land

MDPI

Article

Evaluating the Causal Impact of Greenbelt Planning on Regional Ecological Quality: An Integrated RSEI-PSM-DID Framework

Yu Wang [1], Chengwei He [1] and Zhijun Zhang [2,*]

[1] Department of Urban and Rural Planning, Tianjin University, Tianjin 300072, China; yu_wang2019@tju.edu.cn (Y.W.); hechengwei@tju.edu.cn (C.H.)
[2] Tianjin Forestry Affairs Center, Tianjin 300191, China
* Correspondence: zhangzhijun@whu.edu.cn

Abstract

Greenbelt planning is a key spatial instrument for optimizing ecological space, yet whether its implementation genuinely improves regional ecological quality remains insufficiently examined. This study evaluates the causal impact of the Tianjin Green Ecological Barrier Zone (TGEBZ) planning on regional ecological quality using an annual 1 km grid panel from 2014 to 2023. Ecological quality was measured by the Remote Sensing Ecological Index (RSEI), and the effect was identified using a propensity-score matching difference-in-differences (PSM-DID) design under two-way fixed effects. The RSEI inside the zone rose from 0.5231 in 2014 to 0.5894 in 2023 (+12.66%), against 2.26% for the grids outside; it fell to its lowest value in 2016, recovered gradually, peaked in 2021, and remained relatively high from 2020 to 2023. The planning is estimated to have raised the RSEI inside the zone by 0.022 ($p < 0.01$, 95% CI [0.016, 0.027]); restricting the controls to predetermined climate variables gives a value of 0.016. The estimate held steady through a range of robustness evaluations. This study links spatial planning to measurable ecological outcomes and extends the RSEI from descriptive monitoring to causal evaluation, offering a replicable framework for ecological governance in metropolitan regions.

Keywords: greenbelt; Tianjin Green Ecological Barrier Zone; remote sensing; ecological quality; RSEI; PSM-DID

1. Introduction

Urbanization is one of the most profound land surface transformation processes. The United Nations projects that by 2050, two-thirds of the global population will reside in cities, while the expansion of built-up areas worldwide has outpaced population growth [1]. Over the past four decades, China's urbanization rate has increased from below 20% to above 65%. While rapid urban expansion has generated development dividends, it has also caused a series of ecological and environmental problems, including proliferation of construction land, vegetation loss, surface hardening, rising thermal environmental pressure, and fragmentation of ecological space. These problems have become key bottlenecks that constrain sustainable urban development and residents' well-being.

The greenbelt is a classical planning tool employed to safeguard ecological areas and has been extensively adopted worldwide since its inception. Policy performance has long been a central topic in urban planning research. By delineating continuous green open spaces on the urban periphery, greenbelts establish rigid boundaries that limit the expansion of construction land while carrying multiple ecological functions, such as recreational amenity

Check for updates

Academic Editor: Jane Southworth

Received: 24 October 2026
Revised: 16 November 2026
Accepted: 19 November 2026
Published: 26 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

provision, climate regulation, and biodiversity conservation. However, although greenbelts are expected at the planning level, whether their implementation can genuinely improve regional ecological quality remains to be empirically tested through long-term observations and rigorous causal inferences.

Regarding the effectiveness of greenbelts in curbing sprawl, Pourtaherian and Jaeger compared 60 European cities and confirmed that greenbelts generally mitigate urban sprawl by lowering per capita land consumption and promoting compact built-up development [2]. Mosharafian and Jaeger further proposed targets, limits, and warning values for urban sprawl, although their study focused only on Montreal and was exclusively concerned with sprawl control itself [3]. Han et al. examined greenbelt counties in the United States from a carbon sequestration perspective and revealed lower land conversion rates and higher carbon storage potential [4]. Evidently, the mainstream concern of existing greenbelt research remains whether greenbelts curb sprawl, with little attention devoted to their impacts on regional ecological quality after implementation. In terms of urban greenbelts in China, Ma and Jin analyzed the difficulties in maintaining Beijing's greenbelts from an economic perspective [5], while Yuan focused on the implementation process of the outer ring road greenbelt in Shanghai [6]. In South Korea, Kim et al. and Han et al. concentrated on the effects of greenbelt deregulation or relaxation on land development [7,8]. None of these studies directly link greenbelt policies to a quantitative causal assessment of ecological quality. Although recent studies have shifted toward the ecological patterns of greenbelts, including the measurement of fragmentation and connectivity [9], morphological analysis of greenbelt forest habitats [10], and remote sensing–based evaluations of greenbelt policy performance [11], the latter still characterizes ecological change using single vegetation indicators such as the normalized difference vegetation index (NDVI), lacking a systematic depiction of regional comprehensive ecological quality.

In the remote sensing assessment of ecological quality, the Remote Sensing Ecological Index (RSEI) proposed by Xu integrates four components, namely greenness, wetness, dryness, and heat, and compresses them into a single index through principal component analysis (PCA), thereby providing an effective tool for the objective and rapid monitoring of urban and regional ecological conditions [12,13]. In the field of remote sensing–based ecological quality assessment, studies on the spatiotemporal changes in vegetation ecological quality in nature reserves [14], wetland ecological risk assessment [15], ecological policy benefit evaluation [16], and implementation evaluation of ecological spatial planning [17] all corroborate the objectivity and reliability of remote sensing–based assessment methods. However, these studies mostly remain at the descriptive level of spatiotemporal characteristics and fail to integrate remote sensing measures of ecological quality with policy causal inference, making it difficult to answer the question of whether planning implementation genuinely improves ecological quality. The PSM-DID framework has been shown to effectively mitigate selection bias and endogeneity in policy evaluations, such as low-carbon cities [18–20] and smart cities [21,22], thereby providing a mature methodological toolkit for the causal inference of policy effects.

