100.00
OA = 77.93%, Kappa coefficient = 0.75														

Table S7. Urban land use composition in Ningbo, 2018. Noted that the statistics were calculated using the stage-1 mapping results of EULUC-seg. Definition for each land use category can be seen in Table S3.

Level I			Level II		
Category	Area (km²)	Proportion (%)	Category	Area (km²)	Proportion (%)
01	434.83	30.17	0101	293.09	20.34
			0102	141.75	9.83
02	38.61	2.68	0201	16.61	1.15
			0202	22.00	1.53
03	288.71	20.03	0301	288.71	20.03
04	14.22	0.99	0401	14.22	0.99
05	664.90	46.13	0501	10.47	0.73
			0502	34.59	2.40
			0503	2.71	0.19
			0504	3.99	0.28
			0505	597.42	41.45
			0506	15.71	1.09
Total	1441.27	100.00	Total	1441.27	100.00

Table S8. Extracted ten components from PCA and their variable weights. Red colors indicate greater importance of variables while green colors indicate less importance of variables.

Variable	Component									
	PC ₁	PC ₂	PC ₃	PC ₄	PC ₅	PC ₆	PC ₇	PC ₈	PC ₉	PC ₁₀
<i>elevation</i>	0.05	0.05	0.15	0.02	0.07	0.05	0.04	0.10	0.11	0.07
<i>ndvi</i>	0.06	0.07	0.09	0.90	0.15	0.10	0.03	0.35	0.04	0.07
<i>fra_clay</i>	0.44	0.42	0.09	0.05	0.30	0.48	0.19	0.06	0.12	0.03
<i>fra_sand</i>	0.40	0.44	0.08	0.04	0.05	0.00	0.00	0.07	0.05	0.02
<i>fra_silt</i>	0.27	0.40	0.06	0.02	0.30	0.67	0.27	0.07	0.06	0.08
<i>business</i>	0.00	0.00	0.03	0.05	0.05	0.00	0.19	0.06	0.07	0.20
<i>commercial</i>	0.01	0.00	0.02	0.03	0.02	0.00	0.10	0.02	0.04	0.10
<i>pop</i>	0.01	0.02	0.01	0.11	0.11	0.08	0.12	0.21	0.01	0.22
<i>ntl</i>	0.04	0.02	0.02	0.14	0.17	0.12	0.11	0.27	0.01	0.34
<i>house_price</i>	0.18	0.05	0.12	0.16	0.25	0.08	0.15	0.57	0.53	0.03
<i>dis_bus</i>	0.00	0.01	0.02	0.05	0.01	0.07	0.16	0.03	0.01	0.07
<i>dis_subway</i>	0.64	0.52	0.09	0.10	0.00	0.05	0.07	0.01	0.08	0.42
<i>dis_rail</i>	0.30	0.41	0.54	0.03	0.01	0.07	0.04	0.05	0.15	0.54
<i>dis_major_road</i>	0.01	0.01	0.00	0.04	0.17	0.30	0.65	0.21	0.55	0.26
<i>dis_minor_road</i>	0.01	0.01	0.01	0.07	0.00	0.00	0.02	0.02	0.06	0.12
<i>dis_track_road</i>	0.19	0.08	0.80	0.06	0.07	0.05	0.04	0.14	0.16	0.45
<i>area</i>	0.01	0.01	0.01	0.03	0.01	0.00	0.09	0.07	0.04	0.05
<i>shape</i>	0.03	0.04	0.02	0.30	0.53	0.16	0.38	0.56	0.31	0.03
<i>richness</i>	0.05	0.00	0.05	0.12	0.61	0.39	0.42	0.14	0.46	0.03

Figures S1-4.

