Academic Journal of Science and Technology
ISSN: 2771-3032 | Vol. 15, No. 2, 2025

A Simple Method for Backdating High-Rise Building Areas from Time Series Data Using Contemporary Spatial Constraints

Zhenzhu Wang[1, 2], Liwei Li[2], Gang Cheng[1, *]

[1] School of Mapping and Land Information Engineering, Henan Polytechnic University, Jiaozuo, Henan 454000, China
[2] Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China
[*] Corresponding author: Gang Cheng (Email: chenggang@hpu.edu.cn)

Abstract: In the past two decades, China has experienced rapid urbanization, resulting in the widespread emergence of High-Rise Building Areas (HRBAs). However, open public data on the completion times of HRBAs in large scale is still lacking due to various reasons. To fill this gap, this study proposes a simple backdating method for extracting the construction years of HRBAs using historical Landsat-5/7/8 imagery. Using the HRBAs results of the 2020 as constraints, the method constructs change features by combining spectral and spatial information to accurately capture pixel-level changes in HRBAs. Additionally, a window-based statistical replacement technique is used to minimize the influence of outlier on prediction accuracy. The proposed method is validated using Landsat data from 2000 to 2020, optimized in Zhengzhou, and tested in diverse regions including Jiaozuo, Xinyang, and Nanyang. The results demonstrate that, within a ±2-year threshold, the method achieves extraction accuracies of 80.82% (Zhengzhou), 87.31% (Jiaozuo), 66.95% (Nanyang), and 70.95% (Xinyang), confirming its effectiveness in accurately tracking HRB construction years. The proposed approach offers an efficient solution for backdating HRBAs construction times in large scale and can help provide valuable insights into the urbanization process in China.

Keywords: High-rise building areas, backdating, composite impervious surface extraction index, long term Landsat data.

1. Introduction

According to the Seventh National Population Census of China, by the end of 2020, the urban resident population had reached 901.99 million, accounting for 63.89% of the total population [1]. Along with rapid urbanization, High-Rise Building Areas (HRBAs) have become symbols of development in a region. HRBAs are defined as building areas with spatial connectivity, where the individual building heights are similar and the average height is greater than 25 meters [2]. The impact of HRBAs is complex and diverse. They create more usable space within limited areas, ease land use constraints [3-5], but also provide conditions for the large-scale outbreak and spread of diseases due to population concentration [6,7]. Additionally, high-rise buildings directly interact with the lower atmosphere, altering the urban surface heat loss rate and forming heat island effects, which subsequently change the urban climate [8,9]. Furthermore, HRBAs enhance the city's sense of hierarchy and three-dimensionality, becoming an integral part of the urban landscape [10,11]. HRBAs differ fundamentally from individual buildings as shown in Table 1: spatially, HRBAs exhibit continuous planar distributions compared to discrete building polygons; spectrally, their mixed pixel composition generates higher heterogeneity than uniform structural responses; functionally, they represent urban planning-driven aggregates (e.g., business districts) rather than owner-specific modifications. These distinctions necessitate specialized methodologies for accurate HRBAs characterization. Therefore, understanding the construction years and spatial distribution of HRBAs is of significant importance for research in land resource utilization, disease transmission, urban climate, and landscape studies.

Table 1. Comparison of HRBAs in Sentinel-2 and ESRI Imagery.

Sentinel-2 Image	ESRI Image	Sentinel-2 Image	ESRI Image
			

Time-series satellite imagery has been widely used to monitor and analyze long-term land use and environmental changes. The backdating of HRBAs construction timelines (2000–2020) demands satellite data with both multi-decadal continuity and operational feasibility. High spatial resolution imagery, despite its superior spatial granularity, faces critical limitations for large-area long-term studies: prohibitive acquisition costs, fragmented archives (typically post-2010), and insufficient temporal resolution. Landsat data addresses these challenges through its 40-year continuous global coverage, free accessibility, and predictable 16-day revisit cycle. These attributes enable cost-effective monitoring of building area evolution across urban agglomerations while maintaining the temporal density required for precise construction year estimation—a decisive factor in selecting Landsat as this study's primary data source. The integration of its multi-decadal archives (2000–2020) with advanced algorithms has proven instrumental in backdating long-term surface changes, as evidenced by successful applications in monitoring cropland abandonment [12], urban sprawl [13], forest degradation [14], surface water dynamics [15], slow-moving landslides [16], urbanization [17], and surface water dynamics [18].

Previous studies have employed the deep learning method to extract the yearly dynamic of HRBAs from time-series optical imagery [19,20]. They first extract the spatial distribution of HRBAs at each time points and then determine the construction years of HRBAs across these time points. However, when applied to long-term datasets, the single time point deep learning method may lead to repeated extractions on same HRBAs and largely reduce the efficiency. This is because HRBAs once built will normally stay for a long time. Furthermore, the application of deep learning to long-term HRBAs analysis is limited by reliance on long-term annotated data and opaque decision mechanisms. In contrast, spectral indices offer physics-driven thresholds, cross-sensor compatibility, and interpretable time-series trajectories, aligning with long-term HRBAs analysis needs. In a recent work, Uhl and Leyk proposed a backdating method to estimate the construction years of buildings [21]. This method combines time-series Landsat data with building footprint data to estimate the construction year of each building by curve fitting using spectral indices, and verifies the feasibility of estimating building years based on Landsat data. In the method, the interference of non-target areas is effectively avoided by limiting the interested area to the range of building footprint data. However, the accuracy of the building construction year estimation results in the method is very low, with an average error of 7 years. Additionally, HRBAs have rather different image characteristics compared to buildings. Thus the method can not be easily extended to accurately backdate HRBAs.

To overcome the limitations of existing methods in backdating HRBAs from time-series optical imagery, this study introduces a streamlined backdating framework that prioritizes simplicity and computational efficiency, specifically designed for HRBAs in long-term satellite data. The contribution is fourfold:

1. A threshold-driven workflow requiring only a fixed threshold of 0 for initial HRBAs construction year prediction, eliminating complex parameter tuning while maintaining robustness across diverse urban landscapes.