Three gaps remain in the existing literature. First, with regard to assessment content, greenbelt research either focuses on sprawl control and economic effects or depicts ecological change using single indices; systematic quantitative evaluations of comprehensive ecological quality after greenbelt implementation are still lacking. Second, most studies stop at spatiotemporal descriptions without linking them to causal inferences about policy effects. Third, existing studies predominantly use administrative units as the unit of analysis and often rely on only a few snapshot years, leaving ecological effects at fine grid scales and over continuous annual time series largely unexamined. To address these gaps, this study uses the RSEI for comprehensive measurement, PSM-DID for causal attribution, and an

annual 1 km grid panel from 2014 to 2023 for fine-scale resolution before and after the plan, thereby examining whether greenbelt-related spatial planning has substantially improved ecological quality within the greenbelt area.

Beyond these empirical gaps, the study carries both theoretical and practical significance. It reframes the greenbelt from a normative planning ideal into a testable causal proposition—whether a designated ecological barrier actually changes regional ecological quality—and thereby links planning-instrument theory to measurable ecological outcomes; it also extends the RSEI, which is largely treated as a descriptive monitoring tool, into a component of causal policy evaluation. The RSEI-PSM-DID procedure developed here offers planners and policymakers an operational template for judging whether greenbelt designation works, and evidence for calibrating boundary delimitation, construction-land control, and long-term ecological monitoring in metropolitan regions.

2. Study Area and Data Sources

2.1. Study Area

Tianjin is situated in northern China, covering 11,966.45 km^2 with 16 municipal districts. By 2025, its permanent resident population reached 13.63 million, and the urbanization rate was 86.21%. Tianjin presents a corridor-shaped spatial pattern; its development corridor stretches roughly 40 km from the old downtown toward the coastal area. The interstitial undeveloped ecological functional zone between them is our study object: the Tianjin Green Ecological Barrier Zone (TGEBZ). It constitutes a key component of Tianjin's spatial pattern of "three ecological protection zones, two ecological service belts, and the central key ecological barrier".

The Tianjin Green Ecological Barrier Zone Plan (2018–2035) was approved and enacted in 2019, targeting a 70% blue-green-space proportion within its total area of 736 km^2. Preparatory and pilot construction began in parts of the zone in 2018; the baseline retains 2018 in the pre-policy period, with 2019 as the year of approval and full implementation, and Section 4.6 tests the sensitivity of the estimate to treating 2018 as a transition year. The target has been pursued through regulatory and planning instruments: two local regulations, a master plan with special and district plans, a leading group established in 2019, and standards for high-standard pond and farmland construction, together with planning control over construction, industrial, and service projects. Different from ring-shaped greenbelts surrounding monocentric metropolises such as London, Seoul and Beijing, the TGEBZ acts as a typical "Green Valley" sandwiched between multi-urban clusters to prevent contiguous urban sprawl and protect regional ecological space. It provides a valuable case for greenbelt-policy impact research and can deliver quantitative evaluation frameworks and empirical evidence for ecological spatial governance.

The study area comprises the four near-suburban districts of Tianjin—Beichen, Xiqing, Dongli, and Jinnan—which together cover approximately 1890 km^2. It lies on the alluvial plain of the Haihe River system at elevations of approximately 2–5 m above sea level, gently sloping southeast toward the Bohai coast. The climate is warm-temperate and semi-humid under a monsoon regime: the mean annual temperature is about 14 °C, and annual precipitation is approximately 550–600 mm, roughly three-quarters of which falls between June and August. The area belongs to the warm-temperate deciduous broadleaved forest zone; its present cover is cultivated vegetation—cropland under a winter wheat–summer maize rotation and shelterbelt plantations.

As shown in Figure 1, these four districts constitute the study area, which ensures that the treatment and control groups are drawn from comparable administrative and physical settings. The entire study area was divided into 1 km × 1 km grid cells. After excluding grids with missing data, 2083 valid grid cells intersecting the districts were retained for analysis.

Grids within the study area that intersected the TGEBZ planning area were assigned to the treatment group (orange area in Figure 1), while the remaining grids were assigned to the control group (blue area in Figure 1). Because the plan boundary is a regulatory line rather than a land-cover boundary, any intersection with it qualifies a grid as treated, and the sensitivity of the estimate to this rule is examined in Section 4.6.

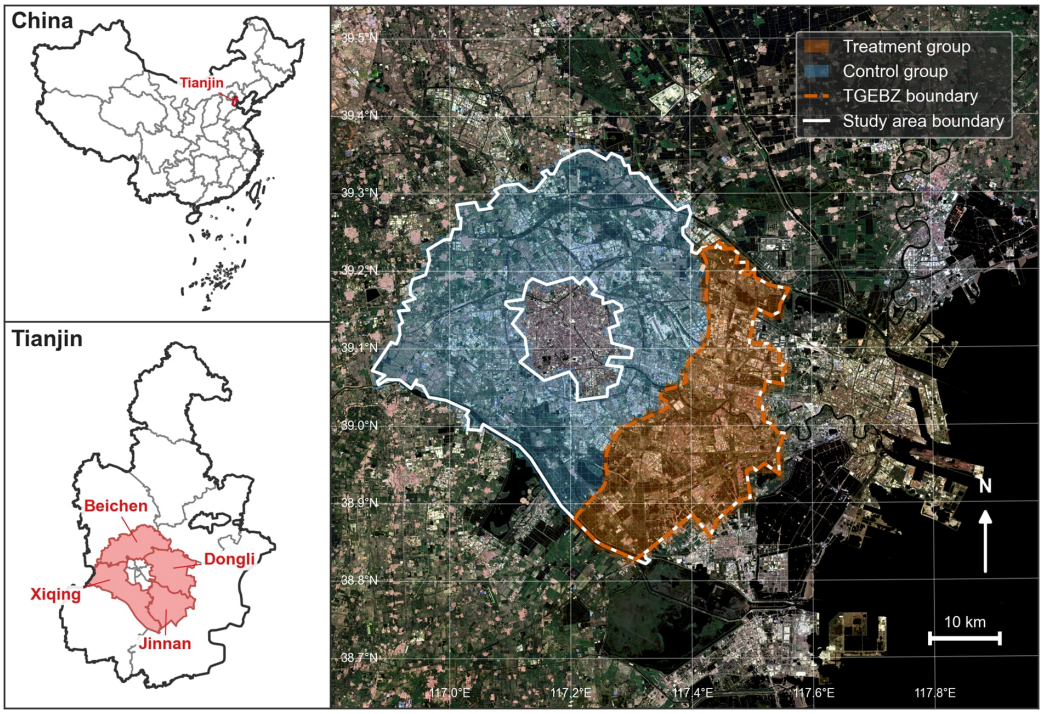

Figure 1. The study area and TGEBZ boundary.