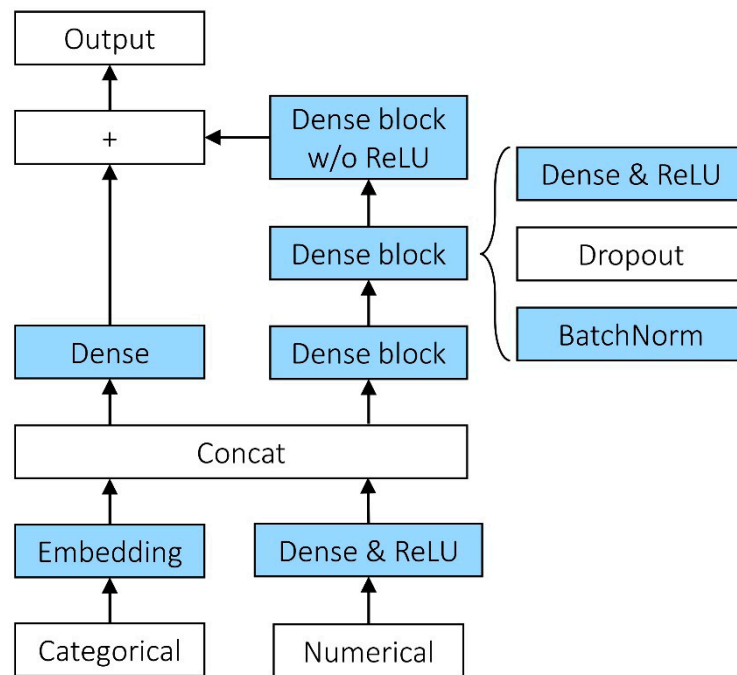


Figure S1. The architecture of Neural Networks in ensemble learning. Layers with learnable parameters are marked as blue. See Erickson et al. (2020) for more information on the structure and parameters of Neural Networks.

