Science Bulletin 68 (2023) 2115–2124

Contents lists available at ScienceDirect

Science Bulletin

journal homepage: www.elsevier.com/locate/scib

Article

Regularity of rural settlement changes driven by rapid urbanization in North China over the three decades

Yansui Liu [a,b,*], Cong Ou [a,*], Yurui Li [a,b], Liqiang Zhang [c], Jianhua He [d]

[a] Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China
[b] College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100190, China
[c] State Key Laboratory of Remote Sensing Science, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China
[d] School of Resource & Environment Science, Wuhan University, Wuhan 430079, China

ARTICLE INFO

Article history:
Received 26 March 2023
Received in revised form 9 July 2023
Accepted 10 July 2023
Available online 4 August 2023

Keywords:
Rural settlement changes
Rapid urbanization
Cost-effective approach
Rural areal system
Beijing-Tianjin-Hebei region

ABSTRACT

The systematic decline of rural areas in the process of rapid urbanization has become a global trend, creating greater challenges for sustainable rural development. As the spatial projection of socio-economic development and living environment in rural areas, the continuous tracking of rural settlements (RUS) is crucial to quantify the imbalance of rural development. However, consistent information on RUS is highly needed but is quite deficient in current research. In this study, a cost-effective mapping model was proposed to produce an annual RUS dataset in the rapid urbanization region of Beijing-Tianjin-Hebei (BTH) in North China during 1990–2020, and the temporal-spatial regularity of RUS changes was further analyzed. The location-based and the area-based comparison verified the effectiveness of our model, with a mean overall accuracy of 85% and a mean correlation value of 0.88, respectively. The total area of RUS in the BTH region increased by 2561 km^2 from 1990 to 2020, while the average size of RUS remained stable after 2005. The annual change trends in RUS appeared with increasing and decreasing accounting for 76.33% and 23.67%, respectively. The centroids of RUS in Tianjin and Hebei have moved closer to Beijing, while those in Beijing have moved away from the former. Notably, we have identified 56.3% counties in the BTH region belong to the "Convex-I" change type in RUS. In general, our work can help to consistently quantify the spatiotemporal patterns of RUS in a cost-effective way, providing more explicit spatial information and continuous temporal information for rural residential land management.

1. Introduction

Rural settlements (RUS) are the main spatial form and the concentrated expression of the cooperative situation of residence as well as the industry in the rural areal system [1], whose spontaneous organization and formation mechanisms reflect the economic development [2], biodiversity conservation [3], and cultural transmission in rural areas [4]. Increasing RUS development with the declining rural population has occurred in many developing countries due to accelerated urbanization [5,6], which may accelerate the process of rural hollowing [7], development weakening [8], and the loss of ecosystem services in the rural areal system [9], creating great barriers to achieve the sustainable development goals (e.g., SDG 1, 2, and 11) promoted by United Nations. Therefore, it is of great significance to quantify and characterize the spatiotemporal patterns of RUS for promoting the implementation of the rural revitalization strategy and achieving sustainable rural development.

Previous literature on RUS focused on a range of aspects, from the dependence of RUS on the natural environment [10,11], the growth of RUS associated with social-economic development [12], and changes in urban–rural relations and sustainable approaches to RUS [13], to the landscape pattern [14], hierarchical classification [15], and differentiation characteristics [16]. Despite various datasets that included censuses, satellite imagery, and land use maps that have been used to capture the dynamics of RUS in the above-mentioned studies, some limitations still remained. First, from the perspective of spatial explicitness of RUS, the quantitative capability of census data from the statistical yearbook or household survey is insufficient due to the coarse resolution. Second, in view of the temporal consistency of RUS, either manually interpreted satellite imagery or existing land use maps are unable

* Corresponding authors.
E-mail addresses: liuys@igsnrr.ac.cn (Y. Liu), oucong@igsnrr.ac.cn (C. Ou).

https://doi.org/10.1016/j.scib.2023.08.006

to meet the need for continuous tracking of RUS due to laborious human intervention. These limitations pose a significant obstacle to understanding the characteristics and mutations of RUS dynamics that vary with the encryption of spatial and temporal scales. Meanwhile, with the rapid but uneven urbanization and the integrated urban–rural development, rural reconstruction and transformation development caused great attention in rural geography, land management, and regional planning communities [17–19]. Accurate information on the dynamics of RUS is highly needed but is quite deficient, which restricts a consistent process analysis of the spatiotemporal patterns of RUS and provides limited scientific references for rural spatial planning and related policy making.

Considering the related censuses at a coarse spatial resolution (usually at the county-level) and the related field survey with a low temporal frequency (usually with a 5- or 10-years interval) [20], an automatic mapping model based on satellite images with temporal and spatial consistency for RUS is critical to reduce human interference and enhance the objectivity of the produced dataset, which can deliver more cost-effective service than traditional censusing or field surveying [21]. Although it is possible to delineate residential habitats for human beings within a clear spatial extent (i.e., with a specific boundary) thanks to the development of Earth observation technology [22], existing works mainly attempted to map urban settlements through multi-sources remote sensing imagery [23–27], while an emphasized focus on rural areas was rare [28,29]. For example, the Global Human Settlement Layer (GHSL) database presents the application of the degree of urbanization via a logic of cell clusters' population size, population and built-up area densities [30]. The global annual impervious area (GAIA) provides a paradigm for mapping annual urban extents dynamics during 1985–2018, using the state-of-art cloud computing platform and the full archive of Landsat images [31]. Also, there are several urban settlement products that have been developed in specific countries and even globally [22,24,32]. After all, compared to consistent and concentrated urban settlements, the spatial distribution of rural settlements is more dispersed and fragmented, which leads to a great challenge in delineating RUS boundaries from fine-resolution and consistent processes, especially within a large region over a long-term period. Consequently, it is urgent to develop a cost-effective approach to mapping RUS at a fine-resolution and produce a long-term dataset of RUS at a continuous time scale.

Therefore, to better understand the spatiotemporal regularity of rural settlement changes from a consistent process analysis perspective, the objective of this study is two-fold: (1) to propose a spatial explicit mapping model for RUS that can capture their dynamics annually and produce the annual RUS dataset in the rapid urbanization region of Beijing-Tianjin-Hebei (BTH) in North China during 1990–2020, and (2) to investigate the temporal and spatial dynamics of RUS at multi-scales over the three decades in BTH region. To our limited knowledge, this study provides the first cost-effective and objective way to quantify and characterize the spatiotemporal regularity of RUS changes.

2. Materials and methods

2.1. Study area

The Beijing-Tianjin-Hebei (BTH) region, located in the northern part of the North China Plain (36°5′–42°37′N, 113°11′–119°45′E), was selected as the case study region (Fig. 1). This region contains two municipalities (Beijing and Tianjin) and Hebei Province, with a total area of approximate 21.8 × 10^4 km^2. The geomorphology of this region includes plains, mountains, hills, and a few plateaus

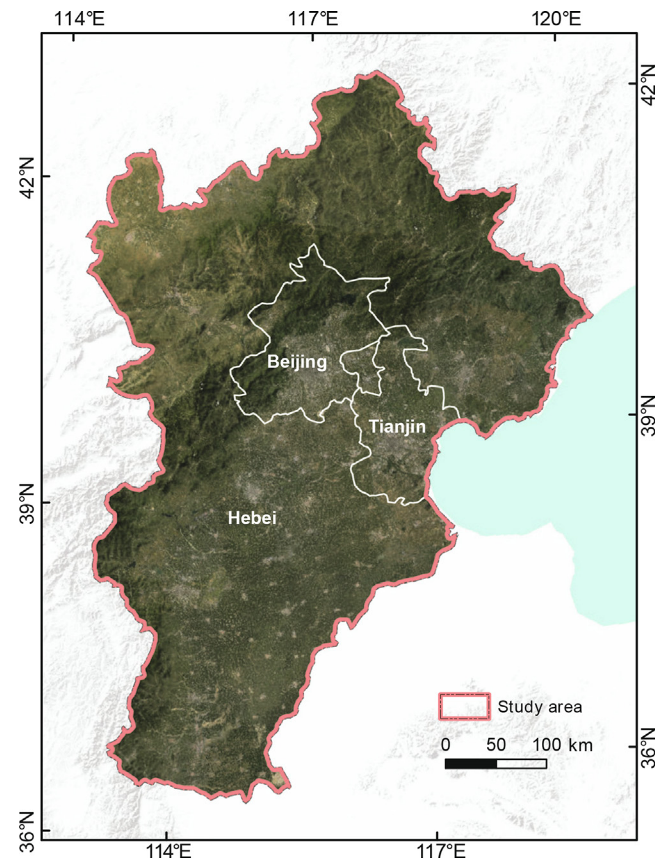


Fig. 1. The location of the BTH region, China.