2. Spatiotemporal fusion that integrates spectral indices with localized window statistics, enabling rapid detection of HRBAs emergence without reliance on ancillary building footprint data.

3. Computational lightweight design compatible with multi-sensor Landsat archives (5/7/8), facilitating large-scale applications through minimal preprocessing steps.

4. Validating the proposed method on four diverse urban areas, showing the effectiveness of the method at backdating HRBAs in large scale.

2. Methodology

2.1. Overall Technical Process

In order to extract the construction time of HRBAs, a simple backdating method with contemporary spatial constraints is proposed in this paper. The process can be divided into three main steps: (1) Data Preprocessing; (2) Algorithm Construction; and (3) Accuracy Evaluation. The flowchart of the proposed method is illustrated in Figure 1.

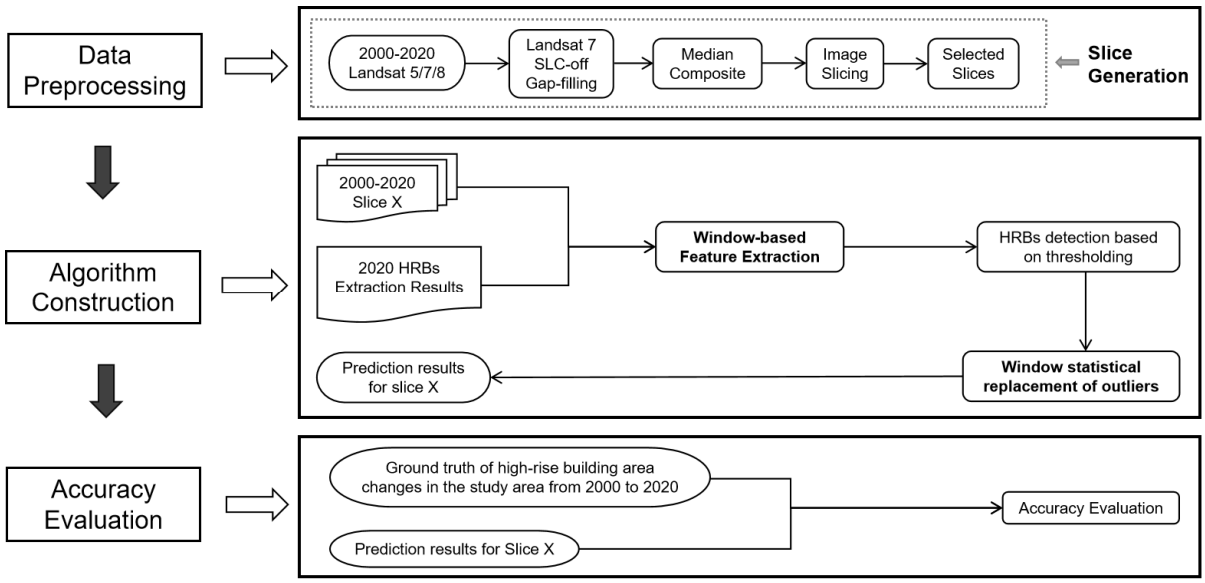


Figure 1. The flowchart of the proposed method.

2.2. Data Preprocessing

Landsat-5, Landsat-7 and Landsat-8 time-series data were selected as data sources, and the data were downloaded through the Google Earth Engine platform (https://code.earthengine.google.com/). In order to minimize the influence of cloud cover on the data, data with less than 5% cloud cover from October to December each year were selected to construct a complete time-series. Given that current cloud removal techniques often result in the loss of original information in the images, this study did not process the data for cloud removal, but instead used median-based temporal synthesized images directly [22]. After temporal synthesis, the remotely sensed images were divided into slices, from which the slices containing HRBAs were selected for subsequent analysis.

2.3. Algorithm construction

2.3.1. Window-based Feature Extraction

As aggregated building clusters, HRBAs constitute critical components of impervious surfaces [23,24], thus making them detectable through impervious surface indices that capture their spectral-spatial characteristics at scale. HRBAs encompass diverse impervious surfaces including buildings, roads, and concrete pavements. The composite impervious surface index (CBI) [25] was selected for its capacity to characterize these heterogeneous features through synergistic integration of spectral responses from built-up areas (NDBI), vegetation suppression (NDVI), and water exclusion (MNDWI). By leveraging negative inter-index correlations, CBI effectively isolates impervious surfaces from complex urban backgrounds while mitigating interference from non-target land covers. The calculation formula is defined as:

$$CBI = NDBI - NDVI\ MNDWI, \tag{1}$$

$$NDBI = \frac{SWIR1 - NIR}{SWIR1 + NIR}, \tag{2}$$

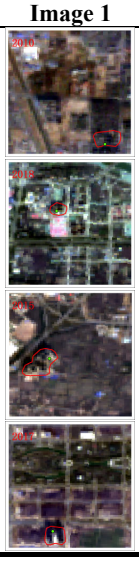
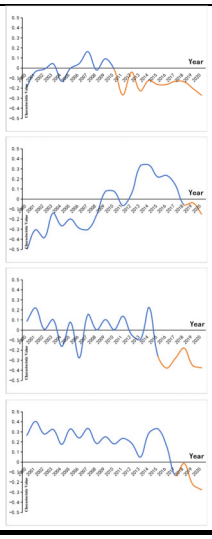
$$NDVI = \frac{NIR - RED}{NIR + RED}, \tag{3}$$

$$MNDWI = \frac{GREEN - SWIR1}{GREEN + SWIR1}, \tag{4}$$

Among them, NDBI reflects information about built-up areas. NDVI is used to reflect vegetation coverage in the area. MNDWI is used to detect water bodies. SWIR1 refers to the shortwave infrared band 1, NIR is the near-infrared band, RED is the red band, and GREEN is the green band.

