

# TEMPORAL AND SPATIAL ANALYSIS OF CLASSIFICATION TREE FOR IMPERVIOUS SURFACE MAPPING FROM SENTINEL-2 MSI DATA

J. Gao\*, Y. Chen, Y. Guo, S. Yang

School of Geographic and Biologic Information  
Nanjing University of Posts and Telecommunications, Nanjing, China - gaoj@njupt.edu.cn

Commission III, WG III/1

**KEY WORDS:** Impervious surface, Land cover classification, Classification tree, Ensemble learning, Temporal-spatial analysis, Sentinel-2 MSI.

## ABSTRACT:

For studies of urban development, it is an important method for obtaining the distribution of impervious surface (IS) areas and their dynamic change from remote sensing data. The dilemma of the same spectrum for different features and different spectrums for the same features, posed by the complexity of the IS objects, is the fundamental obstacle encountered in the extraction of urban IS areas. In this study, an automatic extraction method for urban IS areas is proposed and analyzed, based on classification and regression tree (CART) and ensemble learning strategies. The Sentinel-2 MSI data of 30 cities in China from 2018 to 2020 were selected for IS extraction experiments. We perform temporal and spatial modeling of the splitting threshold offset in the classification model to explore the effect of time and space on IS extraction. The obtained offset models show that the temporal variation is not significant, while the spatial offsets have more obvious linear relationships.

## 1. INTRODUCTION

Impervious surface (IS) is a sort of artificial covering that does not allow water to penetrate fast and is a common feature of metropolitan settings. Building tops, walkways, highways, squares, airports, parking lots, and other sorts of artificial building spaces are examples of the low permeability covered ground features. Urban IS areas are the result of human activity, and variations in the areas directly reflect the changing state of cities. IS makes it difficult for rain to penetrate the soil, reducing local infiltration and soil moisture, while increasing IS areas increases surface runoff. Furthermore, there is a strong positive relationship between IS coverage and surface temperature, and increasing IS coverage increases the urban heat island effect. For studies of urban development, it is a vital method for obtaining the distribution of IS areas and their dynamic change using remote sensing data. The dilemma of the same spectrum for different features and different spectrums for the same features, posed by the complexity of the IS features, is the fundamental obstacle encountered in the extraction of urban IS areas. Existing research has provided many strategies for the extraction of IS areas from various angles.

Different definitions of indexes have been proposed by researchers in the study of index-based approaches. Normalized difference built-up index (NDBI), normalized difference bare-soil index (NDBaI), and albedo are taken for improved linear spectral mixture analysis method was explored to estimate urban IS areas (Fan et al., 2015). And the combinational build-up index (CBI), principal component analysis, normalized difference water index (NDWI), and soil-adjusted vegetation index (SAVI), are proposed to extract IS (Sun et al., 2016) (Sun et al., 2017a). Another index, the normalized IS index (NISI), integrates DMSP-OLS and Moderate Resolution Imaging Spectroradiometer (MODIS) normalized difference vegetation index (NDVI) data, was provided (Guo et al., 2017). A modified

normalized difference IS index (MNDISI) is proposed, and a Gaussian-based automatic threshold selection method is used to identify the optimal MNDISI threshold for delineating IS from background features (Sun et al., 2017b). For the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) spectral data, a perpendicular IS index (PISI) was proposed (Tian et al., 2018) for IS extraction. Considering the topography of the target area, some scholars have proposed indexes to extract IS areas and shaded areas separately (Luo and Chen, 2021), even considering LiDAR data (Luo et al., 2018).

There are different ways to construct IS extraction models with the proposed indexes. The MODIS NDVI and DMSP/OLS nighttime light imageries are taken to establish regression models that estimated the fraction of IS area in urban green space (Kuang, 2019). The random forest model is employed for direct and comprehensive IS extraction on subpixel estimation (Deng et al., 2017). The IS extracting methods based on machine learning, fully convolutional neural networks (FCNN), are provided at pixel level with high-resolution satellite imagery (McGlinchy et al., 2019) (Lin et al., 2020). A method called Continuous Subpixel Monitoring (CSM) was developed to map and monitor urban IS area change continuously at the subpixel level (Deng and Zhu, 2020). To alleviate the difficulties in the conventional scheme, the geospatial distribution knowledge of IS areas at large scales was taken (Huang et al., 2021). The combined use of the data from Sentinel-1 dual-polarized Synthetic Aperture Radar (SAR) and Sentinel-2 Multispectral Sensor Images (MSI) have been successfully applied in many remote sensing applications. Scholars also propose a method for IS mapping by synergetic fusion of dual-polarized SAR and multispectral information (Sun et al., 2022), (Shrestha et al., 2021), (Chen et al., 2021), (Feng and Fan, 2021), (Li et al., 2021).

