remote sensing

Supplementary Materials: Accuracy Improvements to Pixel-Based and Object-Based LULC Classification with Auxiliary Datasets from Google Earth Engine

Le'an Qu, Zhenjie Chen, Manchun Li, Junjun Zhi and Huiming Wang

1. Relationships of land cover types in different land classification systems

Table 1. Relationships of land cover types in different land classification systems.

Land Cover Class	MCD12Q1	GlobCover 2009	Land Survey Data
Cultivated land	12 Croplands 14 Cropland/natural vegetation mosaic	11 Post-flooding or irrigated croplands 14 Rainfed croplands 20 Mosaic cropland (50–70%) / vegetation (grassland, shrubland, forest) (20–50%) 30 Mosaic vegetation (grassland, shrubland, forest) (50–70%)/ cropland (20–50%)	11 irrigated croplands 12 nonirrigated croplands
Forest	1 Evergreen needleleaf forest 2 Evergreen broadleaf forest 3 Deciduous needleleaf forest 4 Deciduous broadleaf forest 5 Mixed forest	40 Closed to open (>15%) broadleaved evergreen and/or semi-deciduous forest (>5m) 50 Closed (>40%) broadleaved deciduous forest (>5m) 60 Open (15–40%) broadleaved deciduous forest (>5m) 70 Closed (>40%) needleleaved evergreen forest (>5m) 90 Open (15–40%) needleleaved deciduous or evergreen forest (>5m)	21 forest 23 open forest 24 other forest
Grassland	8 Woody savannas 9 Savannas 10 Grasslands	110 Mosaic forest-shrubland (50-70%) / grassland (20–50%) 120 Mosaic grassland (50-70%) / forest-shrubland (20–50%) 140 Closed to open (>15%) grassland 150 Sparse (>15%) vegetation (woody vegetation, shrubs, grassland)	31 High coverage grassland (>50%) 32 Medium coverage grassland (20%–50%) 33 Low coverage grassland (5%–20%)

2. Scripts for image pre-processing

https://code.earthengine.google.com/?accept_repo=users/qulean/Landsat8_preprocessing

3. Scripts for training points selection

https://code.earthengine.google.com/?accept_repo=users/qulean/samples

4. Scripts for image segmentation

https://code.earthengine.google.com/?accept_repo=users/qulean/Segmentation

5. Scripts for image classification, accuracy assessment and features' importance

https://code.earthengine.google.com/?accept_repo=users/qulean/Classification

6. Accuracy of pixel-based RF models

Table S2. Accuracy of M1–M7 RF models.

	M1		M2		M3		M4		M5		M6		M7	
	UA	**PA**	**UA**	**PA**	**UA**	**PA**	**UA**	**PA**	**UA**	**PA**	**UA**	**PA**	**UA**	**PA**
Cropland	89.54	96.19	91.60	**96**.02	90.43	97.26	90.81	96.11	91.85	97.61	91.26	96.74	93.12	97.79
Grassland	55.56	34.08	33.33	19.62	27.91	22.22	36.36	15.56	29.17	33.87	42.59	37.70	50.00	40.91
Woodland	95.17	97.79	95.54	97.58	95.45	99.55	95.80	97.88	95.97	99.61	95.44	98.47	96.28	99.60
Water bodies	95.58	89.64	95.86	93.04	96.89	91.22	94.57	96.14	96.02	96.44	95.68	89.89	96.60	95.99
Built-up land	75.32	60.10	78.30	69.53	76.41	59.90	81.56	61.33	84.46	60.53	80.43	69.58	84.11	68.53
kappa	**0.87**		**0.89**		**0.89**		**0.89**		**0.90**		**0.89**		**0.91**	
OA	**91.51**		**92.73**		**92.36**		**92.62**		**93.53**		**92.73**		**94.20**	

Note: PA = producer's accuracy, UA = user's accuracy, OA = overall accuracy.

7. Accuracy of object-based RF models.

Table S3. Accuracy of M1'–M7' RF models.

	M1'		M2'		M3'		M4'		M5'		M6'		M7'	
	UA	**PA**	**UA**	**PA**	**UA**	**PA**	**UA**	**PA**	**UA**	**PA**	**UA**	**PA**	**UA**	**PA**
Cropland	92.70	98.59	94.29	98.28	94.53	99.17	94.59	97.97	94.69	98.55	94.34	98.48	95.48	98.30
Grassland	50.00	37.25	50.00	39.09	72.73	54.41	38.46	29.55	66.66	48.33	35.29	32.79	74.07	53.85
Woodland	95.68	98.91	96.53	98.80	96.66	99.28	95.86	99.28	96.85	99.50	96.54	99.18	96.94	99.34
Water bodies	97.16	91.53	95.92	93.55	98.39	95.61	96.25	94.38	96.73	93.88	97.03	91.48	97.99	96.17
Bulit-up land	86.31	74.86	87.71	78.92	93.82	81.59	90.91	81.84	89.04	78.86	87.98	84.51	89.32	85.95
kappa	**0.91**		**0.92**		**0.94**		**0.92**		**0.93**		**0.92**		**0.94**	
OA	**94.03**		**94.67**		**95.73**		**94.91**		**95.27**		**94.95**		**96.01**	

Note: PA = producer's accuracy, UA = user's accuracy, OA = overall accuracy.