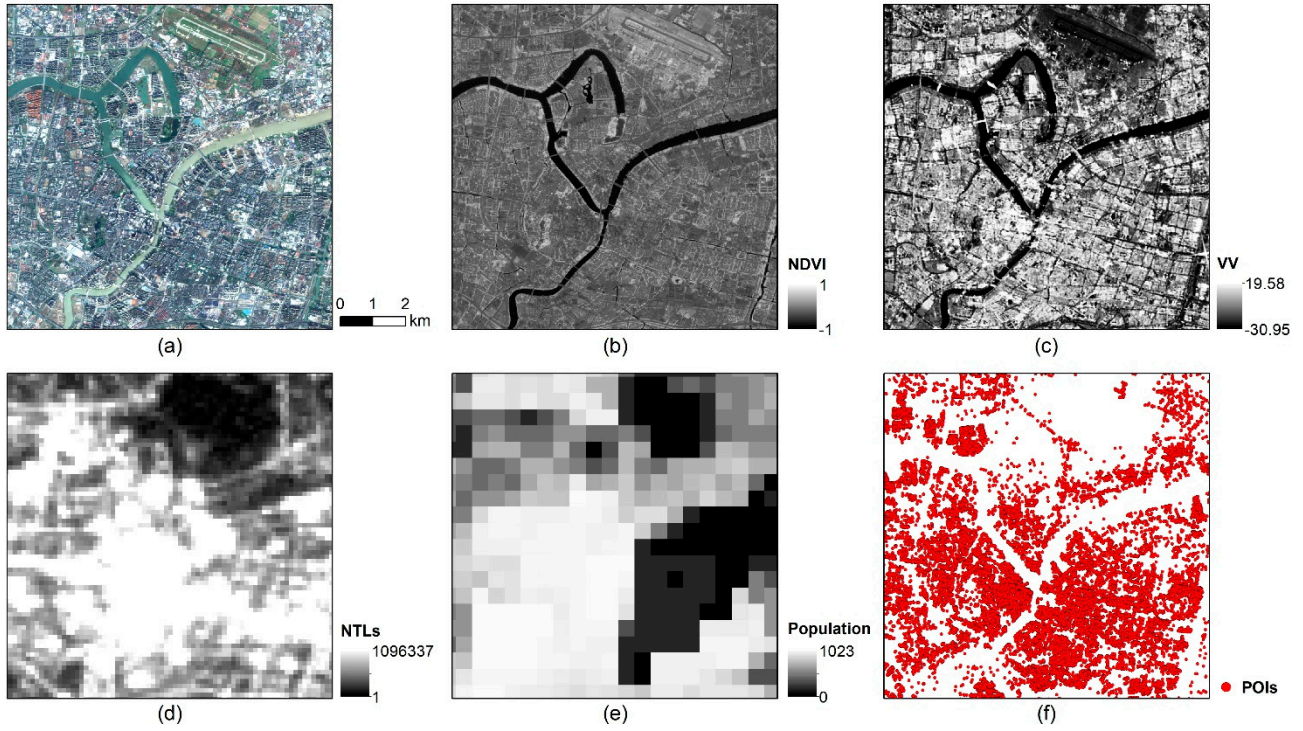


Figure S2. Comparison of multi-source data with different spatial resolutions used for urban land use classification in this study. (a) True color composition of Sentinel-2 multispectral imagery (10 m). (b) Normalized difference vegetation index (NDVI) band of Sentinel-2 (10 m). (c) VV dual-polarization band of Sentinel-1 (10 m). (d) 130-m Luoja-1 nighttime lights (NTLs). (e) 500-m WorldPop population distributions. (f) Baidu points of interest (POIs). The extent in (a-f) was the city center of Ningbo.