This paper presents our work on IS extraction modeling and temporal and spatial analysis using sentinel-2 satellite data,

\* Corresponding author

based on the ensemble learning model of classification trees. The Sentinel-2 satellites deliver high-resolution multispectral data, which is critical for IS studies. To create training samples, we make a combination of existing land-cover data and Sentinel-2 data. For classification and regression tree (CART) learning, we use repeated sampling and training of the IS extraction tree model for ensemble learning. IS extraction experiments are conducted for 30 cities in China in different periods of sentinel-2 data from 2018 to 2020. Then, the temporal and spatial relationships of classification trees are studied and analyzed.

## 2. IS EXTRACTION BASED ON CART

### 2.1 Data Sources

In the study, we used Sentinel-2 MSI acquisitions and global land cover from Copernicus Global Land Service. Sentinel-2 MSI delivers multi-spectral imageries with a wide field of view that can be used to monitor plant, soil, and water cover, as well as observe interior waterways and coastal areas. Sentinel-2 MSI detects the Earth's reflected brightness in 13 spectral bands, ranging from visible and near-infrared (VNIR) to short-wave infrared (SWIR). We select four bands (Green, Red, NIR and SWIR1) from Sentinel-2 data for the study. The land cover data provides a global dynamic land cover map at 100 m spatial resolution, which includes a continuous field for all basic land cover classes that provide proportional estimates for vegetation/ground cover for the land cover types. Band "urban-coverfraction" of the Land Cover maps (v3.0.1) in 2019 over the entire Globe, which reaches an accuracy of 80% at Level 1, is taken for generating IS samples. Sentinel-2 and land cover data are available on Google Earth Engine (GEE) platform.

### 2.2 Experimental Sites

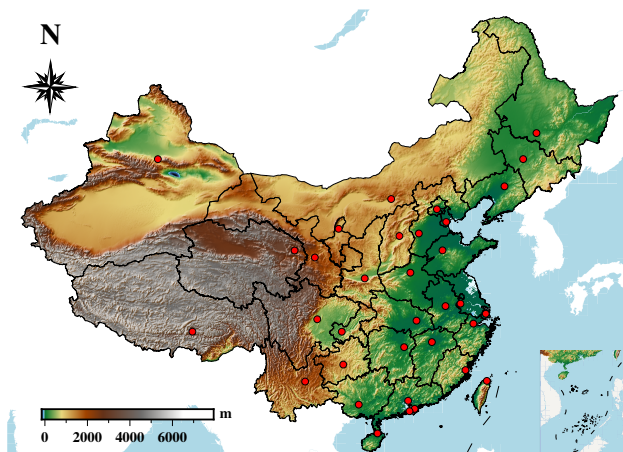


Figure 1. The sites (red points) of Chinese cities for IS extracting experiments in the height map.

The national capital, municipalities directly under the central government, and the capital cities of provinces and autonomous regions in mainland China, Hong Kong, Macao, and Taipei were selected as experimental cities for impermeable surface extraction, as shown in Figure. 1. We use Sentinel-2 data from January 2018 to December 2020, and corresponding land cover data in 2019. Multispectral data with less than 20% cloud coverage in the range of two months around January 1 (winter), March 1 (spring), June 1 (summer), and September 1 (autumn)

were taken as experimental data. The Sentinel-2 data from Nanning in Guangxi Province, Guiyang in Guizhou Province, Xining in Qinghai Province, and Lhasa in Tibet Autonomous Region were excluded from the experiment because their data did not meet the requirements.

### 2.3 Decision Tree Base Learners

The method of decision tree for IS extraction is based on the nonlinear relationship

$$h = T(\rho), \quad (1)$$

where,  $\rho$  refers to attribute values form the description data, namely the four bands form Sentinel-2 in the study,  $h$  refers to the class labels, 1 for IS and 0 for non-IS.  $T$  refers to the classification tree model. The class labels of samples are set by

$$h_s = \begin{cases} 1 & \rho_c \geq t \\ 0 & \rho_c < t \end{cases}, \quad (2)$$

where,  $t$  refers to the IS rate threshold, it is set to 30% in the experiments. We combine class labels and the four band values at random positions to generate sample data for training IS extraction models.

The decision tree model is taken for the base learners. A decision tree can learn to split the training samples into subsets based on attribute values. A decision tree is constructed top-down by choosing an attribute at each step according to how well the attribute splits the set into homogeneous subsets with the same class label. Here we take the Gini purity as the measure.