2.2. Data Sources and Processing

We utilized the Remote Sensing Ecological Index (RSEI) covering multiple periods to thoroughly assess ecological quality. Additionally, we applied the Propensity Score Matching–Difference-in-Differences (PSM-DID) model to evaluate the effectiveness of the TGEBZ plan's implementation, achieving an integration of remote sensing ecological measurements with policy causal identification.

The study period was 2014–2023, with 2019—the year the TGEBZ plan was approved by the Tianjin Municipal People's Government—set as the policy intervention point. The periods of 2014–2018 and 2019–2023 were defined as the pre-policy and post-policy periods, respectively, with five years in each period. As shown in Table 1, the data used in this study include remote sensing imagery, land use, nighttime light, points of interest (POI), temperature, and precipitation. Remote sensing imagery was used to calculate the RSEI to characterize ecological quality, whereas the other five variables served as control variables. The remote sensing imagery was sourced from the Landsat Collection 2 Level-2 dataset provided by NASA and USGS via the Microsoft Planetary Computer website, filtered by the following criteria: images acquired between 1 May and 1 October each year, cloud cover equal to or less than 30%, Landsat 8 and Landsat 9 platforms, path/row numbers 122/123 and 032/033. Land use data were sourced from Yang and Huang [23], nighttime light data from Chen et al. [24], point of interest data from the OpenStreetMap open platform, temperature data from Peng [25], and precipitation data from Peng [26]. Both climate products are monthly fields at a resolution of approximately 1 km; the twelve-monthly fields of each year were averaged, so that precipitation is the mean monthly value and temperature is the annual mean of the monthly means.

Table 1. Key data sources.

Data	Source	Resolution
Remote sensing imagery	NASA/USGS	30 m
Land use	Yang and Huang [23]	30 m
Nighttime light	Chen et al. [24]	~ 30 m
POI	OpenStreetMap	—
Temperature	Peng [25]	~1 km
Precipitation	Peng [26]	~1 km

Note: POI = point of interest; NASA/USGS = National Aeronautics and Space Administration/United States Geological Survey.

First, the RSEI was calculated for each image from multiple remote sensing images that met the criteria within a given year. The images were then clipped and mosaicked based on the administrative boundaries of Tianjin, and the mean RSEI value was computed for overlapping areas within the same year to obtain the RSEI data. POI data were used to calculate the facility density within each grid cell. In the land use data, only the impervious surface area was retained as a land use control variable to avoid multicollinearity arising from the coexistence of multiple land use types. Subsequently, the RSEI, land use, nighttime light, POI density, temperature, and precipitation data were unified to the same coordinate system and extracted at a 1 km grid scale. The complete chain from the six data sources to the analysis panel is summarized in Figure 2.

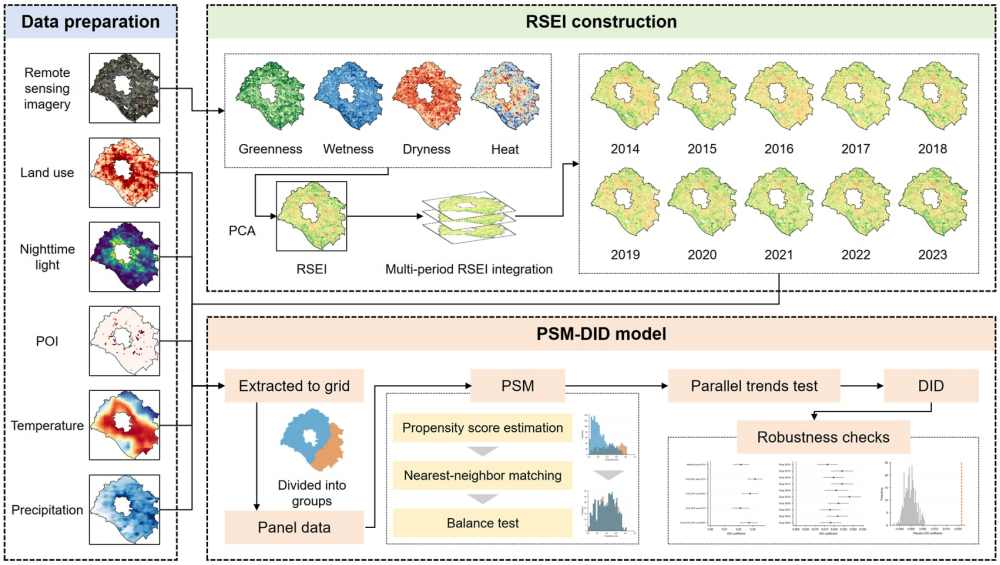

Figure 2. The research analytical workflow.

3. Methods

3.1. Research Framework

This study aims to answer the following two core questions. First, how did the ecological quality of the TGEBZ change over the study period? Second, has the implementation of TGEBZ planning significantly improved ecological quality within the zone? To answer these questions, an analytical framework comprising two components was constructed: RSEI construction for ecological quality measurement and PSM-DID modeling for causal inference.

As shown in Figure 2, the first part addresses the first research question through the RSEI construction. Four ecological components, namely greenness, wetness, dryness, and heat, were computed from remote sensing imagery, synthesized into the RSEI through

principal component analysis, and integrated across years to construct the RSEI, thereby depicting the spatiotemporal dynamics of ecological quality in the study area from 2014 to 2023.

The second part addressed the second research question using the Propensity Score Matching–Difference-in-Differences (PSM-DID) model. The RSEI and control variables of climatic conditions, human activities, and land use were uniformly extracted to the 1 km grid scale to form a grid-level panel dataset, and the treatment and control groups were delineated according to the TGEBZ boundary. Given that the TGEBZ was not randomly delineated, the Propensity Score Matching (PSM) technique was employed to mitigate selection bias by constructing comparable samples through one-to-one nearest-neighbor matching within the region of common support.