and the terrain of the northwest is high, whereas that of the southeast is low. It belongs to a semi-humid and semi-arid continental monsoon climate zone with an annual average temperature, ranging from 10.4 to 11.9 °C and an annual average precipitation, ranging from 375.5 to 684.7 mm. As the economic development pole of northern China, the BTH region has two metropolises with a population of over 10 million, Beijing and Tianjin, as well as the less developed mountainous areas in northern and western Hebei and the traditional agricultural production areas in central Hebei. For a long time, the promotion of the BTH cooperative development strategy highly valued by the central government has been restricted due to the uncoordinated urban–rural development and obvious differences in rural development in this region. Therefore, the selection of this region is of representative significance for understanding the spatiotemporal regularity of rural settlement changes from a consistent process analysis perspective.

2.2. Data sources

Annual mapping of RUS requires input data that are continuous in the temporal dimension, completed coverage in the spatial dimension, and accessible from public sources. Therefore, the Landsat-derived annual China land cover dataset (CLCD) from 1990 to 2020 produced by Wuhan University was used in this study to obtain the original impervious surface area (ISA) [33]. The overall accuracy of CLCD achieved 79.31% and a mean F_1 score of the impervious area over 72% via visually interpreted independent samples, which exhibited a better and more stable accuracy concerning the existing annual LC products such as MCD12Q1 and ESACCI_LC [34,35]. Besides, a prolonged artificial Nighttime-light (NTL) dataset of China during 1984–2020 based on the Long Short-Term Memory (LSTM) network with an average root mean square error (RMSE) of 0.73 and the coefficient of determination (R^2) of 0.95 [36], which is an important indicator of human activities including socio-economic and energy consumption, was used in this study to distinguish RUS from urban extent. And the Point of Interest (POI) data within the built-up area of each city in the BTH region was also used in this study (download via https://www.resdc.cn/data.aspx?DATAID=302) to sample NTL value. In addition, the remote sensing-based dataset of land use land cover change over multiple periods in China (CNLUCC), which provided the spatial distribution of rural settlements since the 1980s with a five-year interval [37], was used to compare with the RUS dataset we mapped.

2.3. Mapping model for RUS

The RUS mapping model proposed in this study on the basis of referring to similar studies consisted of three basic steps [24]. First, a kernel density derived ISA (ISA-KD) map was calculated by using the kernel density estimation (KDE) method from 100 m aggregated ISA data, which reflected the neighboring density of impervious surfaces at an upscaling pixel level. Second, we performed the cellular automata (CA) based approach for the ISA data, combined the ISA-KD map and then the delineation of the initial boundary for each patch was completed after morphological processing, in which most line-type objects were eliminated and the holes within the target patches were refilled. Ultimately, we improved the final RUS patches by applying a double filter based on the dynamic thresholds of NTL as well as an object-based cluster derived patch size, which differentiated the rural patches from urban and other patches in view of the spatial adjacencies and human activities. Besides, it should be pointed out that all the steps of the proposed mapping model were designed and implemented based on the Google Earth Engine (GEE) platform [38], which ensures the extensibility of this model. Fig. 2 illustrates the general technical flow of the RUS mapping model, and the paragraphs that follow provide a detailed description of each step.

2.3.1. Kernel density estimation

KDE, a non-parametric statistical method for generating probability densities [39], has been widely used in describing the spatial heterogeneity of many geographical phenomena [40,41]. In this step, we first aggregated the ISA data at the spatial resolution of 30 m into the upscaled ISA data in terms of the percentage of ISA areas within the 100 m grid. Then, the KDE method was implemented for each upscaled ISA pixel to generate the ISA-KD map, in which the kernel density of each pixel that represents the accumulative influence of its neighborhood was calculated.

2.3.2. Initial boundary delineation

Combined with the derived ISA-KD map that provided an approximate extent for patch boundary delineation based on describing the spatial heterogeneity of the ISA data, some Non-ISA pixels of which (due to the coarse resolution) within areas with high kernel densities (KD value > 80%) were identified and should be refilled to avoid the fragmentation of boundary delineation. As a "bottom-up" model to simulate a complex geographical phenomenon, the CA-based approach can determine the state conversion of a cell under the specific transition rule [42], which is suitable for our needs at this step. Therefore, the CA-based approach was adopted in this step to refill those Non-ISA pixels with high KD values by using a 7 × 7 Moore neighborhood. Given that the implementation of CA only produced a more homogeneous extent of ISA compared with the initial ISA data, we performed two morphological approaches before the boundary delineation, including dilation and erosion. Among them, the dilation processing was used to convert Non-ISA pixels around areas with high kernel densities (which may be missed in the CA implementation) into ISA pixels, and some line-type objects (e.g., roads, rivers, or drains) were removed by the erosion processing. Finally, the delineation of the initial boundary for each patch was completed and converted into a vector-based format.

2.3.3. Post processing

In the third step of the RUS mapping model, the final RUS patches were derived after the double filter processing, including an area-based filter and an NTL-based filter. In terms of the area-based filter, we clustered the initial patches to obtain their sizes by using an object-based approach based on their spatial adjacency [43]. And large clustered patches (> 10 km^2) that were recognized as "urbanized patches" were further filtered, in which the threshold was determined on the minimum area of county-level cities in relevant statistics. In terms of the NTL-based filter, we first calculated the mode values of NTL within 200 urban POI samples from 1990 to 2020, then treated the mode value in each year as the dynamic threshold to mask areas higher than it in the corresponding year. The NTL-based filter was proposed by referring to the use of NTL data in some urban studies [44,45], and adopting the idea of the reverse mask to better screen out some small urbanized patches that may be missed in the area-based filter.

2.3.4. Technical validation

In order to evaluate the validity of the proposed mapping model and the accuracy of the resultant product, a location-based and an area-based comparison were both carried out with the CNLUCC dataset in terms of the spatial details of RUS and their overall patterns. As for the location-based comparison, we randomly collected ~10,000 test points within the extent of rural settlements of CNLUCC for seven years (i.e., 1990, 1995, 2000, 2005, 2010, 2015, 2020), and then the overall accuracy (OA) of our product for each representative year was calculated. As for the area-based comparison, a cross-scale statistical method that included three different aggregated grids (i.e., 10 km × 10 km, 30 km × 30 km, and 50 km × 50 km) was adopted to obtain the RUS areas at different scales, and the scatter plot and the linear regression with the quantitative metrics of the correlation coefficient were used to verify the agreement of the total area of RUS between our product and the CNLUCC.

2.4. Spatiotemporal analyses of RUS

2.4.1. Piecewise linear regression

The piecewise linear regression (PLR) method was implemented to detect trends of the annual changes in the total area of RUS over the period of 1990–2020 in the study area. The PLR method, which is also called segmented linear regression [46], is a special case of switching regression whereby the independent variable is segmented and the regression analysis is performed separately for these segments [47]. This method determines the segmentation reasonableness (detected break-points between the segments) by using a significance test to reduce the subjective error caused by visual inspection, which can be expressed as

$$\begin{aligned} &Y = a_1 + b_1X,\ X \le J_1, \\ &Y = a_2 + b_2X, J_1 \le X \le J_2, \\ &\vdots \\ &Y = a_n + b_nX, X \ge J_n, \end{aligned} \tag{1}$$

where X and Y are the independent and dependent variables, respectively. a_1 to a_n and b_1 to b_n are the intercepts and slopes of the linear segments, respectively. And J_1 to J_n are the break-points between the linear segments. Notably, two statistics, the coefficient