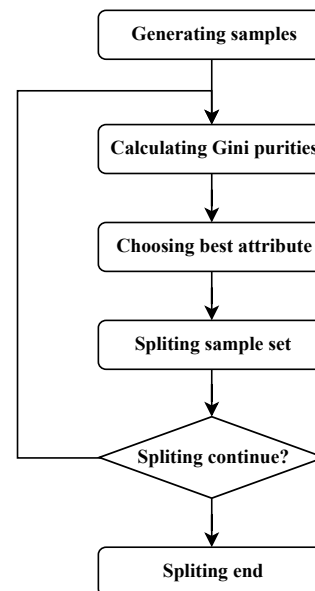


Figure 2. Flow chart of CART base learner training.

The Gini impurity is calculated for different spectral data of the sample data as follows,

$$I_G = 1 - (p_0^2 + p_1^2), \quad (3)$$

where  $p_0$  refers to the fraction part of  $h_s = 1$  and  $p_1$  refers to the non-IS part with  $h_s = 0$ . The minimum Gini impurity is taken as the priority classification attribute. The optimal bisection

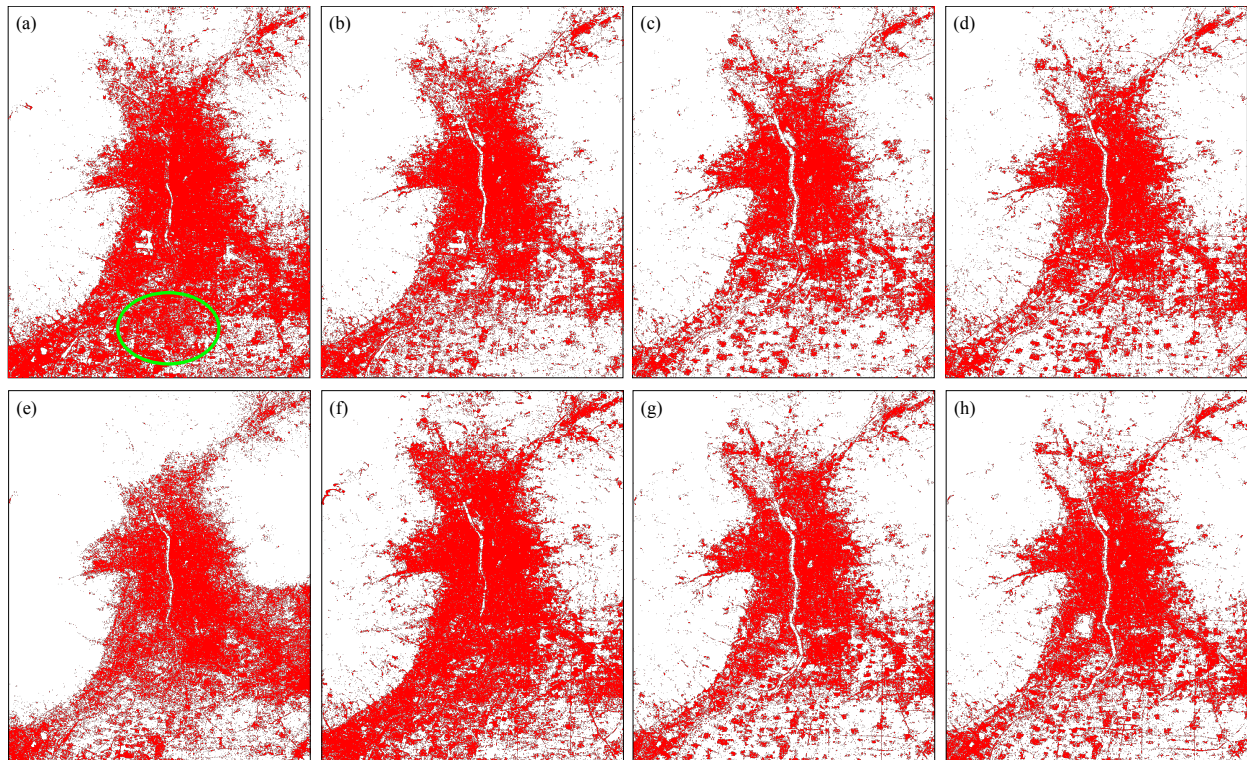


Figure 3. Extracted IS data of Taiyuan City, Shanxi Province, in 2018, (a) for January, (b) for March, (c) for June, (d) for September, and in 2019, (e) for January, (f) for March, (g) for June, (h) for September.

parameter of the classification tree is determined as shown in Figure. 2.

This process is repeated on each derived subset recursively. When subsets all have a same class label, or when there is no gain to the prediction, recursion is done. When the splitting recursion is completed, a decision tree has been built up. An IS extracting result example in 2018 of Taiyuan City in Shanxi Province is shown in Figure. 3.

However, due to the limitations of accuracy as well as resolution, there are imperfect or incorrect labels inevitably exist in the training samples. The quality of the extracting results can be compromised from multiple trees. As shown in Figure 1, the marked area in (a) is significantly different from the other subplots. We employ the ensemble learning model to reduce the impact.