The region of common support is the range of propensity scores where both the treatment and control units are found. A two-way fixed-effects difference-in-differences (DID) model was then estimated on the matched sample with cluster-robust standard errors at the grid level. The reliability of the core conclusions was further verified through a battery of robustness checks, including policy time window adjustment, exclusion of transition years, time placebo tests, leave-one-year-out tests, and randomization placebo tests.

3.2. Construction of RSEI and Annual Integration

The RSEI integrates four ecological components into a single composite index through principal component analysis, which objectively reflects regional ecological quality. The four components are calculated as follows:

(1) Greenness component. The Normalized Difference Vegetation Index (NDVI) was used to characterize the vegetation cover and growth conditions:

$$\mathrm{NDVI} = \frac{\rho_{\mathrm{NIR}} - \rho_{\mathrm{Red}}}{\rho_{\mathrm{NIR}} + \rho_{\mathrm{Red}}} \tag{1}$$

where ρ_{NIR} and ρ_{Red} denote the surface reflectance in the near-infrared and red bands, respectively.

(2) Wetness component. The Tasseled Cap wetness component (Wet) was used to characterize surface moisture conditions.

$$\mathrm{Wet} = 0.0382\rho_{\mathrm{Blue}} + 0.2137\rho_{\mathrm{Green}} + 0.3536\rho_{\mathrm{Red}} + 0.2270\rho_{\mathrm{NIR}} - 0.6108\rho_{\mathrm{SWIR1}} - 0.6351\rho_{\mathrm{SWIR2}} \tag{2}$$

where ρ_{Blue}, ρ_{Green}, ρ_{Red}, ρ_{NIR}, ρ_{SWIR1}, and ρ_{SWIR2} denote the surface reflectance in the blue, green, red, near-infrared, shortwave infrared 1, and shortwave infrared 2 bands, respectively. The band coefficients were obtained according to the research by Zhai et al. [27] and Xu et al. [28], who confirmed the consistency of these coefficients between Landsat-8 and Landsat-9.

(3) Dryness component. The Normalized Difference Built-up and Soil Index (NDBSI), constructed as the average of the Index-based Built-up Index (IBI) and Soil Index (SI), was used to characterize surface hardening.

$$\mathrm{NDBSI} = \frac{\mathrm{IBI} + \mathrm{SI}}{2} \tag{3}$$

$$\mathrm{IBI} = \frac{\frac{2\rho_{\mathrm{SWIR1}}}{\rho_{\mathrm{SWIR1}}+\rho_{\mathrm{NIR}}} - \frac{\rho_{\mathrm{NIR}}}{\rho_{\mathrm{NIR}}+\rho_{\mathrm{Red}}} - \frac{\rho_{\mathrm{Green}}}{\rho_{\mathrm{Green}}+\rho_{\mathrm{SWIR1}}}}{\frac{2\rho_{\mathrm{SWIR1}}}{\rho_{\mathrm{SWIR1}}+\rho_{\mathrm{NIR}}} + \frac{\rho_{\mathrm{NIR}}}{\rho_{\mathrm{NIR}}+\rho_{\mathrm{Red}}} + \frac{\rho_{\mathrm{Green}}}{\rho_{\mathrm{Green}}+\rho_{\mathrm{SWIR1}}}} \tag{4}$$

$$\mathrm{SI} = \frac{\left(\rho_{\mathrm{SWIR1}} + \rho_{\mathrm{Red}}\right) - \left(\rho_{\mathrm{NIR}} + \rho_{\mathrm{Blue}}\right)}{\left(\rho_{\mathrm{SWIR1}} + \rho_{\mathrm{Red}}\right) + \left(\rho_{\mathrm{NIR}} + \rho_{\mathrm{Blue}}\right)} \tag{5}$$

where IBI is the index-based built-up index, which primarily reflects building and impervious surface characteristics; SI is the soil index, which primarily reflects bare soil and other surface characteristics; and the meanings of the band reflectance values are the same as those in Equations (1) and (2), respectively.

(4) Heat component. Land surface temperature (LST) is used to characterize the surface thermal environment. This component can be obtained directly from remote sensing imagery and is denoted as Heat.

(5) Synthesis of the RSEI. To eliminate the dimensional differences among the four components, each component was normalized separately as follows:

$$\xi^* = \frac{\xi - \xi_{min}}{\xi_{max} - \xi_{min}} \quad (6)$$

where ξ denotes the original value of a given ecological component, ξmin and ξmax denote the minimum and maximum values of that component, respectively, and ξ* denotes the normalized value.

Principal component analysis was then performed on the four normalized components, and the first principal component (PC1) was extracted to construct the RSEI:

$$PC_1 = w_1 NDVI^* + w_2 Wet^* + w_3 NDBSI^* + w_4 Heat^* \quad (7)$$

where NDVI*, Wet*, NDBSI*, and Heat* denote the normalized greenness, wetness, dryness, and heat components, respectively, and w_1, w_2, w_3, and w_4 are the weights corresponding to the first principal component, respectively.

In this study, the weights of NDVI and Wet in the first principal component were positive, whereas those of NDBSI and Heat were negative. Thus, PC1 positively reflects ecological quality, and the RSEI was obtained by normalizing PC1 as follows:

$$RSEI = \frac{PC_1 - PC_{1,min}}{PC_{1,max} - PC_{1,min}} \quad (8)$$

where $PC1_{,min}$ and $PC1_{,max}$ denote the minimum and maximum values of PC1, respectively. The RSEI ranges from 0 to 1, with larger values indicating better ecological quality. Both the normalization and PCA were performed scene by scene; PC1 explained 60.6–88.0% of the variance across the 193 scenes of 2014–2023 (mean 79.0%), and Table S1 lists the annual means of the variance share and of the four loadings.