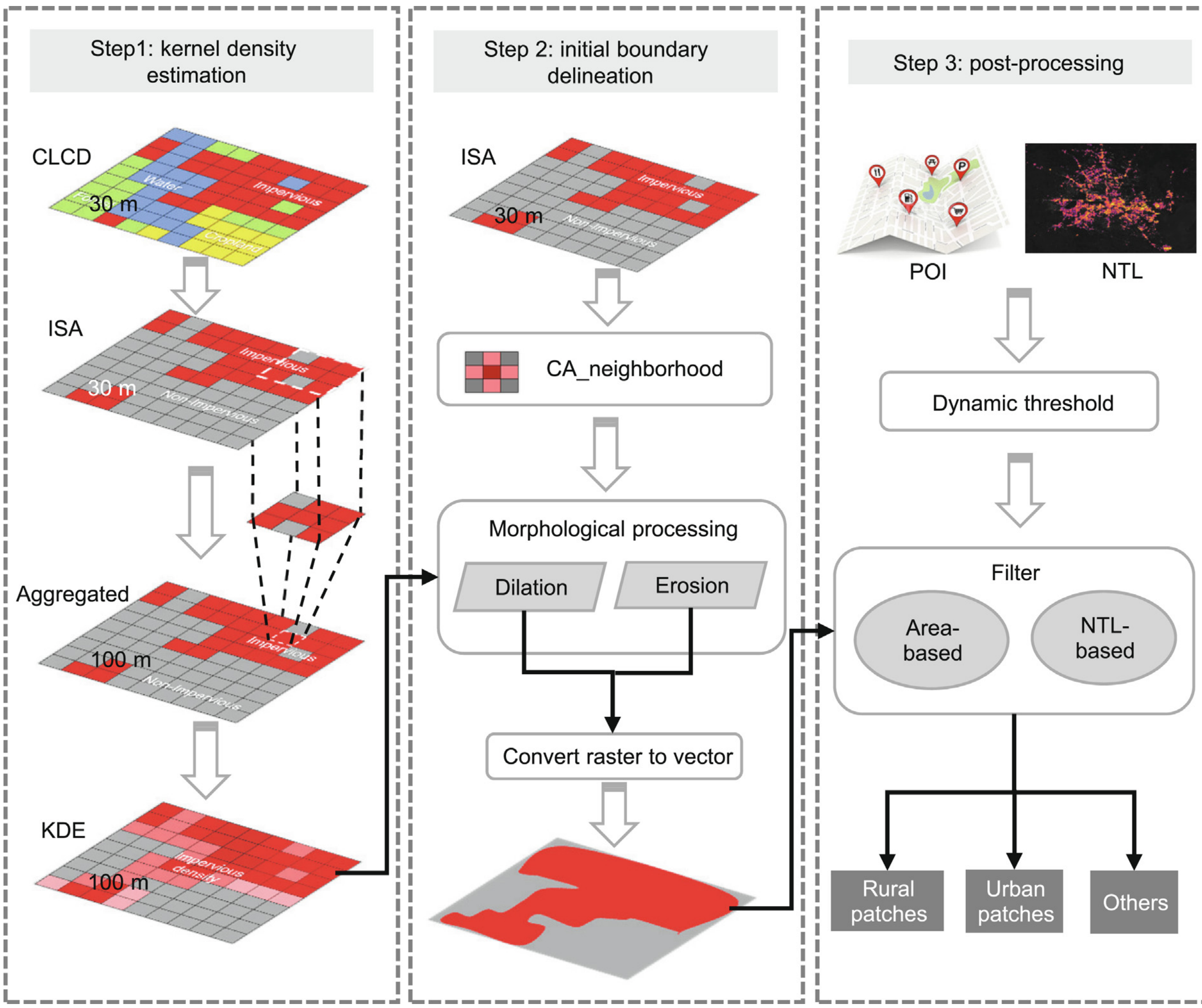


Fig. 2. Illustration of the general technical flow of the RUS mapping model.

of determination (R^2) and the estimate of error variance (σ^2), are used to reflect the performance of the regression analysis. In other words, the break-point between each two interval segments with the largest R^2 or the smallest σ^2 will be selected. Here the PLR method was performed using the segmented R package [48]. If Davies' test was significant, a segmented regression specifying the best break-point was determined on the time-series data.

2.4.2. Identification of change types

Considering land use transition is a mirror of socio-economic development, especially in rural areas, the change of rural residential land can directly reflect the evolution and transformation process of the rural territorial system. Therefore, we first aggregated the annual RUS maps into their fractional area maps within a reduced grid of 1 km $\times$ 1 km, and a spatially explicit map of the linear trend of RUS areas change was calculated based on the following formula [49]:

$$slope = \frac{n \times \sum_{i=1}^{n}(i \times A_i) - \sum_{i=1}^{n} i \times \sum_{i=1}^{n} A_i}{n \times \sum_{i=1}^{n} i^2 - \left(\sum_{i=1}^{n} i\right)^2}, \quad (2)$$

where *slope* is the change rate of RUS areas; i represents the ith year from 1990 to 2020, and n is the number of years; A_i is the area percentage of RUS in the ith year. If the *slope* of the reduced grid is greater than 0, which RUS area increased. If *slope* < 0, then RUS areas are decreased within this grid. Then, we identified different change types of the total area in RUS at the county-level based on time series analysis. Specifically, combined with the linear trend of RUS areas change, the deviation value between the change curve and linear regression line was calculated by the following equation:

$$V = \frac{\sum_{i=1}^{n} A_i}{n} - \frac{A_n - A_1}{2}, \quad (3)$$

where V represents the deviation value between the change curve and linear regression line, i represents the ith year, n is the number of years, and A_i is the total area of RUS in the ith year. In our case, four change types of RUS areas for all counties of the BTH region were identified by the following criteria (Fig. S1 online): convex increased (Convex-I, $slope > 0 \cap V > 0$); concave increased (Concave-I, $slope > 0 \cap V < 0$); convex decreased (Convex-D, $slope < 0 \cap V > 0$); concave decreased (Concave-D, $slope < 0 \cap V < 0$).

3. Results

3.1. Performances of the mapping model

Based on the proposed mapping model, we generated the annual RUS dataset from 1990 to 2020 in the BTH region. Fig. 3 illustrates an example (and its enlarged views in Fig. 3e) of how to generate the final results from the initial ISA distribution. Compared with the initial ISA data (Fig. 3a), the step 1 derived results provided more information about the heterogeneity of impervious

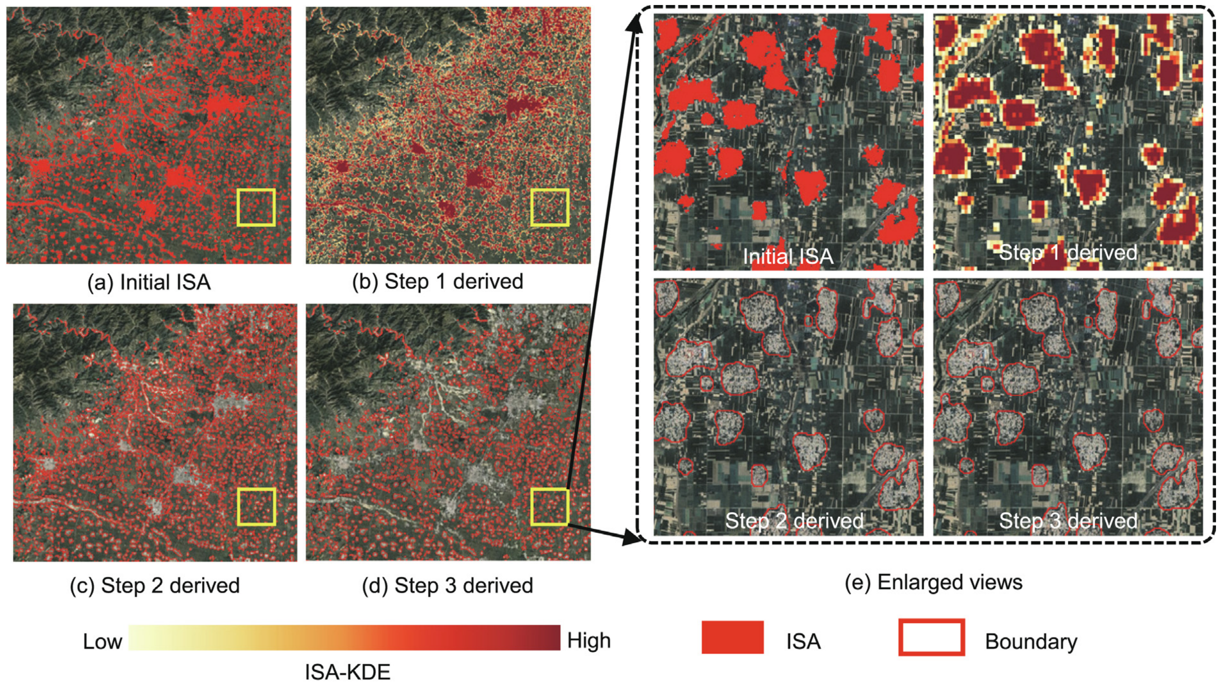


Fig. 3. Illustration of the RUS mapping model in each step: (a) the initial ISA distribution of example in 2020, (b) the step 1 derived results, (c) the step 2 derived results, (d) the step 3 derived results, and enlarged views from (a)–(d). Each subplot is overlaid with Google Earth (Basemap data © 2020 Google).