8. Classification results of pixel-based RF models.

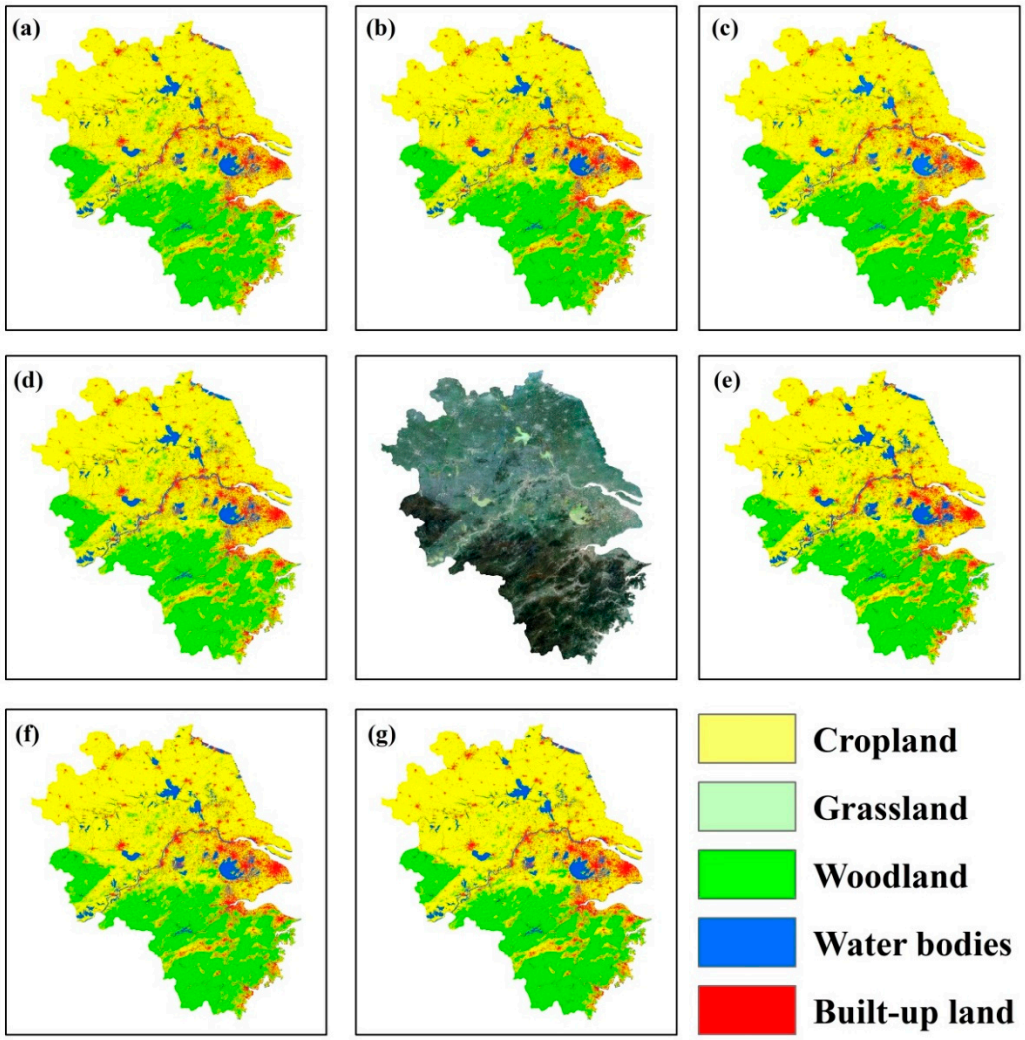
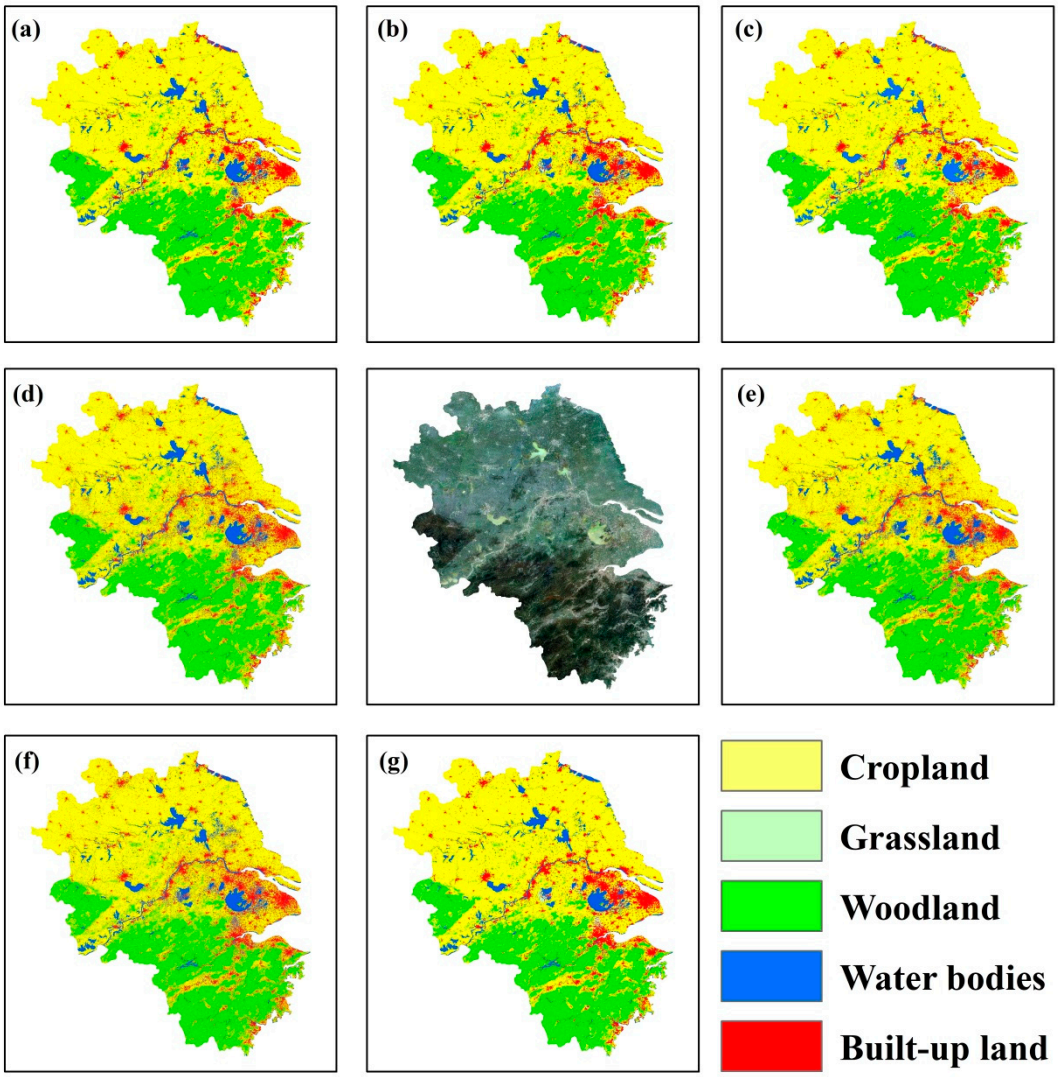