(6) Annual integration of the RSEI. To improve the assessment accuracy of the RSEI, the RSEI values computed from multiple images within the same year were integrated to obtain the annual RSEI value.

$$RSEI_{i,t} = \frac{1}{N_{i,t}} \sum_{m=1}^{M_t} I_{i,m,t} \times RSEI_{i,m,t} \quad (9)$$

$$N_{i,t} = \sum_{m=1}^{M_t} I_{i,m,t} \quad (10)$$

where i indexes grids, t indexes years, M_t is the total number of valid images within year t, $RSEI_{i,m,t}$ is the RSEI value of the m-th image in year t at grid i, $I_{i,m,t}$ is a valid coverage indicator that equals 1 when the m-th image has a valid RSEI value at grid i and 0 otherwise, $N_{i,t}$ is the number of valid images at grid i within year t, and $RSEI_{i,t}$ is the RSEI value of grid i in year t.

3.3. Propensity-Score Matching

(1) Variable specifications. The dependent variable is the grid-level RSEI. The core explanatory variable is the interaction between the treatment group dummy and the policy time dummy.

$$DID_{i,t} = Treat_i \times Post_t \quad (11)$$

where $Treat_i$ is the treatment group dummy, which equals 1 when grid i intersects the TGEBZ and 0 otherwise, and $Post_t$ is the policy time dummy, which equals 1 when $t \geq 2019$ and 0 otherwise.

The vector of the control variables is as follows:

$$X_{i,t} = \left(Tmean_{i,t}, Prec_{i,t}, NLI_{i,t}, POI_{i,t}, ISA_{i,t}\right) \tag{12}$$

where $Tmean_{i,t}$ and $Prec_{i,t}$ denote the annual mean temperature and monthly mean precipitation of grid i in year t, respectively, which control for natural climatic conditions; the nighttime light index (NLI) and point-of-interest density, denoted $NLI_{i,t}$ and $POI_{i,t}$, respectively, control for the intensity of human activities; and the impervious surface area (ISA), denoted $ISA_{i,t}$, controls for land use conditions.

(2) Propensity score estimation. The propensity score is defined as the conditional probability of a grid entering the treatment group, given its pre-policy characteristics:

$$p_i = P(Treat_i = 1 \mid Z_i) \tag{13}$$

where p_i is the propensity score of grid i, and Z_i is the vector of matching variables, comprising the pre-policy mean RSEI, pre-policy RSEI trend, and pre-policy means of the control variables, that is, $Z_i = (RSEI_{i,pre}, \Delta RSEI_i, X_{i,pre})$.

A logit model was employed to estimate the propensity score:

$$p_i = \frac{\exp\left(c + Z_i^T \delta\right)}{1 + \exp\left(c + Z_i^T \delta\right)} \tag{14}$$

where c is a constant term and δ is the vector of parameters to be estimated.

The pre-policy mean RSEI, pre-policy means of the control variables, and pre-policy RSEI trends were defined as follows:

$$\bar{RSEI}_{i,pre} = \frac{1}{5}\sum_{t=2014}^{2018} RSEI_{i,t} \tag{15}$$

$$\bar{X}_{i,pre} = \frac{1}{5}\sum_{t=2014}^{2018} X_{i,t} \tag{16}$$

$$\Delta RSEI_i = RSEI_{i,2018} - RSEI_{i,2014} \tag{17}$$

where $RSEI_{i,pre}$ is the mean RSEI of grid i over the pre-policy period; $X_{i,pre}$ is the mean of the control variables of grid i over the pre-policy period; and $\Delta RSEI_i$ is the pre-policy RSEI trend of grid i, that is, the difference in RSEI between 2018 and 2014.

(3) The region of common support and nearest-neighbor matching. To ensure matching quality, only the samples within the overlapping region of propensity scores between the treatment and control groups, namely, the region of common support, were first retained. One-to-one nearest-neighbor matching was then employed to match each treatment grid with the control grid having the closest propensity score, and a caliper of 0.05 was imposed such that matching was allowed only when the difference in propensity scores between the treatment and control grids did not exceed 0.05. This approach retains sufficient samples while ensuring matching quality.

$$j(i) = \arg\min_{j:Treat_j=0} |p_i - p_j| \tag{18}$$

$$\left|p_i - p_{j(i)}\right| \leq 0.05 \tag{19}$$

where $j(i)$ is the control grid matched to treatment grid i, and p_j is the propensity score of control grid j. Matching without replacement is adopted, meaning that each control grid

can be matched to at most one treatment grid to avoid estimation bias arising from the repeated use of control samples.

(4) Balance tests. After matching, the standardized mean difference (SMD) was used to test the balance of each covariate between groups.

$$SMD = \frac{\bar{X}_T - \bar{X}_C}{\sqrt{\left(S_T^2 + S_C^2\right)/2}} \quad (20)$$

where X_T and X_C denote the means of a given covariate in the treatment and control groups, respectively, and $S^2{}_T$ and $S^2{}_C$ denote the variances of the covariate in the treatment and control groups, respectively. A covariate was considered balanced between the two groups when $|SMD| < 0.1$. In addition, Welch's two-sample *t*-test was employed to examine whether the mean differences in each covariate between the matched treatment and control groups were statistically significant, thereby validating the comparability of the matched samples across multiple dimensions.

3.4. Difference-in-Differences Model

(1) Pre-policy parallel trends test. The validity of the DID model depends on the treatment and control groups having similar time trends before the policy implementation. Using the 2014–2018 pre-policy sample, a time trend variable was constructed, and the parallelism of the two groups' trends was tested:

$$Trend_t = t - 2014 \quad (21)$$

$$RSEI_{i,t} = \theta(Treat_i \times Trend_t) + \gamma X_{i,t} + \mu_i + \lambda_t + \varepsilon_{i,t} \quad (22)$$

where $Trend_t$ is the time trend variable with 2014 as the base year, θ is the coefficient of the trend interaction term, γ is the vector of control variable coefficients, μ_i is the grid fixed effect, λ_t is the year fixed effect, and ε_{it} is the random error term. The statistical insignificance of θ indicates the absence of significant linear trend differences between the two groups before the policy implementation, which is consistent with the parallel trends assumption. It should be noted that because the treatment group dummy $Treat_i$ does not vary over time and the time trend $Trend_t$ varies only over time, the two variables are absorbed by the grid fixed effects and the year fixed effects, respectively; thus, only the interaction term $Treat_i \times Trend_t$ is retained in Equation (22).