#### 2.4 Ensemble Learning Based on Decision Trees

Ensemble learning is the process of combining the outputs of multiple learners to give the final result. To keep the samples highly differentiated for each learner, samples are generated in a repeated random sampling mode. Each sampling brings about a 1000-sample set, including 500 IS samples and 500 non-IS samples in our experiments. Each sample set is used to train a decision tree. These decision trees are used to vote for the results of classifying. The voting results  $H$  can be calculated as

$$H = \begin{cases} 1 & s \geq 0.5 \\ 0 & s < 0.5 \end{cases}, \quad (4)$$

where

$$s = \sum_{i=1}^n \frac{h_i}{n},$$

and  $h_i$  refers to the output form decision tree  $i$ .

Since the definition of classification trees is usually based on the attribute values, it is not convenient to handle the splitting thresholds when performing classification tree synthesis. To simplify the operation, we convert the splitting-based continuous tree structure model into a discrete model based on a regular grid. The discrete attribute values are defined as

$$\rho \in \{\rho_0 + n \cdot \Delta\rho \mid n \in \mathbf{Z}\}, \quad (5)$$

where,  $\rho_0$  refers to an attribute reference value,  $\Delta\rho$  is the attribute grid interval, and  $\mathbf{Z}$  refers to the integer set. The discretized tree models are synthesized according to Equation. 4, i.e., the extraction model corresponding to ensemble learning. The flow is described as shown in Figure. 4.

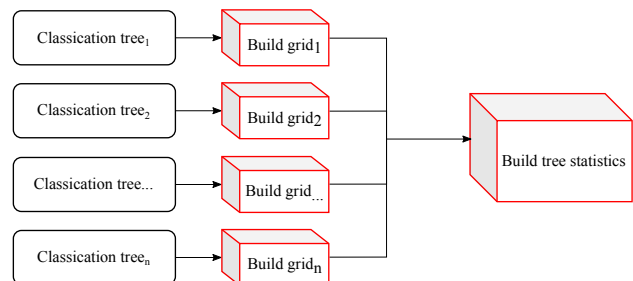


Figure 4. The flow of classification trees in ensemble learning.

In our experiments, based on the statistics of the four bands of the sample data, we define the discrete attribute  $\rho_i$  in range of five times the standard deviation as

$$\rho_i \in \left\{ \mu_i + n \cdot \frac{5\sigma_i}{N} \mid n = -N, -N + 1, \dots, N \right\},$$

where,  $i$  refers to the band index,  $\mu_i$  refers to the mean value of samples from band  $i$ , and  $N$  refers to dimension radius of the grid ( $N = 100$  in the experiments).

We consider that there is no significant change in IS in urban cities within three years, and the classification tree models for the same period are combined as an integrated model for that time. For example, a part of the tree models for January 2018, 2019, 2020, and the integrated model are shown in Figure. 5.

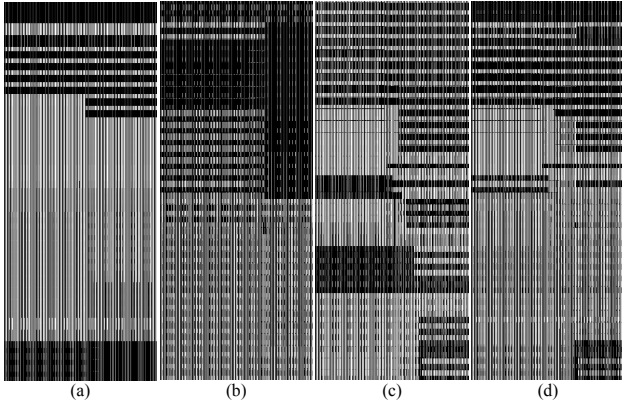


Figure 5. A part of the tree models of Taiyuan City, Shanxi Province, for January 2018 (a), 2019 (b), 2020 (c), and the integrated result (d). The horizontal and vertical axes correspond to two bands, respectively.

### 3. TEMPORAL ANALYSIS OF CLASSIFICATION TREE MODELS

#### 3.1 Temporal Tree Offset Modeling

The classification tree model is built based on attribute splitting values. The study here focuses on the analysis of the variation of attribute splitting thresholds in time and space dimensions. It is assumed that the model of the classification tree is largely consistent, which can be ensured by the quality of the experimental results. The analysis of the temporal dimension focuses on the changes in the offsets of the splitting points between periods. For analysis purposes, we define a measure of the IS distance of attributes,  $S$ , as follows,

$$S(\rho) = \begin{cases} d(\rho) & H(\rho) = 1 \\ -d(\rho) & H(\rho) = 0 \end{cases}, \quad (6)$$

where,  $d(\rho) = \|\rho - \rho_t\|^2$  refers to the distance from  $\rho$  to  $\rho_t$ , and  $\rho_t$  refers to the nearest attribute splitting point.  $S$  represents the signed distance from the target to the splitting point, taking a positive sign in the IS areas and a negative sign in the non-IS areas.