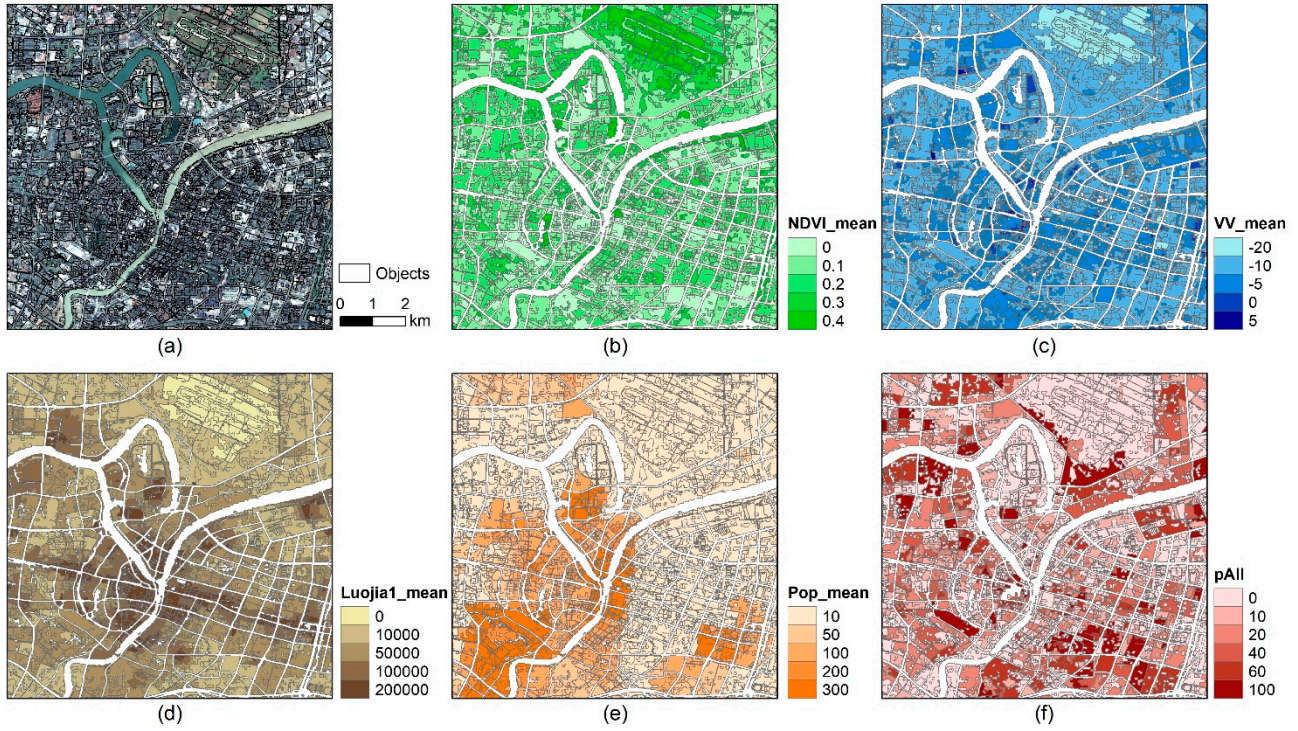


Figure S3. Comparison of object-based features derived from multi-source data in the city center of Ningbo. (a) True color composition of Sentinel-2 multispectral imagery (10 m) overlapped with objects derived from the segmentation results. (b) Mean of normalized difference vegetation index (NDVI) band of Sentinel-2. (c) Mean of VV dual-polarization band of Sentinel-1. (d) Mean of Luojia-1 nighttime lights. (e) Mean of WorldPop populations. (f) Total number of Baidu points of interest. Noted values in (b-f) were all calculated at the object scale (as shown in (a)).