(2) PSM-DID baseline model. A two-way fixed effects DID model was estimated on the matched sample to estimate the policy effect of the TGEBZ planning:

$$RSEI_{i,t} = \beta DID_{i,t} + \gamma X_{i,t} + \mu_i + \lambda_t + \varepsilon_{i,t} \quad (23)$$

where β is the core coefficient of interest, representing the policy effect of TGEBZ planning; a positive and significant β indicates that the planning significantly increased the RSEI of the treatment grids. The meanings of the other symbols are the same as those in Equation (22). Similarly, $Treat_i$ and $Post_t$ were absorbed by the grid fixed effects and year fixed effects, respectively, and only the interaction term $DID_{i,t}$ was retained in Equation (23).

Three progressive models were specified to examine the sensitivity of the core coefficient estimate to the specification of the control variables. Model 1 is the baseline model without any control variables and contains only the core explanatory variable, DID. Model 2 adds the two climate controls (Tmean and Prec), which are predetermined with respect to the planning, and therefore identifies the total policy effect. Model 3 further adds human activity controls, NLI and POI, and a land-use control, ISA; because these variables may themselves respond to the planning, the Model 3 coefficients are read as associations conditional on contemporaneous conditions. Model 3 is retained as the benchmark specification,

and Model 2 is reported alongside it so that the total effect and the conditional association can be compared. Considering the serial correlation among observations of the same grid across years and the potential correlation of disturbances at the grid level, the standard errors in all regressions are clustered at the grid level to obtain robust statistical inferences.

(3) Robustness checks. Five robustness checks were conducted to verify the reliability of the baseline results. First, policy time window adjustment: the post-policy period was redefined as 2020–2023 to examine whether the policy effect exhibited a release lag and to avoid contamination from the early stages of policy implementation in 2019. Second, the transition years (2018 and 2019) were excluded to rule out policy anticipation effects and early implementation disturbances. Third, time placebo tests: based on the pre-policy sample, 2016 and 2017 were separately designated as pseudo-policy years; insignificant DID coefficients for the pseudo-policy years would indicate the absence of spurious policy effects before the actual policy implementation. Fourth, leave-one-year-out tests: each year from 2014 to 2023 was excluded in turn, and the model was re-estimated to determine whether the results were driven by any particular year. Fifth, randomization placebo tests (randomization inference) were used to rule out the possibility that the true DID coefficient reflected a chance occurrence, and treatment group labels were randomly shuffled while keeping the number of treatment units unchanged. The DID coefficients were re-estimated K times to obtain the distribution of placebo DID coefficients, from which the empirical *p*-value was computed:

$$p^{perm} = \frac{1 + \sum_{k=1}^{K} 1\left(|\hat{\beta}_k| \geq |\hat{\beta}|\right)}{1 + K} \quad (24)$$

where K is the number of random permutations (K = 300 in this study), β_k is the placebo DID coefficient estimated in the kth random permutation, and β is the true DID coefficient estimated on the matched sample. The whole of 1 ($|\hat{\beta}_k| \geq |\hat{\beta}|$) is the indicator function, where 1 is the sign of the indicator function, and $|\hat{\beta}_k| \geq |\hat{\beta}|$ is the condition it judges. An empirical *p*-value below 0.05 indicates that the true DID coefficient significantly differs from the distribution of placebo coefficients under random assignment, further supporting the authenticity and robustness of the greenbelt policy's effect.

4. Results

4.1. Spatiotemporal Dynamics of RSEI

Figure 3 presents the annual changes in the RSEI means of the treatment and control groups from 2014 to 2023, where Figure 3a is a line chart of the annual means of the two groups and Figure 3b is a boxplot of the annual distributions of the two groups. During the study period, the RSEI of the treatment group exhibited a generally fluctuating upward trend, rising from 0.5231 in 2014 to 0.5894 in 2023, an increase of 12.66%. The lowest value in the study period occurred in 2016 (0.4896), followed by a gradual recovery and a peak in 2021 (0.5966). The index remained at a relatively high level of 0.57–0.60 from 2020 to 2023. The RSEI of the control group also increased slightly, from 0.5477 in 2014 to 0.5601 in 2023, an increase of only 2.26%.