CBI as a kind of pixel-level spectral indice for impervious surface may be not stable in characterizing HRBAs with complex geometric structures, thus a new spectral and spatial feature is constructed by calculating feature values on a per-pixel basis, which involves forming a window centered on a pixel. Within this window, the average feature value of all pixels that are both within the object boundary and inside the window is calculated. This average is then assigned as the average feature value assigned to the center pixel. This process is repeated for each pixel to obtain the feature for value distribution for the entire image. Utilizing a window on top of the pixel-by-pixel calculation helps better capture the spatial correlations of changes, rather than just single-pixel variations.

Table 1. Analysis of Images and CBI Derived Feature Value Changes at Specific Points During Construction of HRBAs.

Pre-construction Image	Post-construction Image 1	Post-construction Image 2	Feature Value Change Line Chart
			

To demonstrate the effectiveness of the CBI derived feature for detecting HRBAs, Table 2 presents an analysis of image and feature value changes during the construction phase in four typical samples. In the line graphs, blue lines indicate the period prior to the construction of HRBAs, and orange lines indicate the period after the construction of HRBAs. Table 2 illustrates how the CBI derived feature values change as HRBAs are constructed over time, highlighting the consistency and reliability of this feature in identifying the construction period. It should be noted that the image characteristics of HRBAs change every year. We calculate the point feature value and then arrange the feature values in chronological order. The construction time of HRBAs is determined by analyzing the curve: starting from the most recent feature value, we trace forward until the first significant change point is detected. The year corresponding

to this change point represents the predicted construction year for that specific point.

2.3.2. HRBAs detection based on thresholding

Since HRBAs are nearly irreversible in the span of 20 years, their occurrence in the time series of optical data is extracted by thresholding the window-based features constructed in the previous step. Specifically, we apply a threshold of 0, which has previously yielded ideal results in extracting impervious surfaces using CBI [25], to detect HRBAs. We calculate the point feature value (CBI) for each year at every point and arrange these values in chronological order. Starting from the feature value in the end, we trace forward through the time series until the first point with a value greater than the threshold is detected. The year corresponding to this change point, plus one, represents the predicted construction year for that specific point. It should be noted that the construction of HRBAs in China usually takes more than three months, and it is quite challenging to characterize this process at a fine scale using medium spatial resolution remote sensing data. Thus, we define the emergence of HRBAs as the point in time after which the HRBAs do not show distinct changes in their geometric structure in the image.

2.3.3. Window Statistic Replacement for Outliers

The CBI derived feature employed in the previous step still encounters limitations when identifying HRBAs, as HRBAs encompass built-up areas characterized by spatial connectivity, greenery, and bare soil. Conversely, the CBI derived feature primarily focuses on impervious surfaces. This discrepancy can lead to non-impervious surfaces, like greenery and bare soil, influencing the prediction outcomes for the construction year of HRBAs, despite the incorporation of a window-based method in constructing the CBI derived feature. Since it is not possible to identify in advance which predicted values are influenced by non-impervious surfaces, we propose an outlier detection method based on window statistical replacement to address this issue. Specifically, for each predicted value, we create a window of the same size as the one used for feature computation, centered on the pixel. The most frequent value within the window is used to replace the predicted value of the center pixel, while edge pixels are not processed further. The window size here refers to the size used for feature computation.

2.4. Acknowledgment

Two metrics are used to evaluate the proposed method. They are pixel-based accuracies and root mean square error (RMSE). The former reflects the consistency of the prediction results with the actual situation and assesses the overall performance of the model in different regions. RMSE as a complementary metric is to assess the accuracy of the model in the time dimension. RMSE quantifies the error between the predicted value and the actual value, and is particularly suitable for time-series forecasting to help analyze the accuracy of the model's estimation of the year in which the HRBAs are constructed. By combining pixel-based accuracy and RMSE, we are able to comprehensively assess the classification performance and error of the model.

The pixel-based accuracy evaluation method assesses the prediction accuracy by calculating the ratio of correctly predicted pixels to the total number of pixels. Specifically, correct pixels are those where the predicted value matches the true value, while the total number of pixels refers to all pixels within the HRBAs extraction results for 2020 on the slices. By dividing the number of correct pixels by the total number of pixels, the accuracy of the HRBAs predictions for the period 2000–2020 (within the same period) can be obtained.

The formula for RMSE is as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - T_i)^2}, \tag{5}$$

Where P_i is the predicted value, T_i is the true value, and n is the total number of pixels. The RMSE can intuitively reflect the accuracy of the model's predictions, with smaller values indicating that the predicted results are closer to the true values.

Furthermore, given that the Landsat-5/7/8 images used in this study have a resolution of 30 meters, the lack of spatial details may cause some uncertanties in precisely defining the changing point. Therefore, this study defines correct pixels as those with a one-year or two-year discrepancy between the predicted and true values to assess the accuracy of HRBAs construction years extraction. By doing so, the accuracies for one-year and two-year discrepancies are calculated separately to provide a more comprehensive reflection of the model's detection performance. The specific formulas are as follows:

$$N_{\mathrm{correct}(\Delta t)} = \sum_i 1(|\mathrm{pred}_i - \mathrm{true}_i| \le \Delta t), \tag{6}$$

$$\mathrm{Accuracy}(\Delta t) = \frac{N_{\mathrm{correct}(\Delta t)}}{N_{\mathrm{total}}}, \tag{7}$$

3. Data and experiments

3.1. Landsat data

Four cities including Zhengzhou, Jiaozuo, Xinyang, and Nanyang in Henan Province are selected in the study. The cities have different characteristics. In Zhengzhou the city center developed in early years is dominated by HRBAs, and as the city expands, more high-rise residential and commercial buildings are emerging in the newly developed peripheral areas. In Jiaozuo, the urban area is mainly composed of low- to mid-rise buildings, with high-rise buildings concentrated in commercial zones and some newly developed residential areas. Xinyang is characterized by abundant green belts and water bodies, with buildings primarily being low- to mid-rise. High-rise buildings in Xinyang are mainly located in the commercial center and newly developed residential areas, interspersed within the green environment. In Nanyang, high-rise buildings are mainly concentrated in the city's new urban areas and bustling commercial districts, contrasting with the ancient architecture. True color Imagery of the study areas are shown in Figure 2.