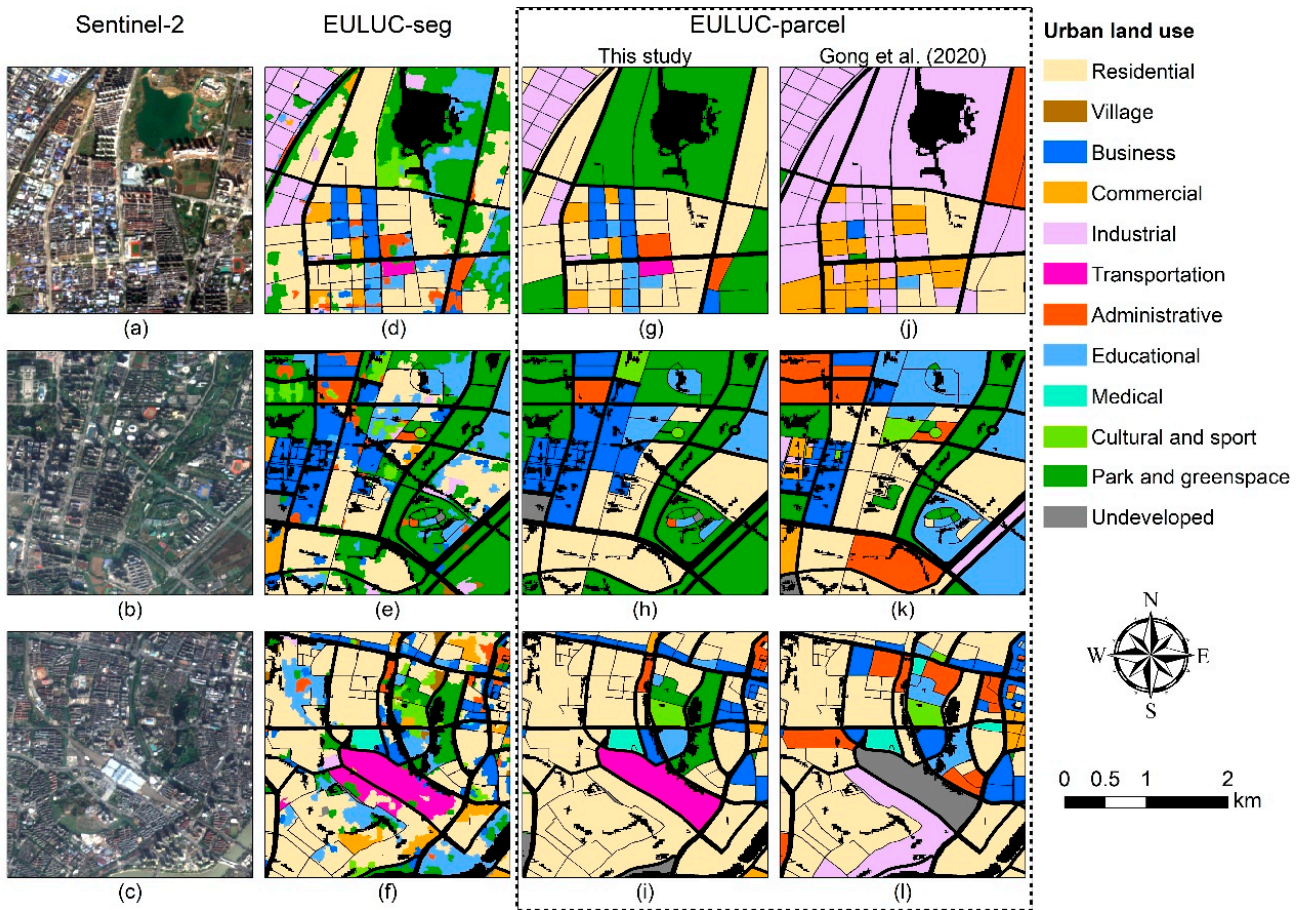


Figure S4. Comparison of remotely sensed images and urban land use mapping results. (a-c) Sentinel-2 multispectral images (true color composition). (d-f) Maps of EULUC-seg in this study. (g-i) Maps of EULUC-parcel in this study. (j-k) Maps of EULUC-parcel produced by Gong et al. (2020).

References

- Achanta, R., & Susstrunk, S. (2017). Superpixels and polygons using simple non-iterative clustering. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4651-4660)
- Al Daoud, E. (2019). Comparison between XGBoost, LightGBM and CatBoost using a home credit dataset. *International Journal of Computer and Information Engineering*, 13, 6-10
- Amani, M., Mahdavi, S., Afshar, M., Brisco, B., Huang, W., Mohammad Javad Mirzadeh, S., White, L., Banks, S., Montgomery, J., & Hopkinson, C. (2019). Canadian wetland inventory using google earth engine: The first map and preliminary results. *Remote Sensing*, 11, 842
- Barrington-Leigh, C., & Millard-Ball, A. (2017). The world's user-generated road map is more than 80% complete. *PLoS ONE*, 12, e0180698
- Breiman, L. (2001). Random forests. *Machine learning*, 45, 5-32
- Brinkhoff, J., Vardanega, J., & Robson, A.J. (2020). Land cover classification of nine perennial crops using sentinel-1 and-2 data. *Remote Sensing*, 12, 96