areas by upscaling the spatial resolution of initial ISA data from 30 to 100 m based on the KDE method (Fig. 3b). After the CA and morphological processing in step 2 (Fig. 3c), we then derived the initial boundary delineation results, in which several unrelated impervious patches (e.g., roads and drains) were effectively eliminated, and some "holes" (e.g., water bodies and green spaces) within the related patches were refilled. Finally, by implementing the double filter that integrated with dynamic thresholds of NTL as well as an object-based cluster in step 3 (Fig. 3d), we identified the extent of RUS from the initial ISA data of the example. Overall, the aforementioned example helps to portray the procedure of the proposed mapping model and its outcomes from a microscopic view.

In terms of temporal dynamics of the RUS dataset, three examples in Beijing, Tianjin, and Hebei are composited by corresponding Google Earth/Landsat images using a natural color (in red–green–blue bands) in the representative years 2020 (Fig. S2 online), 2010 (Fig. S3 online), 2000 (Fig. S4 online), and 1990 (Fig. S5 online). Generally, the overall pattern and spatial details of RUS match well with the actual extent of RUS on the satellite image as well as the CNLUCC in each example, proving the mapping model proposed in this study enables the identification of the rural settlement extent in different areas within different periods. Besides, it can be found that several isolated patches surrounding the urban area were annexed within the process of urbanization, which are the urban–rural fringe that can be characterized by low-income groups with limited amenities like retail, warehousing, or residence [50]. These are human settlements that, although in a blurred zone, are functionally oriented to the urban (so-called "urban villages") and also can be well excluded from RUS in this study. Given the location-based comparison, we also randomly generated ~ 10,000 test points from the CNLUCC in each representative year and the mean overall accuracy of RUS over the past 30 years achieved 84.91%. Considering that the CNLUCC was produced manually, and our results were automatically produced from long-term ISA data that reduced intensive human labor and subjectivity, which means that a little difference in random location-based comparison is acceptable.

Furthermore, a cross-scale statistical method was also adopted to assess the performance of the proposed mapping model in view of area-based comparison. In general, scatter points between RUS and CNLUCC are distributed around the 1:1 line over years and cross scales (Fig. S6 online). Specifically, the mean correlation value that crosses all scales in each representative year was greater than 0.90 (Fig. S6a online), 0.92 (Fig. S6b online), 0.88 (Fig. S6c online), and 0.82 (Fig. S6d online), respectively. Therefore, the total area of RUS derived from our proposed mapping model showed good consistency with the CNLUCC in view of cross-scale comparison. Given the offset between the 1:1 line and the linear fitting line, we found the reduced area of RUS across all scales was slightly higher than those from the CNLUCC, which is likely due to different definitions and approaches used in interpreting the satellite image. Notably, the correlation between RUS and CNLUCC appears a declining trend from 2000 to 2020 in all scales, indicating the distinction between urban and rural areas becomes more complicated due to rapid urbanization.

3.2. Temporal characteristics of RUS

To present the annual changes of RUS over the past three decades, the total area of RUS in the entire BTH region as well as three administrative zones was calculated for each year and the breakpoints for their change trends were detected based on the PLR method (Fig. 4 and Table S1 online), in which the best breakpoint between each two interval segments during 1990–2020 with the largest R^2 or the smallest σ^2 was selected. From the perspective of the entire BTH region, the total area of RUS increased sharply from 13,025 km^2 in 1990 to 14,804 km^2 in 1997 (Fig. 4a), with an annual change rate of 0.258 $\times$ 10^3 km^2/a (P < 0.001), which is nearly 8 times the rate (0.034 $\times$ 10^3 km^2/a, P < 0.001) during 1997–2020 (from 14,804 km^2 to 15,586 km^2). From the perspective of different administrative zones, the annual changes in the total

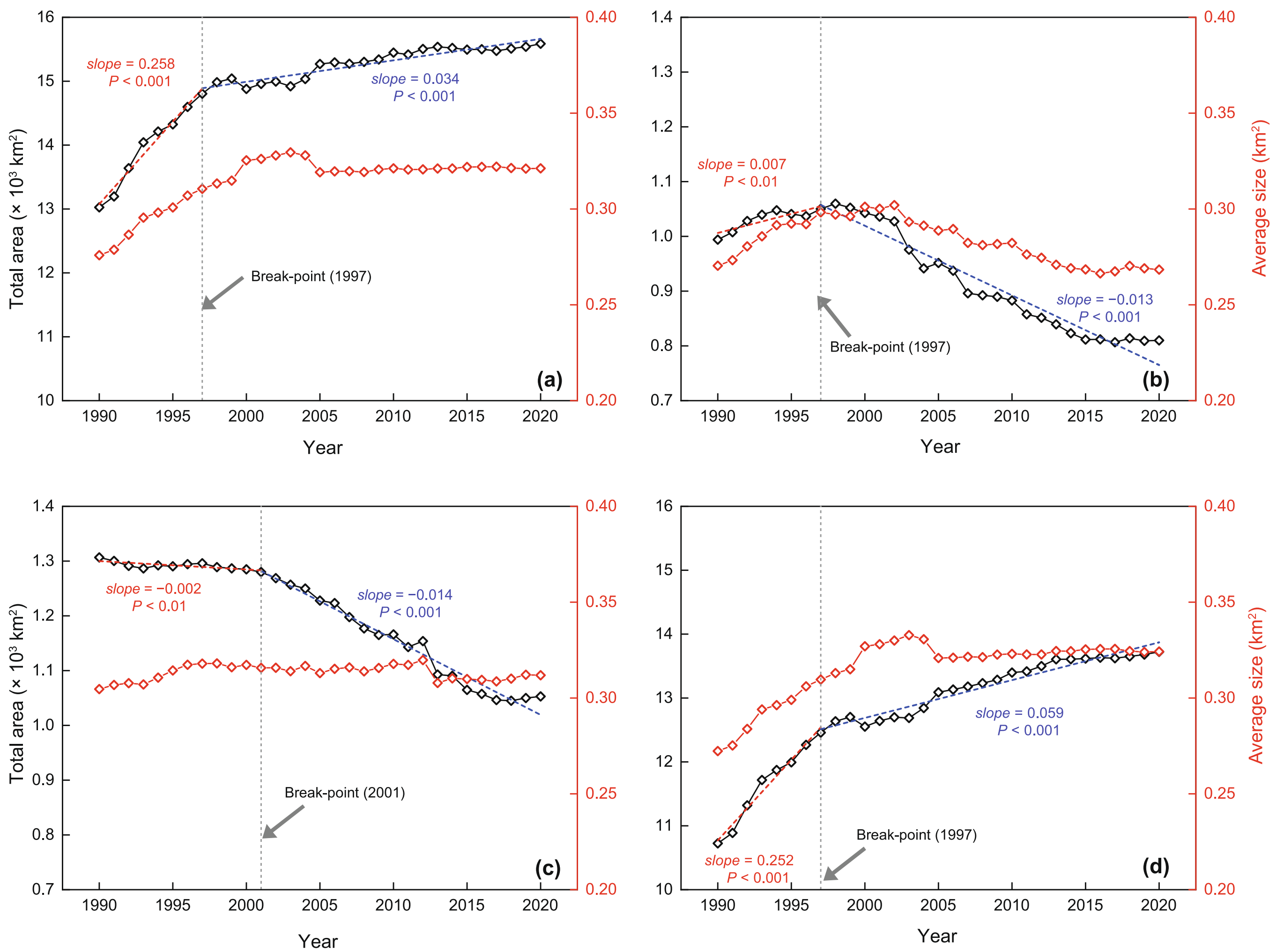


Fig. 4. The annual changes for RUS in the total area and the average size during 1990–2020 in the entire BTH region (a), Beijing (b), Tianjin (c), and Hebei (d).