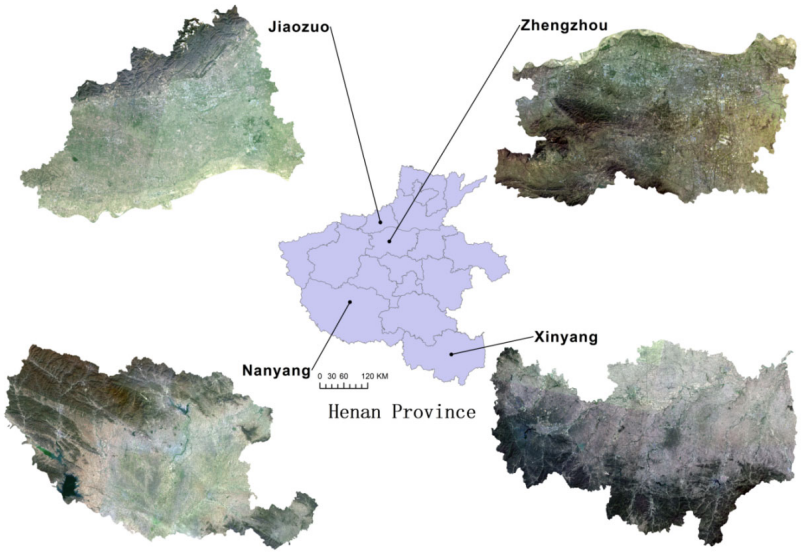


Figure 2. True color Imagery of study area.

Landsat 5, Landsat 7, and Landsat 8 ranging from 2000 to 2020 covering the study area are selected as data sources. To ensure the continuity and consistency of Landsat imagery across different generations, the USGS has standardized the data from Landsat-5, Landsat-7, and Landsat-8, including geometric and radiometric corrections. Six bands were selected from each satellite: red, green, blue, near-infrared, shortwave infrared 1, and shortwave infrared 2, all of which possess a spatial resolution of 30 meters. After the temporal composition, they were divided into tiles of size 128×128. Totally 305 ground truth tiles were created for Zhengzhou, 24 for Jiaozuo, 20 for Xinyang, and 18 for Nanyang. Each tile contains one or more blocks of HRBAs. The tiles for Zhengzhou covered the entire urban area, while the tiles for Jiaozuo, Xinyang, and Nanyang were randomly selected to represent characteristic areas. Slices in Zhengzhou were randomly split into two parts: 60% for training and 40% for testing data.

To improve the accuracy and efficiency of data processing, this study uses the HRBAs extraction results of the study area from 2020 as a spatial constraint. This constraint was obtained from Sentinel-2 data through deep learning method [18] and manually refined to gaurrentee its accuracy. This constraint allows for more precise identification of HRBAs spatial distribution, reducing unnecessary calculations and enhancing the reliability of subsequent analyses.

3.2. Experimental Design

We design a two stage process to validate the proposed method using samples in Zhengzhou, Jiaozuo, Xinyang, and Nanyang.

In the first stage, experiments were conducted in Zhengzhou to validate the proposed method. 60% of the slices from Zhengzhou were randomly selected as training data, and the remaining slices were reserved for testing. The optimal parameters, including window size, fixed threshold, and the post-processing measure of window statistics replacement, were determined using the training data from Zhengzhou and validated through the test data from the same region.

In the second stage, the validated method obtained from the experiments in Zhengzhou is applied to other cities, including Jiaozuo, Xinyang, and Nanyang. This cross-city validation aimed to examine the method's generality and applicability. Additionally, to verify the robustness of the key steps in the algorithm, multiple experiments were designed and conducted, considering variables such as window size and threshold settings. These experiments helped determine the optimal parameter configuration and evaluate the method's performance under different parameter settings.

4. Results and Analysis

The optimal parameters derived from the training samples in Zhengzhou include a window size of 15x15 and a fixed threshold of 0. With the parameters, the model is validated in the test data in Zhengzhou. The accuracy on the test data is shown in the first row of Table 3. The accuracy for predicting the exact construction year is low at 32.32%. However, when ±1 year error is allowed, the accuracy increases dramatically to 61.25%, and further reaches 80.82% within the ±2 year error. This change indicates that there is some delay in detecting the year of construction of HRBAs at shorter time scales, but as the time tolerance increases, the accuracy improves significantly, suggesting that the method is more stable at long-term time scales. Panel (a) in Figure 3 shows the scatter plot of predicted versus actual values for Zhengzhou. Although the RMSE is high at 3.48, indicating a large discrepancy between the predicted and real values, the overall trend agrees well with the real pattern.

Table 3. Comparison of accuracy for HRBAs construction years extraction in the study area.

City	Number of Patches	Accuracy(0)	Accuracy(1)	Accuracy(2)	RMSE
Zhengzhou	305	32.32%	61.25%	80.82%	3.48
Jiaozuo	24	49.72%	74.44%	87.31%	2.84
Xinyang	19	20.96%	46.43%	70.95%	3.53
Nanyang	18	28.01%	54.41%	66.95%	3.50