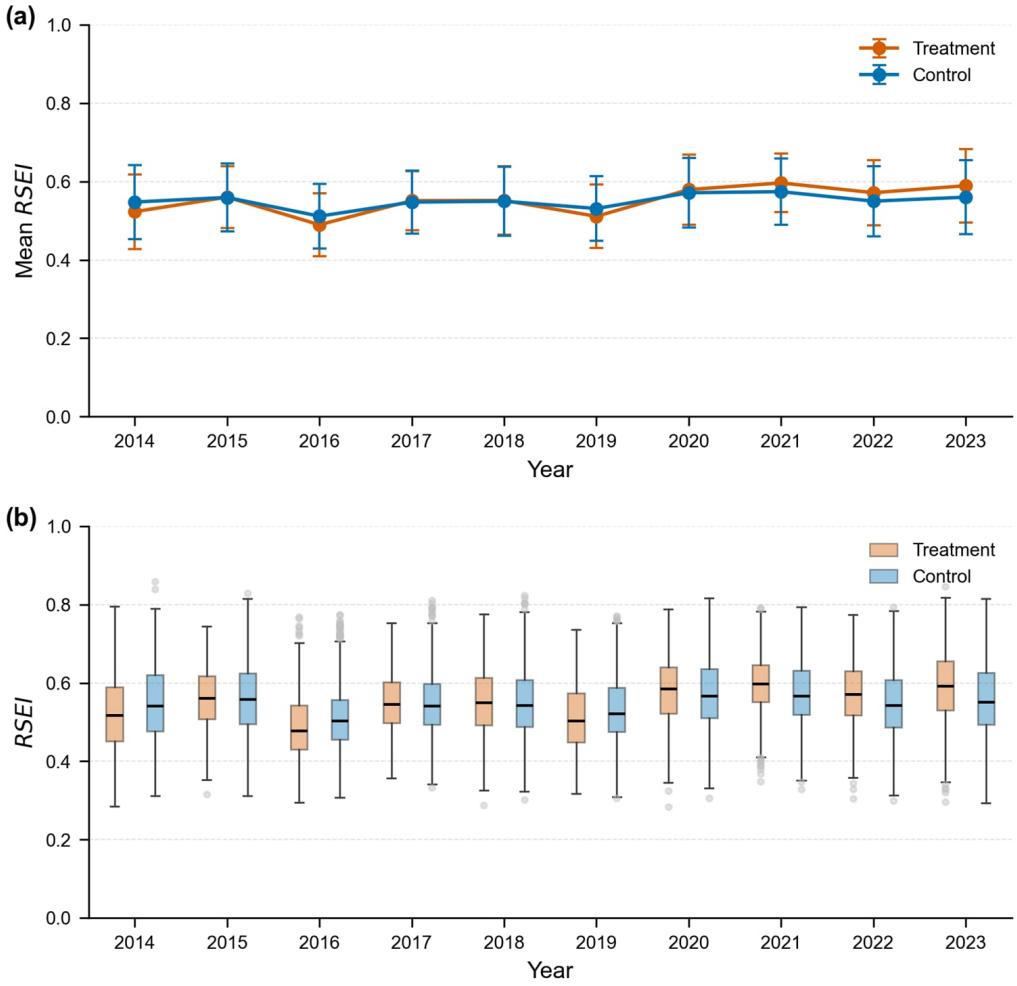

Figure 3. Annual mean RSEI trajectories of the treatment and control groups: (**a**) annual means of the two groups; (**b**) annual distributions of the two groups.

Comparing the two groups, the pre-policy mean RSEI of the treatment group (0.5352) was slightly lower than that of the control group (0.5432), with a between-group difference of −0.0080. In the post-policy period, the treatment group's mean rose to 0.5696, surpassing the control group (0.5574), and the between-group difference reversed to 0.0122. Year by year, the gap between the two groups was mostly close to zero or negative before 2020, whereas after 2020, the treatment group consistently exceeded the control group, with the gap widening from 0.0079 to 0.0292 by 2023. This indicates that the improvement in the ecological quality of the treatment group after policy implementation was substantially greater than that of the control group.

Figure 4 shows the RSEI spatial distribution in the study area between 2014 and 2023. During the study period, the average RSEI level within the TGEBZ increased from 0.5204 to 0.5876, and the proportion of high-value pixels (RSEI > 0.6) increased from 31.23% to 50.42%. These are pixel-level means within the TGEBZ, whereas Figure 3 averages the 716 treated and 1367 control grids, and the estimates use 495 matched pairs. Since 2020, the share of high-value pixels within the TGEBZ has consistently exceeded the overall level of the study area, indicating that the extent of high-ecological-quality areas within the TGEBZ expanded markedly and that ecological quality improved significantly in the latter half of the study period.

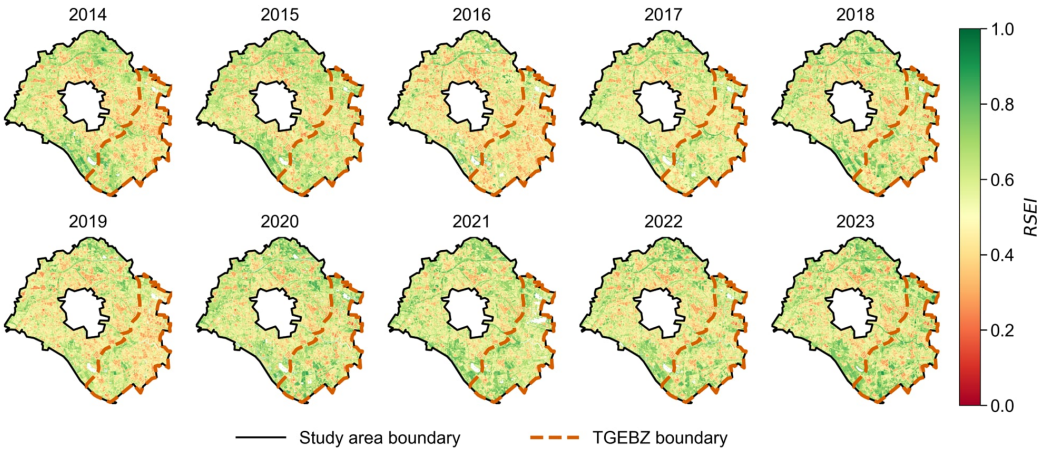

Figure 4. Spatial RSEI distribution snapshots 2014–2023.

4.2. Descriptive Statistics of Matched-Sample Variables

As shown in Table 2, the RSEI mean values of the treatment and control groups were 0.559 and 0.549, respectively, with the treatment group being significantly higher ($t = 5.18$, $p < 0.001$). The NLI of the treatment group was 11.893, which was significantly lower than the control group's value of 13.295 ($t = -6.25$, $p < 0.001$). The Tmean of the treatment group was 14.018, which was significantly higher than the control group's value of 13.974 ($t = 5.87$, $p < 0.001$). The differences in Prec, POI, and ISA between the two groups were all insignificant ($p > 0.1$), indicating that the two groups were largely comparable on most confounding characteristics after matching.

Table 2. Descriptive statistics of the main variables.

Variable	Obs.	Mean	SD	Min	Max	Treatment Mean	Control Mean	t	*p*
RSEI	9900	0.554	0.094	0.291	0.845	0.559	0.549	5.18	<0.001
Tmean	9900	13.996	0.373	13.029	14.944	14.018	13.974	5.87	<0.001
Prec	9900	49.085	8.138	34.669	66.521	49.183	48.988	1.19	0.234
NLI	9900	12.594	11.178	0.000	98.555	11.893	13.295	−6.25	<0.001
POI density	9900	0.081	1.444	0.000	75.000	0.103	0.059	1.54	0.124
ISA	9900	572,435.962	309,621.599	0.000	1,075,872.876	570,834.772	574,037.152	−0.51	0.607

Note: Obs. = number of grid-year observations; SD = standard deviation; RSEI = Remote Sensing Ecological Index; Tmean = annual mean air temperature (°C); Prec = monthly mean precipitation (mm); NLI = nighttime light index ($nW \cdot cm^{-2} \cdot sr^{-1}$); POI density = points of interest per square kilometer ($points/km^2$); ISA = impervious surface area (m^2). The matched sample comprised 990 grids × 10 years. The t and *p* values were obtained from Welch's two-sample *t*-test between the treatment and control groups.