area of RUS differed in the near-metropolitan area, the plain area, and the mountainous area. As shown in Fig. 4b, the total areas of RUS in Beijing fluctuated with the annual change rate of 0.007×10^3 km^2/a (P < 0.001) during 1990–1997, then declined considerably from 1050 km^2 in 1997 to 810 km^2 in 2020 with the annual change rate of -0.013×10^3 km^2/a (P < 0.001). Although the change trend of RUS areas in Tianjin was slightly different from that of Beijing before the break-point (2001), the total area of RUS in Tianjin also had a substantial decrease of ~ 18% from 1280 km^2 in 2001, to 1053 km^2 in 2020 (Fig. 4c). In contrast, the total area of RUS in Hebei (Fig. 4d) increased significantly from 1990 to 1997 with the annual change rate of 0.252×10^3 km^2/a (P < 0.001), then moderately increased after the break-points with the annual change rate of 0.059×10^3 km^2/a (P < 0.001).

To explore temporal changes in RUS patches during 1990–2020 in the BTH region, we also calculated the average size and the standard deviation of RUS patches for each year. As shown in Fig. S7a (online), the average size of RUS patches in the BTH region gradually increased from 0.28 km^2 in 1990, reached the maximum size (about 0.33 km^2) in 2003, then maintained around 0.32 km^2 after 2005. Meanwhile, the standard deviation of RUS patches decelerated increased from 0.32 in 1990 to 0.39 in 2005, then gradually increased from 2005 to 2020. It is important to point out that although the change in the average size of RUS patches has gradually stabilized, the variations in the size of RUS patches have gradually widened in the BTH region over the past 30 years. Furthermore, by identifying spatially disconnected RUS patches using an object-based method, each RUS patch was sorted by its size and given a unique ranked identification. As shown in Fig. S7b (online), the size of RUS patches exhibited an exponential decay trend with the increase in the number of RUS patches. The enlarged view of curves in Fig. S7b (online) for different years further showed that the turning-points in these trends occurred less than 4 km^2 before 2000, while occurred greater than 4 km^2 after 2000. On the other hand, a notable expansion of the size of RUS patches can be observed over the past 30 years in the BTH region. In particular, the total number of large RUS patches (greater than 1 km^2) in 1990 is 1523, which is about doubled in 2020 (3196).

3.3. Spatial variations of RUS

The spatially explicit map that was classified into "Increased (*slope* greater than 0, P < 0.05)" as well as "Decreased (*slope* < 0, P < 0.05)" was calculated based on the slope of linear trend. In view of the entire BTH region (Fig. 5a), the grids with the increasing trend in the RUS area accounted for 76.33% and a 23.67% decrease of all the reduced grids. In view of different administrative zones, the proportions of increased and decreased grids in Beijing were 58.13% and 41.87%, respectively, and it can be found that contiguous decreased grids were mainly distributed around the central urban area of Beijing (Fig. 5e). As for Tianjin, the grids with the increasing trend in the RUS area accounted for 58.33%, of which

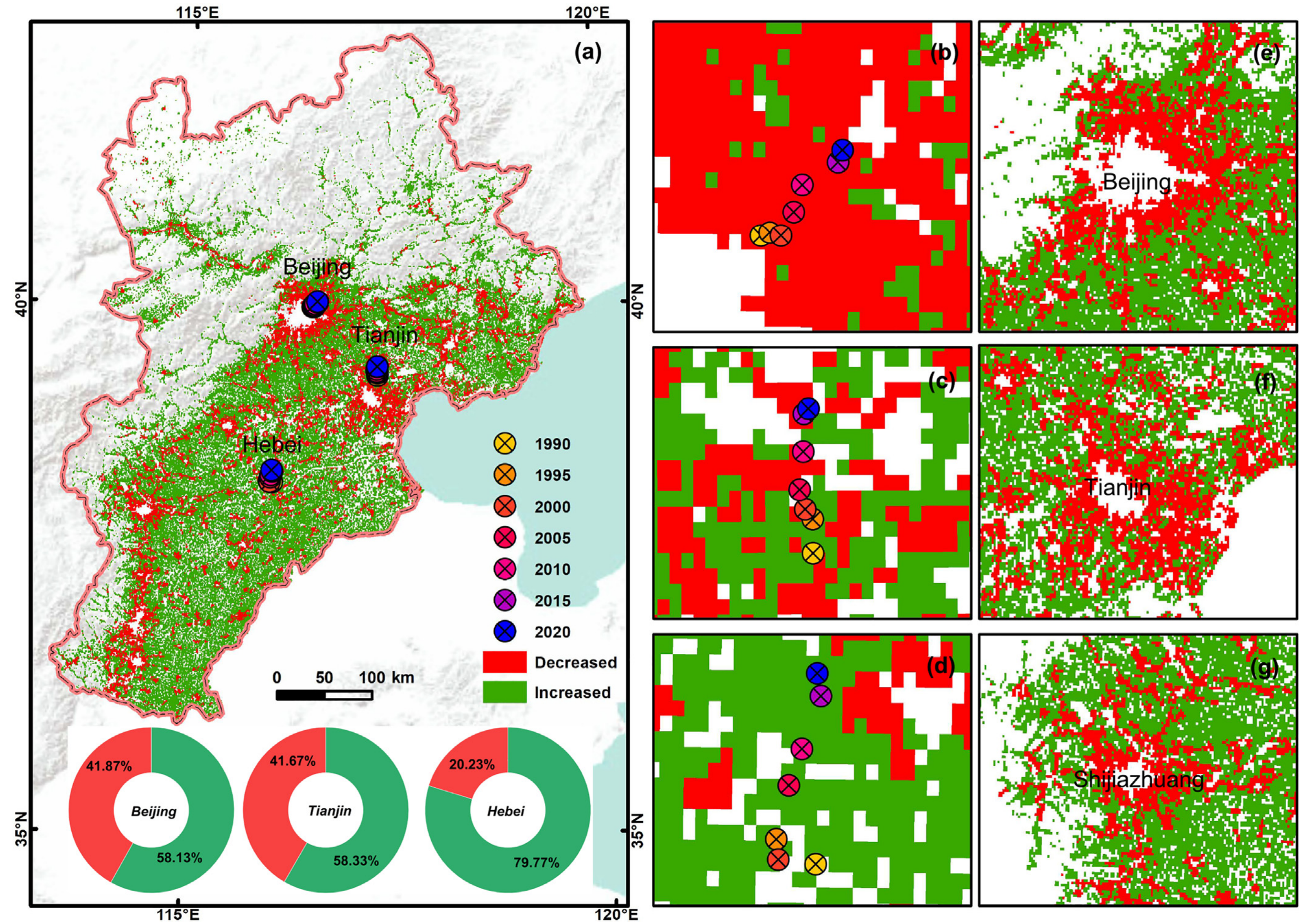


Fig. 5. (a) Spatially explicit map of RUS areas changes trend at 1 km × 1 km grids. The enlarged views show the proportion of increased and decreased grids in different administrative zones. The trajectory of the RUS centroid from 1990 to 2020 is shown in subplot (a); subplots (b), (c), and (d) are the zoom-in subplots for the trajectories of the RUS centroid in Beijing, Tianjin, and Hebei, respectively. Subplots (e), (f), and (g) are the zoom-in figures for the main distribution of decreased grids in Beijing, Tianjin, and Shijiazhuang, respectively.