- Cheng, H.-T., Koc, L., Harmsen, J., Shaked, T., Chandra, T., Aradhye, H., Anderson, G., Corrado, G., Chai, W., & Ispir, M. (2016). Wide & deep learning for recommender systems. In *Proceedings of the 1st Workshop on Deep Learning for Recommender Systems* (pp. 7-10)
- Crowley, M.A., Cardille, J.A., White, J.C., & Wulder, M.A. (2019). Generating intra-year metrics of wildfire progression using multiple open-access satellite data streams. *Remote Sensing of Environment*, 232, 111295
- Dorogush, A.V., Ershov, V., & Gulin, A. (2018). CatBoost: gradient boosting with categorical features support. *arXiv preprint arXiv:1810.11363*
- Erickson, N., Mueller, J., Shirkov, A., Zhang, H., Larroy, P., Li, M., & Smola, A. (2020). Autogluon-tabular: Robust and accurate automl for structured data. *arXiv preprint arXiv:2003.06505*
- Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., & Alsdorf, D. (2007). The Shuttle Radar Topography Mission. *Reviews of Geophysics*, 45
- Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely randomized trees. *Machine learning*, 63, 3-42
- Ghorbanian, A., Kakooei, M., Amani, M., Mahdavi, S., Mohammadzadeh, A., & Hasanlou, M. (2020). Improved land cover map of Iran using Sentinel imagery within Google Earth Engine and a novel automatic workflow for land cover classification using migrated training samples. *ISPRS Journal of Photogrammetry and Remote Sensing*, 167, 276-288
- Gong, P., Chen, B., Li, X., Liu, H., Wang, J., Bai, Y., Chen, J., Chen, X., Fang, L., Feng, S., Feng, Y., Gong, Y., Gu, H., Huang, H., Huang, X., Jiao, H., Kang, Y., Lei, G., Li, A., Li, X., Li, X., Li, Y., Li, Z., Li, Z., Liu, C., Liu, C., Liu, M., Liu, S., Mao, W., Miao, C., Ni, H., Pan, Q., Qi, S., Ren, Z., Shan, Z., Shen, S., Shi, M., Song, Y., Su, M., Ping Suen, H., Sun, B., Sun, F., Sun, J., Sun, L., Sun, W., Tian, T., Tong, X., Tseng, Y., Tu, Y., Wang, H., Wang, L., Wang, X., Wang, Z., Wu, T., Xie, Y., Yang, J., Yang, J., Yuan, M., Yue, W., Zeng, H., Zhang, K., Zhang, N., Zhang, T., Zhang, Y., Zhao, F., Zheng, Y., Zhou, Q., Clinton, N., Zhu, Z., & Xu, B. (2020). Mapping essential urban land use categories in China (EULUC-China): preliminary results for 2018. *Science Bulletin*, 65, 182-187
- Gong, P., Liu, H., Zhang, M., Li, C., Wang, J., Huang, H., Clinton, N., Ji, L., Li, W., Bai, Y., Chen, B., Xu, B., Zhu, Z., Yuan, C., Suen, H.P., Guo, J., Xu, N., Li, W., Zhao, Y., Yang, J., Yu, C., Wang, X., Fu, H., Yu, L., Dronova, I., Hui, F., Cheng, X., Shi, X., Xiao, F., Liu, Q., & Song, L. (2019). 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. *Science Bulletin*, 64, 370-373
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202, 18-27
- Guo, C., & Berkhahn, F. (2016). Entity embeddings of categorical variables. *arXiv preprint arXiv:1604.06737*
- Haklay, M. (2010). How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. *Environment and Planning B: Planning and Design*, 37, 682-703

- Haralick, R.M., Shanmugam, K., & Dinstein, I.H. (1973). Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, 610-621
- Havard, S., Deguen, S., Bodin, J., Louis, K., Laurent, O., & Bard, D. (2008). A small-area index of socioeconomic deprivation to capture health inequalities in France. *Social Science & Medicine*, 67, 2007-2016
- Ho, T.K. (1995). Random decision forests. In, *Proceedings of 3rd International Conference on Document Analysis and Recognition* (pp. 278-282): IEEE
- Ho, T.K. (1998). The random subspace method for constructing decision forests. *IEEE transactions on pattern analysis and machine intelligence*, 20, 832-844
- Howard, J., & Gugger, S. (2020). Fastai: a layered API for deep learning. *Information*, 11, 108
- Jolliffe, I.T., & Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374, 20150202
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, 30, 3146-3154
- Kennedy, C.M., Oakleaf, J.R., Theobald, D.M., Baruch-Mordo, S., & Kiesecker, J. (2019). Managing the middle: A shift in conservation priorities based on the global human modification gradient. *Global Change Biology*
- Lalloué, B., Monnez, J.-M., Padilla, C., Kihal, W., Le Meur, N., Zmirou-Navier, D., & Deguen, S. (2013). A statistical procedure to create a neighborhood socioeconomic index for health inequalities analysis. *International journal for equity in health*, 12, 1-11
- Langlois, A., & Kitchen, P. (2001). Identifying and measuring dimensions of urban deprivation in Montreal: An analysis of the 1996 census data. *Urban Studies*, 38, 119-139
- Li, X., Gong, P., Zhou, Y., Wang, J., Bai, Y., Chen, B., Hu, T., Xiao, Y., Xu, B., Yang, J., Liu, X., Cai, W., Huang, H., Wu, T., Wang, X., Lin, P., Li, X., Chen, J., He, C., Li, X., Yu, L., Clinton, N., & Zhu, Z. (2020). Mapping global urban boundaries from the global artificial impervious area (GAIA) data. *Environmental Research Letters*, 15
- Li, X., Li, X., Li, D., He, X., & Jendryke, M. (2019). A preliminary investigation of LuoJia-1 night-time light imagery. *Remote sensing letters*, 10, 526-535
- Louis, J., Debaecker, V., Pflug, B., Main-Knorn, M., Bieniarz, J., Mueller-Wilm, U., Cadau, E., & Gascon, F. (2016). Sentinel-2 Sen2Cor: L2A processor for users. In, *Proceedings Living Planet Symposium 2016* (pp. 1-8): Spacebooks Online
- Machado, M.R., Karray, S., & de Sousa, I.T. (2019). LightGBM: An effective decision tree gradient boosting method to predict customer loyalty in the finance industry. In, *2019 14th International Conference on Computer Science & Education (ICCSE)* (pp. 1111-1116): IEEE
- Meijer, J.R., Huijbregts, M.A., Schotten, K.C., & Schipper, A.M. (2018). Global patterns of current and future road infrastructure. *Environmental Research Letters*, 13, 064006
- Messer, L.C., Laraia, B.A., Kaufman, J.S., Eyster, J., Holzman, C., Culhane, J., Elo, I., Burke, J.G., & O'campo, P. (2006). The development of a standardized neighborhood deprivation index. *Journal of Urban Health*, 83, 1041-1062