Based on the IS distances of the attributes of different decision tree models, we define the IS distance difference  $\delta S$  between models as follows,

$$\delta S_i(\rho) = \begin{cases} S_0(\rho) - S_i(\rho) & S'_0(\rho) > 0, S'_i(\rho) > 0 \\ S_i(\rho) - S_0(\rho) & S'_0(\rho) < 0, S'_i(\rho) < 0 \\ 0 & \text{others} \end{cases}, \quad (7)$$

where,  $S_0$  refers to the distance to the reference model,  $S'$  refers to the gradient.  $\delta S$  describes the distance difference between

the attributes of the target model and the reference model. This distance difference will be more accurate if the two models are more similar. The distance difference is directly related to the offset of the splitting point, namely,

$$\delta S_i(\rho) = \rho_{t_i} - \rho_{t_0}$$

As shown in Figure. 6, the offsets  $a$  and  $b$  are related to the area of the difference part.

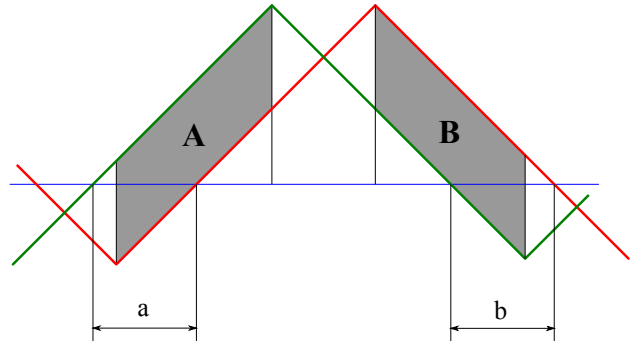


Figure 6. Tree model offset and attribute IS distances. The red and green lines refer to attribute IS distance for the two different tree models, and the blue line refers to the zero distance reference. Offset  $a$  and  $b$  between the models is related to the area  $A$  and  $B$ .

We obtain by integrating the model over the defined range of attribute,  $f$ , namely

$$f_i = \frac{\sum_{n=-N}^N S_i(\rho_n) \Delta \rho}{\sum_{n=-N}^N \Delta \rho}, \quad S_i(\rho_n) \neq 0. \quad (8)$$

For each city, we obtained the IS distances for the attributes during the four different periods of the year. If the attribute IS distances changes over time, there is a corresponding offset in their corresponding classification tree splitting points. We consider the relationship between the corresponding offsets of different bands and periods and establish the following function

$$f(\tau) = \alpha + \beta\tau + \gamma\tau^2, \quad (9)$$

where,  $\tau$  refers to the temporal variable defined in 1, 2, 3, 4 for the four periods,  $\alpha, \beta, \gamma$  refer the offset model coefficients to be determined.

#### 3.2 Temporal Tree Offset Analysis

Based on the offset observation data and the quadratic model, the residual equation can be constructed as follows,

$$V = B\hat{X} - L, \quad (10)$$

where  $B$  = coefficient matrix consisting of 1,  $\tau$ ,  $\tau^2$   
 $\hat{X}$  = parameter vector consisting of  $\alpha$ ,  $\beta$ ,  $\gamma$   
 $L$  = attribute offset values  $f$  at different times  
 $V$  = residual vector of the offset model.

The solution of the equation can be obtained as follows,

$$\hat{X} = (B^T B)^{-1} B^T L. \quad (11)$$

We calculate the average of IS distances for the four bands of data for each city as the reference model. The offset values for four different times for thirty cities were calculated according to Equation. 8. These offset values are brought into the equation as data to solve the parameters of the offset model for each of the four bands. The obtained offset model parameters are listed in Table. 1.

Band	$\alpha$	$\beta$	$\gamma$
Green	-0.0067	+0.0649	-0.0246
Red	+0.0064	-0.0091	+0.0002
NIR	-0.0145	+0.0162	-0.0020
SWIR1	-0.0280	+0.0430	-0.0113

Table 1. Temporal offset model parameters of classification trees.

The continuous curves of the temporal offset corresponding to the four bands are shown in Figure. 7. The figure shows that among the four bands, the trend is not obvious in the other three bands, except for the Green band which has significant negative growth.