41.67% were decreased grids, similar to Beijing, contiguous decreased grids were also mainly distributed around its central urban area (Fig. 5f). For Hebei, which has the largest zonal area, its increased (79.77%) and decreased (20.23%) grids proportion was similar to that of the entire BTH region. Similarly, as the largest city in Hebei, a large number of decreased grids were also distributed around Shijiazhuang, which is the provincial capital city of Hebei (Fig. 5g). In addition, based on the derived RUS patches within each administrative zone during 1990–2020 with a 5-year interval, the centroids of which have been identified. In Beijing, the centroid of its RUS patches had shifted about 7.64 km to the northeast over the past three decades due to the significant decrease RUS trend that occurred in the northeast area (Fig. 5b). However, the centroids of RUS patches in Tianjin and Hebei had shifted to the north about 9.31 km (Fig. 5d), and 12.31 km (Fig. 5d), respectively, which is closer to the administrative boundaries of Beijing.

Furthermore, Fig. 6 illustrates the spatial distribution of four change types of RUS areas for all counties (except for those counties without RUS since 1990) of the BTH region and the detailed temporal changes for different change types in typical counties. As shown in Fig. 6a, counties with the Convex-I in RUS areas accounted for 56.3% of 199 counties, most of which are far away from the well-developed urban zones. 9.5% of 199 counties saw the Concave-I in RUS areas, most of which belong to mountainous zones. 42 counties (21.1% of 199 counties) mainly distributed in suburb zones showed the Convex-D in RUS areas. And those counties that are close to the central urban area and gradually occupied by urban expansion showed the Concave-D in RUS areas (13.1% of 199 counties). Specifically, in view of the detailed temporal changes for different change types in typical counties, Tangxian (Fig. 6b) is a typical Convex-I county located in the middle of Hebei (with a change rate of 1.58 and a deviation value of 16.77), which represents the change type of RUS areas within the traditional agricultural production counties under the context of rural population loss. Kuangcheng (Fig. 6c) is a typical Concave-I county located in the northern mountainous area of Hebei (with a change rate of 1.73 and a deviation value of −7.41), which represents the change type of RUS areas within the natural mountainous counties with the implementation of ecological protection project. Daxing (Fig. 6c) is a typical Convex-D county located in the south of Beijing (with a change rate of −1.56 and a deviation value of 11.02), which represents the change type of RUS areas within the suburb counties under the context of rapid urbanization. Jinnan is (Fig. 6d) a typical Concave-D county located in the southeast of Tianjin (with a change rate of −1.76 and a deviation value of −5.18), which represents the change type of RUS areas within the highly urbanized counties. Interestingly, as a more urbanized area, Beijing has fewer Concave-D counties (districts) than Tianjin, which may be due to the fact that the county-level administrative area of Beijing is too

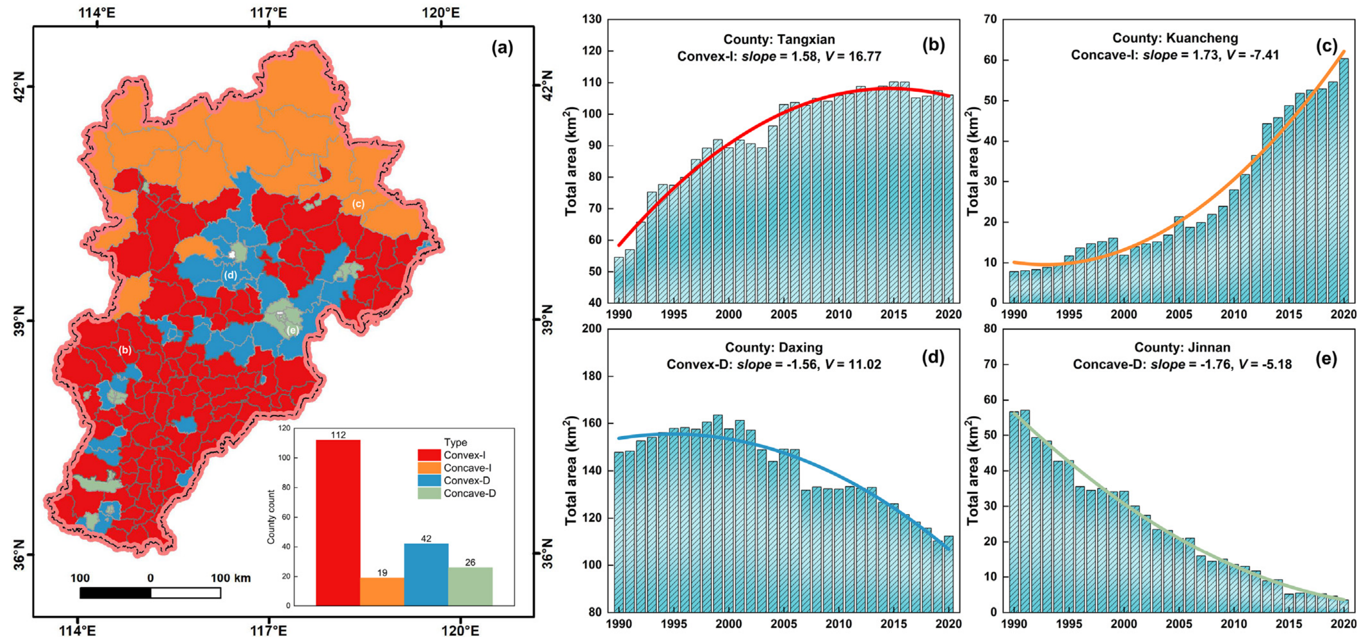


Fig. 6. (a) The spatial distribution of four change types of RUS areas for all counties of the BTH region and the detailed temporal changes of Convex-I (d), Concave-I (c), Convex-D (d), and Concave-D (e) in typical counties. The change rate of RUS areas is marked as "*Slope*" and the deviation value between the change curve and the linear regression line is marked as "*V*".

large compared with Tianjin, and it is not a sign that the RUS in Beijing is still experiencing a rapid decrease. In view of the counties that are similar in socioeconomic status, we also found that they may also belong to different change types in RUS (e.g., Chaoyang and Haidian of Beijing) due to different terrain characteristics.

4. Discussion

4.1. Contributions of the proposed RUS mapping model

Compared with the automatic delineation of physical boundaries by using remotely sensed data in urban-related studies [22,51], a cost-effective approach toward rural settlements mapping is considered insufficient due to the complexity and spatial dispersion of the target objects within the un-emphasized rural areas [28,52], placing restrictions on the quantitative understanding of the evolution of villages that affected by the natural geographical environment and human socio-economic activities [53,54]. In this regard, the proposed RUS mapping model in this study delineated the initial patch boundary by enhancing the spatial heterogeneity of the original ISA data and reducing its spatial fragmentation, using an integrated framework of the KDE method, the CA-based approach, and two morphological approaches. Then, the final RUS patches were successfully identified using the area-based as well as the NTL-based filter. The implementation of the proposed mapping model here can capture the spatial heterogeneity of RUS changes and the structural differences between individual RUS patches over long-term periods and large scales, which is critical information for the quantitative understanding of the rural areal system that is not available from the conventional dataset either censuses from field survey or other remote sensing-derived human settlement mapping products like GHSL, GAIA, etc. Also, it should be noted that both ISA and NTL data required in the proposed model are sufficiently available in related public products and their spatial resolution and accuracy are still improving with the efforts of relevant communities [55]. Therefore, our proposed model not only provides an effective way to automatic mapping RUS that enables the reduction of human labor and subjectivity but also may achieve higher accuracy with the improvement of those input data from related studies.

4.2. Dynamics of RUS from the perspective of consistent process

Prior to this study, to our limited knowledge, the best publicly available dataset of RUS in China was provided by the Resource and Environment Science and Data Center, which recorded the spatial distribution of RUS in China since the 1980s with a five-year interval. Otherwise, the researches on the spatiotemporal patterns of RUS were either carried out based on the county-level statistical data or obtained in a small area by manual sketching. As for a large region over a long-term period, quantifying the dynamics of RUS from the perspective of the consistent process is quite deficient due to the lack of annul RUS dataset, which constrains our ability to provide more detailed information and scientific guidance for related rural planning and policy making. In this study, based on the proposed mapping model, we first produced the annual RUS dataset with a spatial resolution at 30 m in the BTH region, China during 1990–2020 on the GEE platform. Through spatiotemporal analyses of these annual RUS dataset, we not only detected the break-points of its areas change trends, the average size as well as the standard deviation of its patches, and the size-rank relationship of its patches from the temporal dimension, but also revealed the heterogeneity of the annual changes of its areas, the migration of its centroids, and the change types of its areas in different counties from the spatial dimension. Based on a fine-grained delineation of RUS in both spatial and temporal resolutions here, we can summarize the new findings in the regularity of RUS changes compared with those from coarse-grained studies into two aspects. One is promoting the research paradigm of RUS transformation from a stage-wise qualitative description to an annual-wise quantitative detection over a long-term period, and the other is enhancing the understanding of the spatial heterogeneity of RUS at the patch, grid, and zone levels within a large region. Consequently, these analysis results provided spatiotemporally explicit information that can support the simulation of the rural territorial system and the rural revitalization planning [56,57] compared to similar results from census data. Additionally, due to the automatic production process, these analysis results also demonstrated the feasibility of achieving a real-time capacity for tracking the development of RUS in a cost-effective way compared to similar results from discontinuous spatial data.