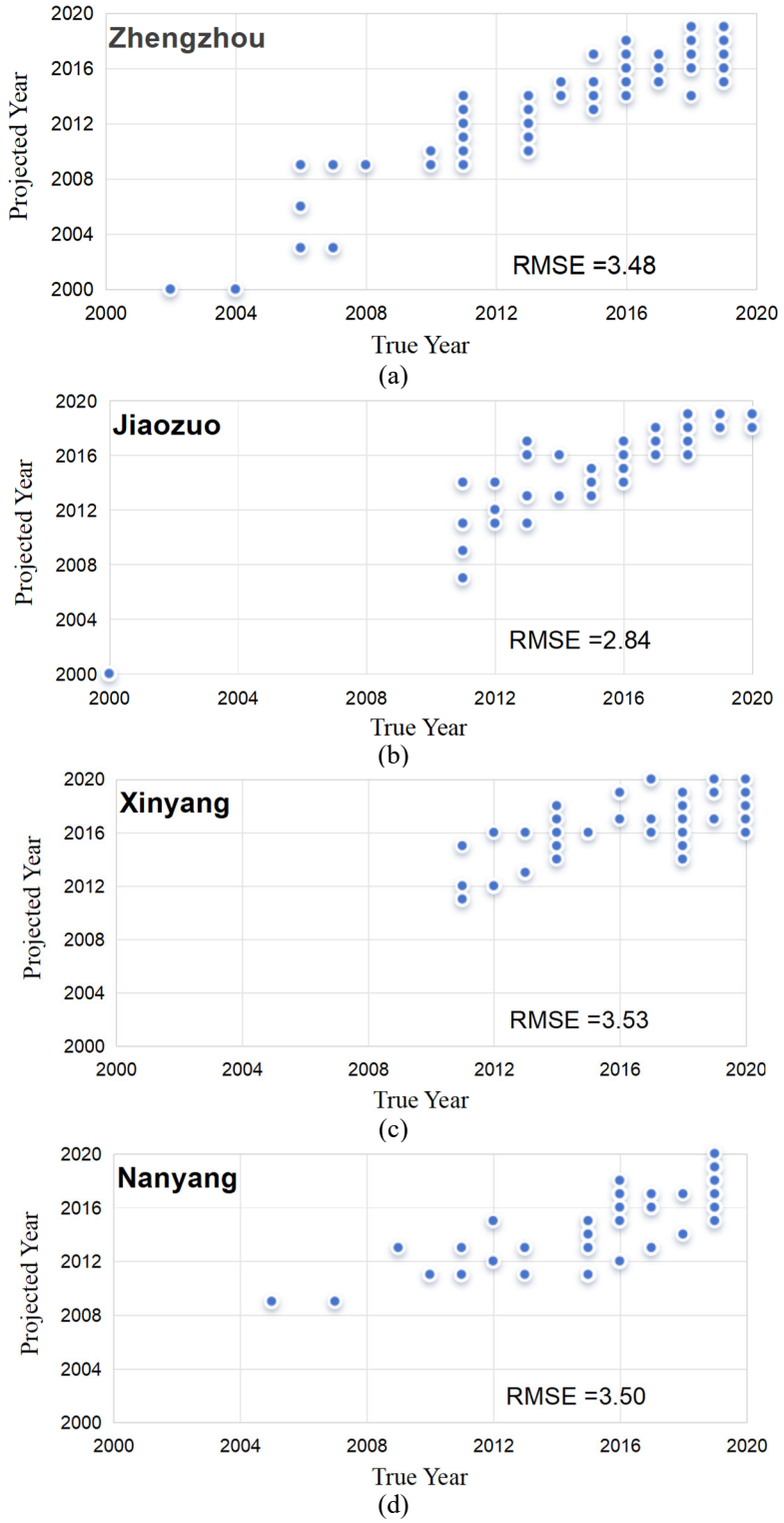


Figure 3. Scatter Plot of Predicted vs. True Values for Study Areas: (a) Zhengzhou, (b) Jiaozuo, (c) Xinyang, and (d) Nanyang.

To further demonstrate the effectiveness of the proposed method, Figure 4 illustrates the predicted construction years of HRBAs within the urban area of Zhengzhou, showing that our method clearly captures the spatial and temporal patterns of HRBAs between 2000 and 2020. In Figure 4, HRBAs originate from the city center and gradually expand to suburban areas, with newer constructions appearing progressively further out. The high land value and real estate appeal in the city center drive initial development here. As central development saturates and land costs rise, construction expands into suburban areas [26].

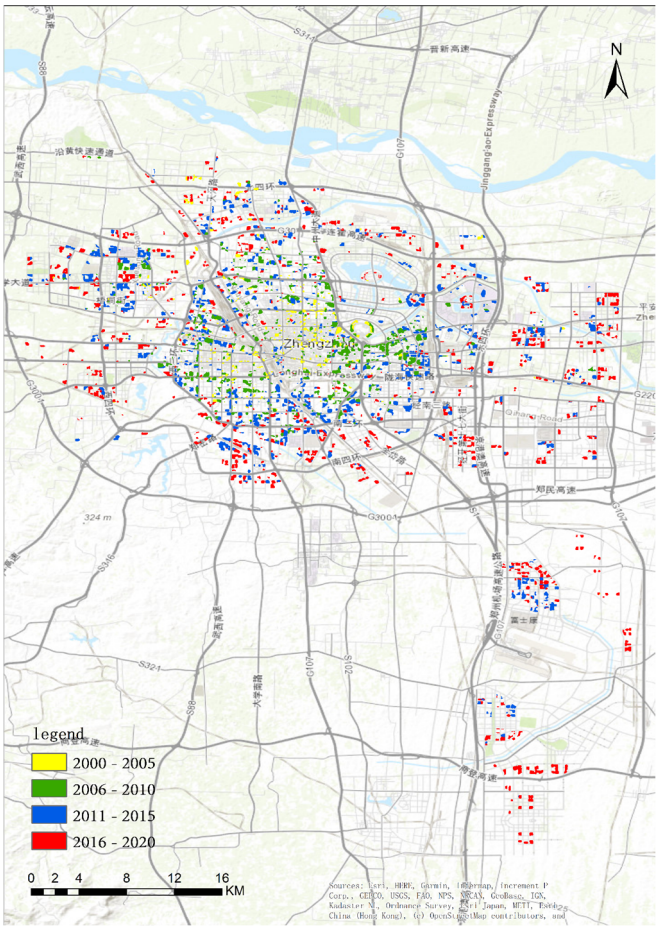


Figure 4. Map of HRBAs in Zhengzhou urban area from 2000 to 2020.