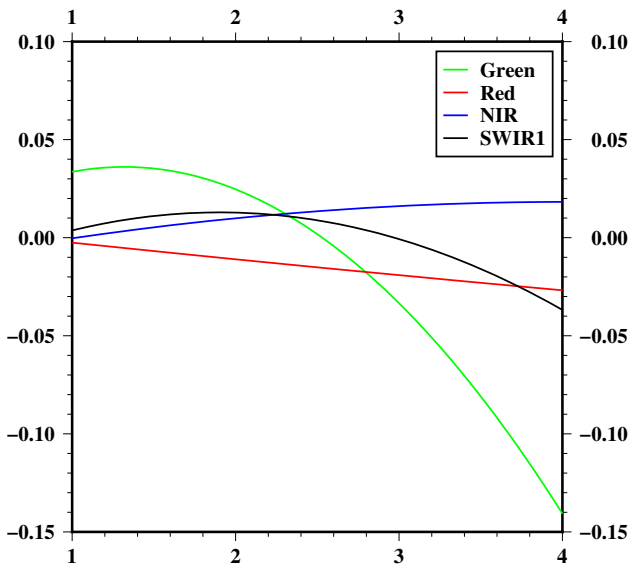


Figure 7. Temporal offset model of classification trees for band Green, Red, NIR, SWIR1 of Sentinel-2 MSI.

The covariance of the estimated parameters,  $D$ , is calculated as follows,

$$D = \frac{V^T V (B^T B)^{-1}}{r}, \quad (12)$$

where  $r$  refers to the degrees of freedom. The standard deviation information of the parameters is extracted from  $D$  and listed in Table. 2.

Band	$\sigma_\alpha$	$\sigma_\beta$	$\sigma_\gamma$
Green	0.0951	0.1527	0.0488
Red	0.0078	0.0125	0.0040
NIR	0.0099	0.0158	0.0051
SWIR1	0.0094	0.0150	0.0048

Table 2. The standard deviation of the estimated temporal model parameter values.

From the standard deviation data in the table, we can find that the confidence level of the model parameters being non-zero is not high. The effect of spatial factors could be mostly ignored.

#### 4. SPATIAL ANALYSIS OF CLASSIFICATION TREE MODELS

We relate the offsets of the tree model to the spatial locations, expressed in the expression as follows,

$$f(\varphi, \lambda) = c_0 + c_1\varphi + c_2\lambda + c_3\varphi^2 + c_4\lambda^2 + c_5\varphi\lambda, \quad (13)$$

where,  $\varphi, \lambda$  refer to centered longitude and latitude coordinate values,  $c_i$  refers to the model coefficients to be determined.  $\varphi, \lambda$  can be expressed as follows,

$$\varphi = \varphi_1 - \varphi_0, \quad \lambda = \lambda_1 - \lambda_0,$$

where,  $\varphi_1$  and  $\lambda_1$  refer to the original longitude and latitude coordinate,  $\varphi_0$  and  $\lambda_0$  refer to the reference longitude and latitude coordinate. In the experiment, we select the latitude and longitude of Xi'an City in Shaanxi Province, (108.93689°, 34.278453°), as the reference. We obtain the parameters of the spatial offset model, which are listed in Table. 3.

Index	$c_0$	$c_1$	$c_2$	$c_3$	$c_4$	$c_5$
$\tau_1 b_G$	38.90	-5.89	-1.61	0.58	-0.17	0.37
$\tau_1 b_R$	5.81	-0.83	5.32	0.08	0.07	-0.10
$\tau_1 b_N$	-54.08	0.14	1.65	0.15	0.22	-0.10
$\tau_1 b_S$	13.98	-0.33	0.68	-0.15	-0.11	0.08
$\tau_2 b_G$	28.68	7.66	14.36	-0.90	0.69	-0.20
$\tau_2 b_R$	0.33	-0.20	0.41	-0.13	0.32	-0.04
$\tau_2 b_N$	-25.66	-0.56	-0.09	-0.02	-0.02	0.01
$\tau_2 b_S$	25.06	-2.01	2.18	-0.10	0.26	0.14
$\tau_3 b_G$	32.46	-9.63	1.40	-0.31	1.46	-0.30
$\tau_3 b_R$	-10.48	-0.06	1.05	0.01	-0.05	-0.02
$\tau_3 b_N$	17.79	-0.63	-0.94	0.08	-0.32	0.05
$\tau_3 b_S$	29.79	-0.95	2.20	-0.02	0.20	0.04
$\tau_4 b_G$	91.93	-0.64	2.49	-0.44	-0.86	-0.26
$\tau_4 b_R$	-17.49	-0.23	0.83	0.03	-0.03	0.02
$\tau_4 b_N$	1.40	0.24	-1.18	0.03	-0.24	0.08
$\tau_4 b_S$	15.30	-1.46	2.72	-0.11	0.29	-0.07

Table 3. Parameters ( $\times 10^{-3}$ ) of the spatial offset model,  $\tau$  for period index and  $b$  for band index.