- Oshiro, T.M., Perez, P.S., & Baranauskas, J.A. (2012). How many trees in a random forest? In, *International Workshop on Machine Learning and Data Mining in Pattern Recognition* (pp. 154-168): Springer
- Paludo, A., Becker, W.R., Richetti, J., Silva, L.C.D.A., & Johann, J.A. (2020). Mapping summer soybean and corn with remote sensing on Google Earth Engine cloud computing in Parana state–Brazil. *International Journal of Digital Earth*, 13, 1624-1636
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., & Dubourg, V. (2011). Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, 12, 2825-2830
- Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A.V., & Gulin, A. (2017). CatBoost: unbiased boosting with categorical features. *arXiv preprint arXiv:1706.09516*
- Ringnér, M. (2008). What is principal component analysis? *Nature biotechnology*, 26, 303-304
- Seabold, S., & Perktold, J. (2010). Statsmodels: Econometric and statistical modeling with python. In, *Proceedings of the 9th Python in Science Conference* (p. 61): Austin, TX
- Tassi, A., & Vizzari, M. (2020). Object-Oriented LULC Classification in Google Earth Engine Combining SNIC, GLCM, and Machine Learning Algorithms. *Remote Sensing*, 12, 3776
- Tatem, A.J. (2017). WorldPop, open data for spatial demography. *Scientific data*, 4, 1-4
- Tu, Y., Chen, B., Zhang, T., & Xu, B. (2020a). Regional Mapping of Essential Urban Land Use Categories in China: A Segmentation-Based Approach. *Remote Sensing*, 12
- Tu, Y., Lang, W., Yu, L., Li, Y., Jiang, J., Qin, Y., Wu, J., Chen, T., & Xu, B. (2020b). Improved Mapping Results of 10 m Resolution Land Cover Classification in Guangdong, China Using Multisource Remote Sensing Data With Google Earth Engine. *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, 5384-5397
- Wold, S., Esbensen, K., & Geladi, P. (1987). Principal component analysis. *Chemometrics and intelligent laboratory systems*, 2, 37-52
- Yang, L., Wang, L., Abubakar, G.A., & Huang, J. (2021). High-Resolution Rice Mapping Based on SNIC Segmentation and Multi-Source Remote Sensing Images. *Remote Sensing*, 13, 1148
- Zhang, G., Wang, J., Jiang, Y., Zhou, P., Zhao, Y., & Xu, Y. (2019). On-Orbit Geometric Calibration and Validation of Luojia 1-01 Night-Light Satellite. *Remote Sensing*, 11, 264