To validate the general applicability of the proposed method, we applied the parameters obtained from the Zhengzhou experiment to three other cities: Jiaozuo, Xinyang, and Nanyang, with the experimental results presented in rows 3 to 4 of Table 4. For Jiaozuo, the accuracy of the exact construction year is 49.72%, which is slightly higher than that of Zhengzhou. With a tolerance of ±1 year, the accuracy improves to 74.44% and further reaches 87.31% with a tolerance of ±2 years. This result suggests that despite the diverse topography of Jiaozuo, HRBAs are concentrated in flat areas, and thus topographic complexity has less influence on the temporal and spatial distribution of HRBAs construction extraction. The RMSE for Jiaozuo is 2.84, which is lower than that of Zhengzhou, indicating a smaller prediction error.

In contrast to Jiaozuo, where the prediction accuracy is relatively high, the prediction performance in Xinyang and Nanyang is lower, suggesting that local environmental factors may play a significant role in the variation of results. In Xinyang, the accuracy of precise construction year is lower at 20.96%, which improved to 46.43% with ±1 year tolerance and 70.95% with ±2 year tolerance. Its RMSE is 3.53. The accuracy of the precise construction year in Nanyang is 28.01%, 54.41% under ±1 year tolerance and 66.95% under ±2 year tolerance with an RMSE of 3.50. The results of Xinyang and Nanyang showed that the prediction errors were large and the overall accuracy could be improved. In particular, the high humidity environment in the southern cities may have affected the imaging quality, while the weakening of the feature expression by the dense vegetation and high water vapor made the detection accuracy of the construction years of HRBAs hard. Therefore, further model optimization or parameter adjustment is needed to improve the detection accuracy.

Overall, the accuracy of the precise construction year is low across cities, suggesting that the timeliness of construction years detection of HRBAs is challenging. However, as the allowed time tolerance increased, the accuracy improved significantly, with all cities achieving more than 65% accuracy within a ±2-year tolerance. This suggests that the overall trend is a valid representation of actual change despite some delayed detection. In addition, the accuracy differences among cities may be closely related to factors such as geographic environment, building distribution and data quality. The above results provide an important reference for further model optimization and parameter adjustment.

5. Discussion

5.1. The Impact of Window Size on Accuracy

The analysis presented in Table 4 evaluates the impact of window size on the accuracy of HRBAs construction year prediction in Zhengzhou. The results demonstrate a clear trend: as window size increases, the detection accuracy improves across all three accuracy metrics. Specifically, accuracy(0) reached its peak at 32.32% with a 15x15 window, accuracy(1) is highest at 61.85% with a 19x19 window, and accuracy(2) peaked at 80.98% with a 27x27 window. Concurrently, the RMSE value decreased from 6.93 to 3.33, highlighting the beneficial effect of larger window sizes on both accuracy and error reduction.

Table 4. Overview of experimental results on the impact of window size on HRBAs construction years extraction accuracy.

Calculation Method	Accuracy(0)	Accuracy(1)	Accuracy(2)	RMSE
Pixel-by-Pixel	19.96%	40.40%	56.24%	6.93
Window-by-Window (3x3)	25.03%	48.03%	64.45%	6.00
Window-by-Window (5x5)	27.78%	52.30%	69.51%	5.26
Window-by-Window (7x7)	29.78%	55.86%	74.19%	4.50
Window-by-Window (9x9)	31.09%	58.30%	77.31%	4.01
Window-by-Window (11x11)	31.98%	60.11%	79.40%	3.74
Window-by-Window (13x13)	32.24%	60.84%	80.40%	3.58
Window-by-Window (15x15)	32.32%	61.25%	80.82%	3.48
Window-by-Window (17x17)	31.96%	61.43%	80.95%	3.42
Window-by-Window (19x19)	31.85%	61.85%	80.97%	3.39
Window-by-Window (21x21)	31.62%	61.80%	80.92%	3.37
Window-by-Window (23x23)	31.46%	61.82%	80.92%	3.37
Window-by-Window (25x25)	31.23%	61.72%	80.97%	3.36
Window-by-Window (27x27)	30.86%	61.45%	80.98%	3.33
Window-by-Window (29x29)	30.37%	61.19%	80.85%	3.35
Window-by-Window (31x31)	30.05%	60.98%	80.66%	3.36

The 15x15 window size was chosen for its optimal balance between computational efficiency and the effective capture of spatial correlations. This size is large enough to capture meaningful spatial relationships between pixels without introducing excessive smoothing or losing image detail, which can occur with larger windows. While increasing the window size (e.g., 17x17 or 19x19) may slightly improve accuracy, it does not significantly reduce RMSE, and it comes at the cost of increased computational complexity and processing time. The 15x15 window, on the other hand, achieves 80.82% accuracy and the RMSE of 3.48, making it the most efficient choice for processing large-scale datasets. This moderate window size ensures efficient computation while maintaining spatial detail, thus providing the best overall performance in terms of both accuracy and resource consumption.

This improvement can be attributed to the window-based approach, which aggregates pixel values within each window, reducing the influence of individual outliers and mitigating noise. This approach enhances the stability and reliability of the results. However, the data also reveal that accuracy does not consistently improve with increasing window size. As window size expands, the accuracy gains diminish. This is due to the fact that larger windows may encompass multiple construction periods, where buildings constructed at different times influence one another, leading to a reduction in accuracy.