4.3. Limitation analysis and future directions

Although the proposed mapping model and the produced annual dataset are suitable for understanding the evolutionary process of RUS across the BTH region over a long-term period, we are also aware of two limitations of this study that need to be further strengthened in future works. One is about the use of ISA data at a finer spatial resolution for a more accurate delineation of the RUS boundaries. Due to the lack of high-resolution historical observation records, we used annual ISA data from 1990 to 2020 in this study derived from Landsat images to generate the RUS dataset at the spatial resolution of 30 m. Given the development of Earth observation such as the Sentinel missions [58] as well as corresponding ISA products, some of the mapping errors in RUS with Landsat-based ISA data due to its coarse resolution can be improved in the future [59,60]. Another one is that some of the parameters and thresholds of the proposed model need to be optimized based on the different contexts of geographical zones when applied at national or global scales [61]. For example, the KD value and Moore neighborhood in the step of initial boundary delineation were set based on repeated attempts of the experimental process, which may not be suitable for regions where the geographical conditions differ significantly from those of the BTH region, and requires figure out how to improve the generalization of model parameters when expanding to a larger scale. Additionally, the NTL-based filter thresholds vary greatly in regions with different levels of economic development and ignorance of their spatial variations in similar studies at the national or global scale may cause filtering bias, which also needs to be further considered in future works.

5. Conclusions

In view of the existing deficiency in the accurate information on the dynamics of RUS, which restricts a consistent process analysis of the spatiotemporal patterns of RUS and provides limited scientific references for rural spatial planning and related policy making. In this study, we proposed an automatic mapping model to consistently delineate RUS boundaries and quantified the spatiotemporal patterns of RUS from the perspective of consistent process in the rapid but uneven urbanization region of BTH during 1990–2020. The results showed the following: (1) from a methodological point of view, the RUS dataset that was generated based on the proposed mapping model showed a good agreement with the CNLUCC, achieving a mean overall accuracy of 84.91% over the three decades in view of the location-based comparison, and a mean correlation value that crosses all scales was greater than 0.8 in view of the area-based comparison. (2) From a temporal point of view, we showed a trend of slower growth in the total area and the average size of RUS in the BTH region over the past three decades, along with a trend of widening differences within all RUS patches, and the size of RUS patches exhibited an exponential decay trend with the increase of the number of RUS patches. (3) From a spatial point of view, we found that the spatial distribution of the annual change of RUS areas in the BTH region is extremely uneven, the grids with the increasing and decreasing trend in the RUS area accounted for 76.33% and 23.67%, respectively. RUS areas in more than half of the counties of the BTH region saw a Convex-I change type, most of which are far away from the well-developed urban zones and belong to the plain areas engaged in agricultural production. In general, the cost-effective mapping model for RUS proposed and the generated results herein are convincing, which extended an annual-wise quantitative detection for RUS transformation and enhanced the understanding of the spatial heterogeneity of RUS across multi-scales, providing more explicit spatial information as well as consistent temporal information for a better understanding the spatiotemporal regularity of RUS changes.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgments

This work was supported by the Major Program of the National Natural Science Foundation of China (42293271), the Key Program of the National Natural Science Foundation of China (41931293), and the Funded Project of China Postdoctoral Science Foundation (2022M713116). We thank all anonymous reviewers for their helpful remarks.

Author contributions

Yansui Liu conceived the idea, designed the study, and revised the paper; Cong Ou performed the study and wrote the original draft; Yurui Li participated in resource and data curation; Liqiang Zhang and Jianhua He reviewed and edited the paper.

Appendix A. Supplementary materials

Supplementary materials to this article can be found online at https://doi.org/10.1016/j.scib.2023.08.006.

Data availability

The annual RUS dataset herein is now available in the public domain at https://doi.org/10.5281/zenodo.8092890 (accessed on 29 June 2023). This dataset was tagged in SHP file format, named "RUS_BTH_CN_v01.zip".

References

[1] Liu Y. Introduction to land use and rural sustainability in China. Land Use Pol 2018;74:1–4.

[2] Hudson JC. A location theory for rural settlement. Ann Assoc Am Geogr 1969;59:365–81.

[3] Baral H, Keenan RJ, Sharma SK, et al. Spatial assessment and mapping of biodiversity and conservation priorities in a heavily modified and fragmented production landscape in north-central Victoria, Australia. Ecol Indic 2014;36:552–62.

[4] Ristić D, Vukoičić D, Milinčić M. Tourism and sustainable development of rural settlements in protected areas-Example NP Kopaonik (Serbia). Land Use Pol 2019;89:104231.

[5] Song W, Liu M. Assessment of decoupling between rural settlement area and rural population in China. Land Use Pol 2014;39:331–41.

[6] Tan M, Li X. The changing settlements in rural areas under urban pressure in China: patterns, driving forces and policy implications. Landsc Urban Plan 2013;120:170–7.

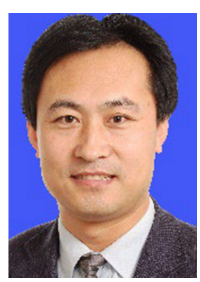
[7] Long H, Li Y, Liu Y. Analysis of evolutive characteristics and their driving mechanism of hollowing villages in China. Acta Geogr Sin 2009;64:1203–13.

[8] Liu Y, Li Y. Revitalize the world's countryside. Nature 2017;548:275–7.

[9] Hou Y, Zhou S, Burkhard B, et al. Socioeconomic influences on biodiversity, ecosystem services and human well-being: a quantitative application of the DPSIR model in Jiangsu, China. Sci Total Environ 2014;490:1012–28.

[10] Li G, Jiang G, Jiang C, et al. Differentiation of spatial morphology of rural settlements from an ethnic cultural perspective on the Northeast Tibetan Plateau, China. Habitat Int 2018;79:1–9.