From the comparison of the model parameters in the table, it can be found that there is a certain regularity in the value of  $c_0$  at different bands, with the band order decreasing before increasing.  $c_1$  and  $c_2$  tend to have larger values in b1 and b4, indicating that the spatially correlated linear shifts are more pronounced in these two bands.  $c_3, c_4$  and  $c_5$  do not have large values, indicating that the spatial quadratic feature is not significant. These 16 spatial offset models are plotted as shown in Figure 8. The linear variation characteristic of the offsets can be found visually.

#### 5. CONCLUSIONS

We employed CART to achieve automatic IS extraction with sentinel-2 MSI and land cover classification data for 30 cities in China from 2018 to 2020, and analyze the temporal and spatial offsets of the classification model splitting thresholds. To reduce the sensitivity of the classification model to imperfect data, we ensemble the three models for the same period of three years. The temporal and spatial offset models were built up by linear regression. The results show that the offsets are not significant in the temporal dimension, while the linear variations in the spatial dimension are relatively more pronounced.

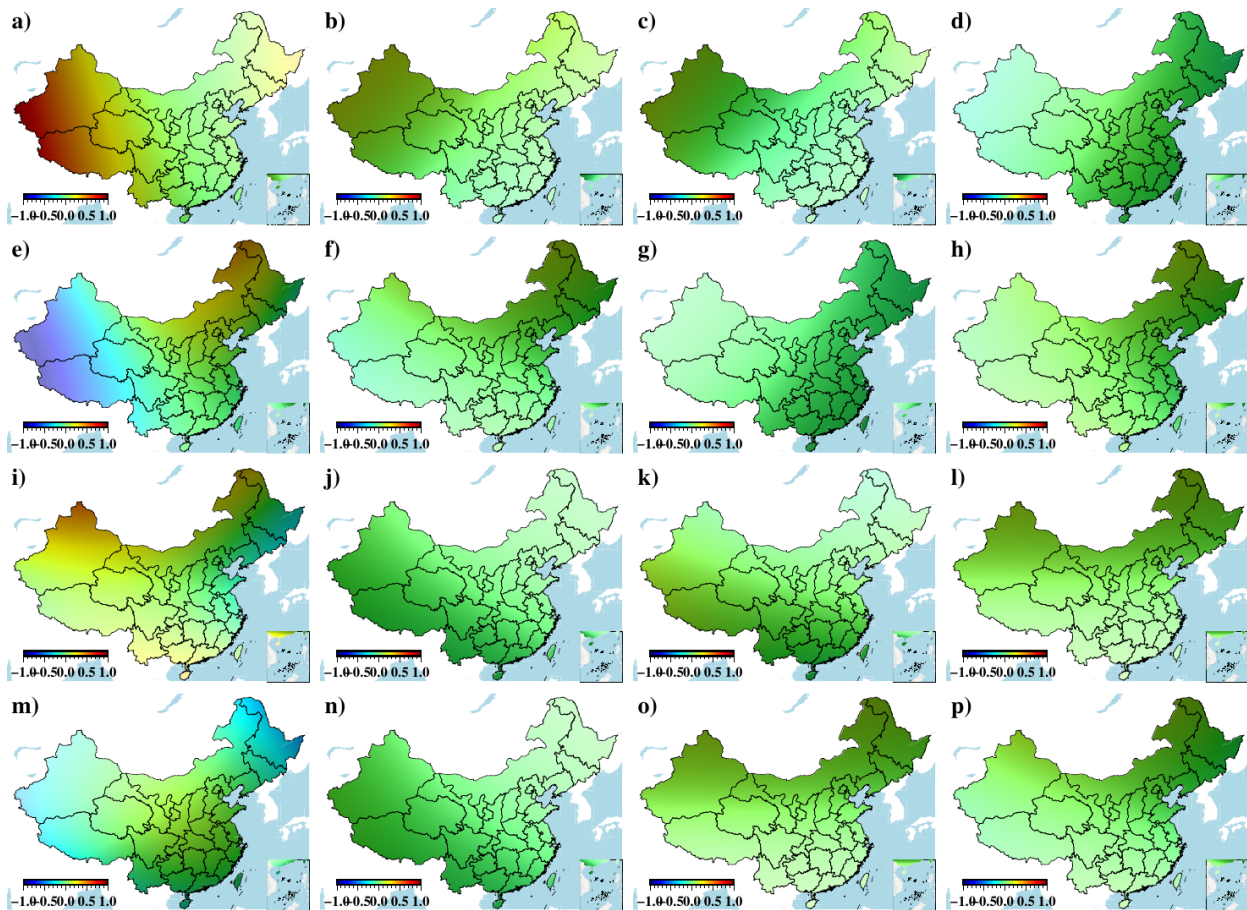


Figure 8. The spatial offset model of the experimental cities. The first row refers to January, the second row to March, the third row to June, and the fourth row to September. The columns from left to right refer to the bands Green, Red, NIR and SWIR1.