5.2. The Impact of Window Statistics Replacement on Accuracy

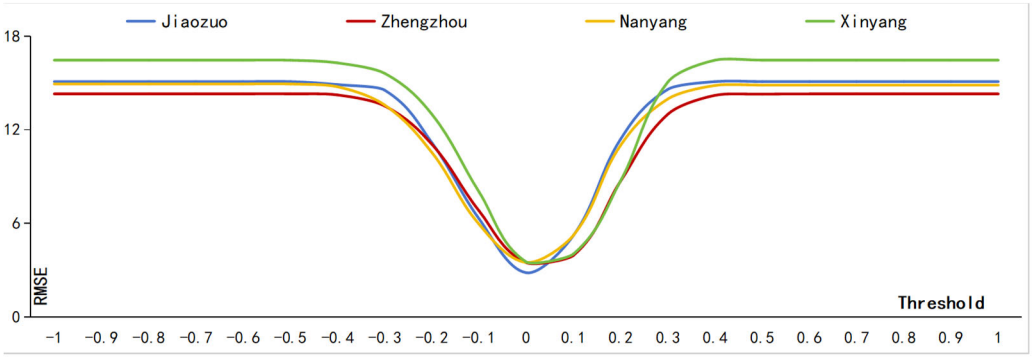
The impact of window statistics replacement on accuracy is shown in Table 5. The analysis reveals that the same-year accuracy increased by 1.19 percentage points, from 31.13% to 32.32%. The one-year tolerance accuracy improved by 1.17 percentage points, rising from 60.08% to 61.25%, while the two-year tolerance accuracy increased by 1.27 percentage points, from 79.55% to 80.82%. Additionally, The overall improvement is reflected in the reduction of RMSE from 3.63 to 3.48, highlighting the method's effectiveness in enhancing prediction stability. While the increase in accuracy is modest, the window-based statistical replacement proves valuable for improving the extraction of HRB construction years, by increasing the overall number of correct predictions and reducing errors.

Table 5. Overview of the impact of window statistical replacement on accuracy of HRBAs construction years extraction.

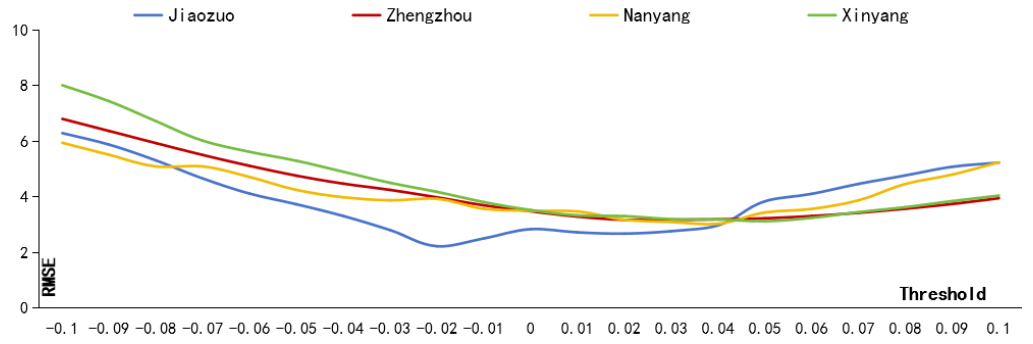
Processing Measures	Accuracy (0)	Accuracy (1)	Accuracy (2)	RMSE
None	31.13%	60.08%	79.55%	3.63
Window Statistics Replacement	32.32%	61.25%	80.82%	3.48

5.3. The Impact of Threshold value on Accuracy

To validate the impact of the setting of threshold, multiple experiments were conducted with multiple thresholds ranging from [-1, 1] in steps of 0.1. The results showed that threshold 0 performed best for the Zhengzhou sample. Therefore, threshold 0 was applied as a fixed threshold for subsequent experiments. In order to further explore the optimal thresholds for extracting the construction year of HRBAs in the four study areas of Zhengzhou, Jiaozuo, Xinyang, and Nanyang, this study conducts threshold exploration experiments using the full samples of the four study areas while keeping all other conditions constant. Figure 5 demonstrates the trend of RMSE with thresholds, highlighting the effect of different thresholds on the prediction error in each region. Table 6 further compares the RMSE and accuracies for each city at specific fixed thresholds, providing insight into the impact of threshold selection on prediction results.



(a)



(b)

Figure 5. Line Chart of RMSE Under Different Fixed Thresholds in the Study Area: (a) Impact of Threshold Range (-1 to 1) on RMSE, (b) Impact of Threshold Range (-0.1 to 0.1) on RMSE.

As shown in Figure 5(a), RMSE generally remains stable at first, then decreases, increases, and stabilizes again as the threshold increases from -1 to 1. Notably, RMSE decreases significantly near 0, suggesting that model prediction accuracy improves in this range. To further refine the analysis, Figure 5(b) limits the threshold range to between -0.1 and 0.1. Within this restricted range, particularly from -0.10 to -0.04, RMSE decreases for all study areas, indicating that lower thresholds may enhance the model's prediction accuracy. As the threshold approaches zero (e.g., -0.02 to 0.02), RMSE continues to decrease for most cities, though the magnitude of change diminishes. Notably, this indicates that a threshold of zero could be a scientifically justifiable choice as a prior selection. However, in the higher threshold range (e.g., 0.04 to 0.10), RMSE generally increases, indicating that higher thresholds may negatively affect model performance.

6. Conclusion

This study presents a computationally efficient backdating method for reconstructing HRBAs construction timelines using Landsat imagery (2000–2020). The method's simplicity stems from two innovations:

1. Threshold-centric logic: A fixed threshold of 0 serves as the baseline for HRB detection, reducing algorithmic complexity while achieving accuracies of 80.82% (Zhengzhou), 87.31% (Jiaozuo), 70.95% (Xinyang), and 66.95% (Nanyang) within ±2 years.

2. Window-based optimization: Dynamic threshold adjustments through localized window statistics streamline post-processing, boosting accuracies to 84.11–88.70% without requiring machine learning infrastructure.

Future work will focus on scaling this lightweight framework via Google Earth Engine, where its minimal data requirements (Landsat Tier 1 inputs) and model-agnostic design can enable nationwide HRBAs mapping within days—a task previously hindered by the computational burdens of deep learning approaches.

References

[1] Cheng, M.; Duan, C. The changing trends of internal migration and urbanization in China: new evidence from the seventh National Population Census. China Population and Development Studies 2021, 5, 275-295.(in Chinese)

[2] Zhu, J.; Li, L.; Cheng, G.; Gao, L.; Zhang, B. Detection and Analysis of High-rising Buildings within The sixth Ring Road of Beijing based on Sentinel-2 and Fully Convolutional Network. Remote Sensing Technology Application 2021, 36, 1436-1445.(in Chinese)

[3] Bahareh, M.; Casanovas-Rubio, M.d.M.; Antequera, A.d.l.F. Sustainability assessment in residential high-rise building design: state of the art. Architectural Engineering Design Management 2022, 18, 927-940.