[11] Leyk S, Maclaurin GJ, Hunter LM, et al. Spatially and temporally varying associations between temporary outmigration and natural resource availability in resource-dependent rural communities in South Africa: a modeling framework. Appl Geogr 2012;34:559–68.
[12] Su S, Zhang Q, Zhang Z, et al. Rural settlement expansion and paddy soil loss across an ex-urbanizing watershed in eastern coastal China during market transition. Reg Environ Chang 2011;11:651–62.
[13] Lang W, Chen T, Li X. A new style of urbanization in China: transformation of urban rural communities. Habitat Int 2016;55:1–9.
[14] Coskun HC. Quantifying landscape pattern and connectivity in a Mediterranean coastal settlement: the case of the Urla district, Turkey. Environ Monit Assess 2013;185:143–55.
[15] Zheng X, Wu B, Weston MV, et al. Rural settlement subdivision by using landscape metrics as spatial contextual information. Remote Sens 2017;9:486.
[16] Li G, Jiang C, Du J, et al. Spatial differentiation characteristics of internal ecological land structure in rural settlements and its response to natural and socio-economic conditions in the Central Plains, China. Sci Total Environ 2020;709:135932.
[17] Yang R, Zhang J, Xu Q, et al. Urban-rural spatial transformation process and influences from the perspective of land use: a case study of the Pearl River Delta Region. Habitat Int 2020;104:102234.
[18] Qu Y, Jiang G, Ma W, et al. How does the rural settlement transition contribute to shaping sustainable rural development? Evidence from Shandong, China. J Rural Stud 2021;82:279–93.
[19] Aiyar A, Rahman A, Pingali P. India's rural transformation and rising obesity burden. World Dev 2021;138:105258.
[20] Pan W, Wang J, Qin X, et al. Trends and types of rural residential land use change in China: a process analysis perspective. Growth Chang 2021;52:2437–52.
[21] Li X. Big Earth Data boost UN SDGs. Sci Bull 2023;68:773–4.
[22] Xu Z, Jiao L, Lan T, et al. Mapping hierarchical urban boundaries for global urban settlements. Int J Appl Earth Obs Geoinf 2021;103:102480.
[23] Henderson M, Yeh ET, Gong P, et al. Validation of urban boundaries derived from global night-time satellite imagery. Int J Remote Sens 2003;24:595–609.
[24] Li X, Gong P, Zhou Y, et al. Mapping global urban boundaries from the global artificial impervious area (GAIA) data. Environ Res Lett 2020;15:094044.
[25] Liang X, Liu X, Li X, et al. Delineating multi-scenario urban growth boundaries with a CA-based FLUS model and morphological method. Landsc Urban Plan 2018;177:47–63.
[26] Yin J, Soliman A, Yin D, et al. Depicting urban boundaries from a mobility network of spatial interactions: a case study of Great Britain with geo-located Twitter data. Int J Geogr Inf Sci 2017;31:1293–313.
[27] Zhen F, Cao Y, Qin X, et al. Delineation of an urban agglomeration boundary based on Sina Weibo microblog 'check-in' data: a case study of the Yangtze River Delta. Cities 2017;60:180–91.
[28] Hoffman-Hall A, Loboda TV, Hall JV, et al. Mapping remote rural settlements at 30 m spatial resolution using geospatial data-fusion. Remote Sens Environ 2019;233:111386.
[29] Ji H, Li X, Wei X, et al. Mapping 10-m resolution rural settlements using multi-source remote sensing datasets with the Google Earth Engine platform. Remote Sens 2020;12:2832.
[30] Pesaresi M, Blaes X, Ehrlich D, et al. A global human settlement layer from optical HR/VHR RS data: concept and first results. IEEE J Sel Top Appl Earth Obs Remote Sens 2013;6:2102–31.
[31] Gong P, Li X, Wang J, et al. Annual maps of global artificial impervious area (GAIA) between 1985 and 2018. Remote Sens Environ 2020;236:111510.
[32] Li X, Zhou Y, Zhu Z, et al. A national dataset of 30 m annual urban extent dynamics (1985–2015) in the conterminous United States. Earth Syst Sci Data 2020;12:357–71.
[33] Yang J, Huang X. The 30 m annual land cover dataset and its dynamics in China from 1990 to 2019. Earth Syst Sci Data 2021;13:3907–25.
[34] Sulla-Menashe D, Gray JM, Abercrombie SP, et al. Hierarchical mapping of annual global land cover 2001 to present: the MODIS Collection 6 Land Cover product. Remote Sens Environ 2019;222:183–94.
[35] Mousivand A, Arsanjani JJ. Insights on the historical and emerging global land cover changes: the case of ESA-CCI-LC datasets. Appl Geogr 2019;106:82–92.
[36] Zhang L, Ren Z, Chen B, et al. A prolonged artificial nighttime-light dataset of China (1984–2020). National Tibetan Plateau Data Center, 2021, https://data.tpdc.ac.cn/en/data/e755f1ba-9cd1-4e43-98ca-cd081b5a0b3e.
[37] Liu J, Kuang W, Zhang Z, et al. Spatiotemporal characteristics, patterns, and causes of land-use changes in China since the late 1980s. J Geog Sci 2014;24:195–210.
[38] Gorelick N, Hancher M, Dixon M, et al. Google Earth Engine: planetary-scale geospatial analysis for everyone. Remote Sens Environ 2017;202:18–27.
[39] De la Riva J, Pérez-Cabello F, Lana-Renault N, et al. Mapping wildfire occurrence at regional scale. Remote Sens Environ 2004;92:363–9.
[40] Shu Y, Li J, Yousif H, et al. Dark-spot detection from SAR intensity imagery with spatial density thresholding for oil-spill monitoring. Remote Sens Environ 2010;114:2026–35.
[41] Xie Z, Yan J. Kernel density estimation of traffic accidents in a network space. Comput Environ Urban Syst 2008;32:396–406.
[42] Liu X, Li X, Liu L, et al. A bottom-up approach to discover transition rules of cellular automata using ant intelligence. Int J Geogr Inf Sci 2008;22:1247–69.
[43] Leichtle T, Geiß C, Wurm M, et al. Unsupervised change detection in VHR remote sensing imagery—an object-based clustering approach in a dynamic urban environment. Int J Appl Earth Obs Geoinf 2017;54:15–27.
[44] Small C, Pozzi F, Elvidge CD. Spatial analysis of global urban extent from DMSP-OLS night lights. Remote Sens Environ 2005;96:277–91.
[45] Chen Z, Yu B, Song W, et al. A new approach for detecting urban centers and their spatial structure with nighttime light remote sensing. IEEE Trans Geosci Remote Sens 2017;55:6305–19.
[46] Chen T, De Jeu R, Liu Y, et al. Using satellite based soil moisture to quantify the water driven variability in NDVI: a case study over mainland Australia. Remote Sens Environ 2014;140:330–8.
[47] Cao Z, Li Y, Liu Y, et al. When and where did the Loess Plateau turn "green"? Analysis of the tendency and breakpoints of the normalized difference vegetation index. Land Degrad Dev 2018;29:162–75.
[48] Muggeo V. Segmented: An R package to fit regression models with broken-line relationships. R News 2008;8:20–5.
[49] Zhang G, Xiao X, Biradar CM, et al. Spatiotemporal patterns of paddy rice croplands in China and India from 2000 to 2015. Sci Total Environ 2017;579:82–92.
[50] Tian G, Wu J, Yang Z. Spatial pattern of urban functions in the Beijing metropolitan region. Habitat Int 2010;34:249–55.
[51] Zhao M, Zhou Y, Li X, et al. Mapping urban dynamics (1992–2018) in Southeast Asia using consistent nighttime light data from DMSP and VIIRS. Remote Sens Environ 2020;248:111980.
[52] Liu Y, Long H, Li Y. Human geography research based on the new thinking of global rural-urban relationship. Acta Geogr Sin 2021;76:2869–84.
[53] Liu Y, Long H, Chen Y, et al. Progress of research on urban-rural transformation and rural development in China in the past decade and future prospects. J Geog Sci 2016;26:1117–32.
[54] Liu Y, Liu J, Zhou Y. Spatio-temporal patterns of rural poverty in China and targeted poverty alleviation strategies. J Rural Stud 2017;52:66–75.
[55] Sun Z, Du W, Jiang H, et al. Global 10-m impervious surface area mapping: a big Earth data based extraction and updating approach. Int J Appl Earth Obs Geoinf 2022;109:102800.
[56] Liu Y, Zhou Y, Li Y. Rural regional system and rural revitalization strategy in China. Acta Geogr Sin 2019;74:2511–28.
[57] Liu Y. The basic theory and methodology of rural revitalization planning in China. Acta Geogr Sin 2020;75:1120–33.
[58] Berger M, Aschbacher J. The Sentinel missions-new opportunities for science. Remote Sens Environ 2012;120:1–276.
[59] Huang X, Yang J, Wang W, et al. Mapping 10 m global impervious surface area (GISA-10m) using multi-source geospatial data. Earth Syst Sci Data 2022;14:3649–72.
[60] Guo H, Dou C, Chen H, et al. SDGSAT-1: the world's first scientific satellite for sustainable development goals. Sci Bull 2023;68:34–8.
[61] Lu D, Weng Q. A survey of image classification methods and techniques for improving classification performance. Int J Remote Sens 2007;28:823–70.



Yansui Liu is the fellow of the World Academy of Sciences (TWAS), a professor of the Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, director of the Key Laboratory of the Regional Sustainable Development Analysis and Simulation, Chinese Academy of Sciences, and chair of International Geographical Union Commission on Agricultural Geography and Land Engineering. His research interest focuses on agricultural and rural geography, land use and urban-rural development, rural areal system, and human-Earth system science.

Cong Ou is a postdoctoral researcher of the Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences. His research interest focuses on remote sensing monitoring of human-Earth system, quantitative diagnosis of the processes & mechanisms of rural areal system, and technical method of dynamic simulation of territorial space.