## REFERENCES

- Chen, R., Li, X., Zhang, Y., Zhou, P., Wang, Y., Shi, L., Jiang, L., Ling, F., Du, Y., 2021. Spatiotemporal Continuous Impervious Surface Mapping by Fusion of Landsat Time Series Data and Google Earth Imagery. *Remote Sensing*, 13(12), 2409. Number: 12 Publisher: Multidisciplinary Digital Publishing Institute.
- Deng, C., Li, C., Zhu, Z., Lin, W., 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., Zhu, Z., 2020. Continuous subpixel monitoring of urban impervious surface using Landsat time series. *Remote Sensing of Environment*, 238, 110929.
- Fan, F., Fan, W., Weng, Q., 2015. Improving Urban Impervious Surface Mapping by Linear Spectral Mixture Analysis and Using Spectral Indices. *Canadian Journal of Remote Sensing*, 41(6), 577–586.
- Feng, S., Fan, F., 2021. Impervious surface extraction based on different methods from multiple spatial resolution images: a comprehensive comparison. *International Journal of Digital Earth*, 14(9), 1148–1174.
- Guo, W., Lu, D., Kuang, W., 2017. Improving Fractional Impervious Surface Mapping Performance through Combination of DMSP-OLS and MODIS NDVI Data. *Remote Sensing*, 9(4), 375. Number: 4 Publisher: Multidisciplinary Digital Publishing Institute.
- Huang, M., Chen, N., Du, W., Wen, M., Zhu, D., Gong, J., 2021. An on-demand scheme driven by the knowledge of geospatial distribution for large-scale high-resolution impervious surface mapping. *GIScience & Remote Sensing*, 58(4), 562–586.
- Kuang, W., 2019. Mapping global impervious surface area and green space within urban environments. *Science China Earth Sciences*, 62(10), 1591–1606.
- Li, C., Shao, Z., Zhang, L., Huang, X., Zhang, M., 2021. A Comparative Analysis of Index-Based Methods for Impervious Surface Mapping Using Multiseasonal Sentinel-2 Satellite Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14, 3682–3694. Conference Name: IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.
- Lin, Y., Zhang, H., Lin, H., Gamba, P. E., Liu, X., 2020. Incorporating synthetic aperture radar and optical images to investigate the annual dynamics of anthropogenic impervious surface at large scale. *Remote Sensing of Environment*, 242, 111757.
- Luo, H., Chen, N., 2021. A Combined Unmixing Framework for Impervious Surface Mapping on Medium-Resolution Images with Visible Shadows. *Photogrammetric Engineering & Remote Sensing*, 87(6), 431–443.

Luo, H., Wang, L., Wu, C., Zhang, L., 2018. An Improved Method for Impervious Surface Mapping Incorporating LiDAR Data and High-Resolution Imagery at Different Acquisition Times. *Remote Sensing*, 10(9), 1349. Number: 9 Publisher: Multidisciplinary Digital Publishing Institute.

McGlinchy, J., Johnson, B., Muller, B., Joseph, M., Diaz, J., 2019. Application of unet fully convolutional neural network to impervious surface segmentation in urban environment from high resolution satellite imagery. *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, 3915–3918.

Shrestha, B., Ahmad, S., Stephen, H., 2021. Fusion of Sentinel-1 and Sentinel-2 data in mapping the impervious surfaces at city scale. *Environmental Monitoring and Assessment*, 193(9), 556.

Sun, G., Chen, X., Jia, X., Yao, Y., Wang, Z., 2016. Combinational Build-Up Index (CBI) for Effective Impervious Surface Mapping in Urban Areas. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(5), 2081–2092. Conference Name: IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.

Sun, G., Chen, X., Ren, J., Zhang, A., Jia, X., 2017a. Stratified spectral mixture analysis of medium resolution imagery for impervious surface mapping. *International Journal of Applied Earth Observation and Geoinformation*, 60, 38–48.

Sun, G., Cheng, J., Zhang, A., Jia, X., Yao, Y., Jiao, Z., 2022. Hierarchical fusion of optical and dual-polarized SAR on impervious surface mapping at city scale. *ISPRS Journal of Photogrammetry and Remote Sensing*, 184, 264–278.

Sun, Z., Wang, C., Guo, H., Shang, R., 2017b. A Modified Normalized Difference Impervious Surface Index (MNDISI) for Automatic Urban Mapping from Landsat Imagery. *Remote Sensing*, 9(9), 942. Number: 9 Publisher: Multidisciplinary Digital Publishing Institute.

Tian, Y., Chen, H., Song, Q., Zheng, K., 2018. A Novel Index for Impervious Surface Area Mapping: Development and Validation. *Remote Sensing*, 10(10), 1521. Number: 10 Publisher: Multidisciplinary Digital Publishing Institute.