[4] Ilgın, H.E. Space efficiency in tapered super-tall towers. Buildings 2023, 13, 2819.

[5] Yang, J.; Yang, Y.; Sun, D.; Jin, C.; Xiao, X. Influence of urban morphological characteristics on thermal environment. Sustainable Cities and Society 2021, 72, 103045.

[6] Guo, Y.; Li, X.; Luby, S.; Jiang, G. Vertical outbreak of COVID-19 in high-rise buildings: The role of sewer stacks and prevention measures. Current Opinion in Environmental Science & Health 2022, 29, 100379.

[7] Guo, G.Z.; Yu, Y.; Kwok, K.; Zhang, Y. Air pollutant dispersion around high-rise buildings due to roof emissions. Building and Environment 2022, 219, 109215.

[8] Mansouri, S.T.; Zarghami, E. Investigating the effect of the physical layout of the architecture of high-rise buildings, residential complexes, and urban heat islands. Energy and Built Environment 2025, 6, 1-17.

[9] Nugroho, N.Y.; Triyadi, S.; Wonorahardjo, S. Effect of high-rise buildings on the surrounding thermal environment. Building and Environment 2022, 207, 108393.

[10] Liu, Y. Analysis of the vertical forest of milan in terms of high-rise architecture and biodiversity. Highlights in Art and Design 2023, 3, 47-52.

[11] Zhang, Z.; Tang, W. Mixed landform with high-rise buildings: A spatial analysis integrating horizon-vertical dimension in natural-human urban systems. Land Use Policy 2023, 132, 106806.

[12] Yin, H.; Brandão Jr, A.; Buchner, J.; Helmers, D.; Iuliano, B.G.; Kimambo, N.E.; Lewińska, K.E.; Razenkova, E.; Rizayeva, A.; Rogova, N. Monitoring cropland abandonment with Landsat time series. Remote Sensing of Environment 2020, 246, 111873.

[13] Deng, C.; Zhu, Z. Continuous subpixel monitoring of urban impervious surface using Landsat time series. Remote Sensing of Environment 2020, 238, 110929.

[14] Bullock, E.L.; Woodcock, C.E.; Olofsson, P. Monitoring tropical forest degradation using spectral unmixing and Landsat time series analysis. Remote sensing of Environment 2020, 238, 110968.

[15] Potapov, P.; Li, X.; Hernandez-Serna, A.; Tyukavina, A.; Hansen, M.C.; Kommareddy, A.; Pickens, A.; Turubanova, S.; Tang, H.; Silva, C.E. Mapping global forest canopy height through integration of GEDI and Landsat data. Remote Sensing of Environment 2021, 253, 112165.

[16] Bekaert, D.P.; Handwerger, A.L.; Agram, P.; Kirschbaum, D.B. InSAR-based detection method for mapping and monitoring slow-moving landslides in remote regions with steep and mountainous terrain: An application to Nepal. Remote Sensing of Environment 2020, 249, 111983.

[17] Zhao, M.; Zhou, Y.; Li, X.; Cheng, W.; Zhou, C.; Ma, T.; Li, M.; Huang, K. Mapping urban dynamics (1992–2018) in Southeast Asia using consistent nighttime light data from DMSP and VIIRS. Remote Sensing of Environment 2020, 248, 111980.

[18] Pickens, A.H.; Hansen, M.C.; Hancher, M.; Stehman, S.V.; Tyukavina, A.; Potapov, P.; Marroquin, B.; Sherani, Z. Mapping and sampling to characterize global inland water dynamics from 1999 to 2018 with full Landsat time-series. Remote Sensing of Environment 2020, 243, 111792.

[19] Li, L.; Zhu, J.; Gao, L.; Cheng, G.; Zhang, B. Detecting and Analyzing the Increase of High-Rising Buildings to Monitor the Dynamic of the Xiong'an New Area. Sustainability 2020, 12, 4355.

[20] Yao, S.; Li, L.; Cheng, G.; Zhang, B. Analyzing long-term high-rise building areas changes using deep learning and multisource satellite images. Remote Sensing 2023, 15, 2427.

[21] Uhl, J.H.; Leyk, S. Towards a novel backdating strategy for creating built-up land time series data using contemporary spatial constraints. Remote Sensing of Environment 2020, 238, 111197.

[22] Skakun, S.; Wevers, J.; Brockmann, C.; Doxani, G.; Aleksandrov, M.; Batič, M.; Frantz, D.; Gascon, F.; Gómez-Chova, L.; Hagolle, O. Cloud Mask Intercomparison eXercise (CMIX): An evaluation of cloud masking algorithms for Landsat 8 and Sentinel-2. Remote Sensing of Environment 2022, 274, 112990.

[23] Hung, C.-L.J.; James, L.A.; Hodgson, M.E. An automated algorithm for mapping building impervious areas from airborne LiDAR point-cloud data for flood hydrology. GIScience & Remote Sensing 2018, 55, 793-816.

[24] Lin, Y.; Zhang, H.; Li, G.; Wang, T.; Wan, L.; Lin, H. Improving impervious surface extraction with shadow-based sparse representation from optical, SAR, and LiDAR data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 2019, 12, 2417-2428.

[25] Zhang, F.; Gao, Y. Composite extraction index to enhance impervious surface information in remotely sensed imagery. The Egyptian Journal of Remote Sensing and Space Science 2023, 26, 141-150.

[26] Cai, E.; Kou, Z.; Meng, K.; Zhang, Y.; Hou, H. Spatio-temporal characteristics of urban expansion in Zhengzhou from 1990 to 2020. Journal of Henan Agricultural University 2022, 56, 674-684.(in Chinese)