Figure S1. Comparison of pixel-based classification results: (**a**)–(**g**) indicate M1–M7 model, respectively.

9. Classification results of object-based RF models.



FigureS2. Comparison of object-based classification results: (**a**)–(**g**) indicate M1′–M7′ model, respectively.

10. Classification results by the 10-fold cross-validation models.

We applied the k-fold cross-validation method to further analyze the stability and sensitivity of the RF classification models with k set to 10 according to the experience of previous research. As is shown in Table S4, the overall accuracy of the 10-fold cross-validation classification is similar to our previous RF models. The major reason for this result is that the sampling proportion significantly differs with the 10-fold cross-validation classification applying 90% of the total sample as training data while our previous RF models using 70%. The parameter settings of our 14 RF models are the same as the 10-fold cross-validation classification except for the percentage of training sample points. Results indicate that our previous RF models are as stable and robust as the 10-fold cross-validation.

Table S4. Classification results by the 10-fold cross-validation models.

Auxiliary features used	10-fold cross-validation models		RF models	
	Pixel-based	Object-based	Pixel-based	Object-based
spectral features	91.63±0.25	94.23±0.16	91.51	94.03
spectral indices	92.38±0.23	94.94±0.22	92.73	94.67
topographic features	93.01±0.19	95.49±0.17	92.36	95.73

distance to water bodies	92.33±0.23	94.82±0.20	92.62	94.91
soil features	93.69±0.22	95.27±0.18	93.53	95.27
spectral-temporal metrics	92.74±0.16	95.05±0.19	92.73	94.95
ALL	94.02±0.28	95.93±0.21	94.20	96.01

Note that 10-fold cross-validation models results are mean overall accuracy with standard deviation, and RF models results are mean overall accuracy.

11. Time-costs of pixel- and object-based models

Table S5. Time-costs of pixel and object-based models.

	Pixel-based	**Object-based**
Image Segmentation	——	About 2 hours
Spectral Features Classification	About 17 minutes	About 30 minutes
Spectral features + Spectral index Classification	About 24 minutes	About 33 minutes
Spectral features + Topographic features Classification	About 18 minutes	About 30 minutes
Spectral features + Distance to water body Classification	About 18 minutes	About 30 minutes
Spectral features + Soil features Classification	About 24 minutes	About 33 minutes
Spectral features + Spectral-temporal metrics Classification	About 24 minutes	About 1 hour
All features classification	About 33 minutes	About 8 hours

Note that it is difficult to give an accurate time-cost because the time-cost varies in every run.