4.3. PSM and Balance-Test Results

Before matching, there were 716 treatment grids and 1367 control grids. The propensity score p_i of the treatment group ranged from 0.0018 to 0.9825, and that of the control group ranged from 0.0002 to 0.8302, yielding the region of common support of 0.0018–0.8302. Within the region of common support, there were 632 treatment grids and 1366 control grids. After one-to-one nearest-neighbor matching with a 0.05 caliper constraint, 495 matched pairs were obtained, comprising 990 grids and 9900 panel observations. The mean matching distance was 0.0010, and the maximum matching distance was 0.0499 (below the caliper of 0.05).

Figure 5 presents the distributions of propensity scores for the treatment and control groups before (a) and after (b) matching. Before matching, the propensity score distributions of the two groups show some misalignment; after matching, the two distri-

butions overlap to a substantial degree, indicating that matching effectively improved the comparability of the treatment and control groups.

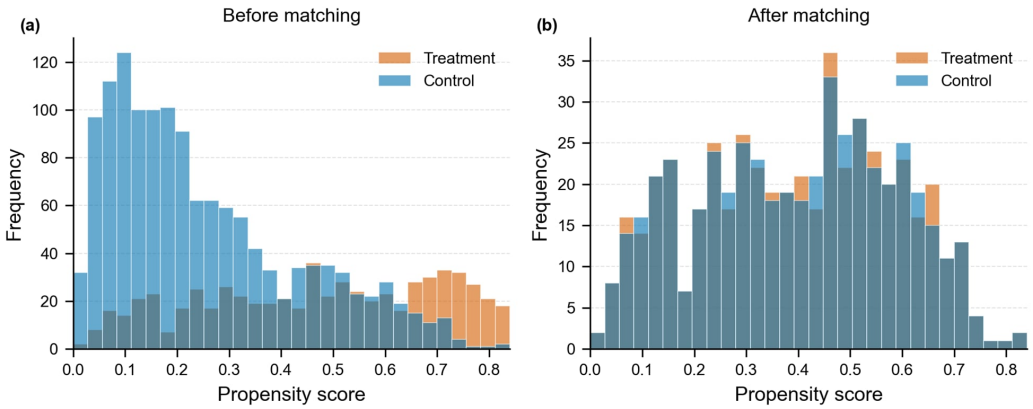

Figure 5. Propensity score distributions before (**a**) and after (**b**) matching.

As shown in Table 3, before matching, the standardized mean differences in the pre-policy RSEI trend, Prec, and NLI were relatively large. The absolute standardized mean difference (|SMD|) of Prec reached 0.833, and that of the pre-policy RSEI trend reached 0.468, indicating notable differences between the two groups in certain pre-policy characteristics. After matching, the |SMD| of all covariates fell below 0.1, with a mean |SMD| of 0.039 and a maximum |SMD| of 0.077, and the between-group mean differences in all covariates were statistically insignificant ($p > 0.1$). This indicates that the differences in the pre-policy characteristics between the two groups were substantially reduced and that the matched samples were highly comparable.

Table 3. Covariate balance test results before and after matching.

Matching Variable	Sample	Treatment Mean	Control Mean	|SMD|	t	*p*	Balanced
Pre-policy mean RSEI	Before	0.535	0.543	0.103	−2.27	0.024	No
	After	0.542	0.541	0.012	0.20	0.845	Yes
Pre-policy RSEI trend	Before	0.028	0.002	0.468	9.63	<0.001	No
	After	0.012	0.013	0.023	−0.36	0.718	Yes
Pre-policy Tmean	Before	13.973	13.957	0.057	1.20	0.231	Yes
	After	13.941	13.930	0.037	0.59	0.556	Yes
Pre-policy Prec	Before	48.050	47.310	0.833	18.35	<0.001	No
	After	47.742	47.800	0.077	−1.22	0.223	Yes
Pre-policy NLI	Before	9.751	12.561	0.292	−6.68	<0.001	No
	After	10.475	11.128	0.072	−1.13	0.261	Yes
Pre-policy POI density	Before	0.015	0.022	0.034	−0.69	0.492	Yes
	After	0.019	0.023	0.013	−0.21	0.834	Yes
Pre-policy ISA	Before	564,063.307	554,089.742	0.033	0.74	0.462	Yes
	After	551,533.200	563,699.368	0.039	−0.62	0.536	Yes

Note: |SMD| = absolute standardized mean difference; |SMD| < 0.1 indicates balance. The t-values and *p*-values were based on Welch's two-sample *t*-test. RSEI = Remote Sensing Ecological Index; Tmean = annual mean air temperature; Prec = precipitation; NLI = nighttime light index; POI = point of interest; ISA = impervious surface area.

4.4. Pre-Policy Parallel-Trends Test

The pre-policy parallel trends test used the 2014–2018 sample, with 990 grids and 4950 observations. The coefficient θ of the core variable Treat × Trend is 0.001 (SE = 0.001) with a *p*-value of 0.156, which is not significant at the 10% level. The test therefore does not reject the null of a common linear pre-policy trend, which is consistent with the parallel trends

condition for DID estimation. The test detects only a differential linear trend, and a five-year pre-policy window limits its power.

Figure 6 displays the mean RSEI trajectories of the treatment and control groups from 2014 to 2018 with 95% confidence interval bands. The mean trends of the two groups are largely parallel, with the treatment group being slightly higher than the control group but maintaining a stable gap, which corroborates the insignificant Treat × Trend result in Table 4.

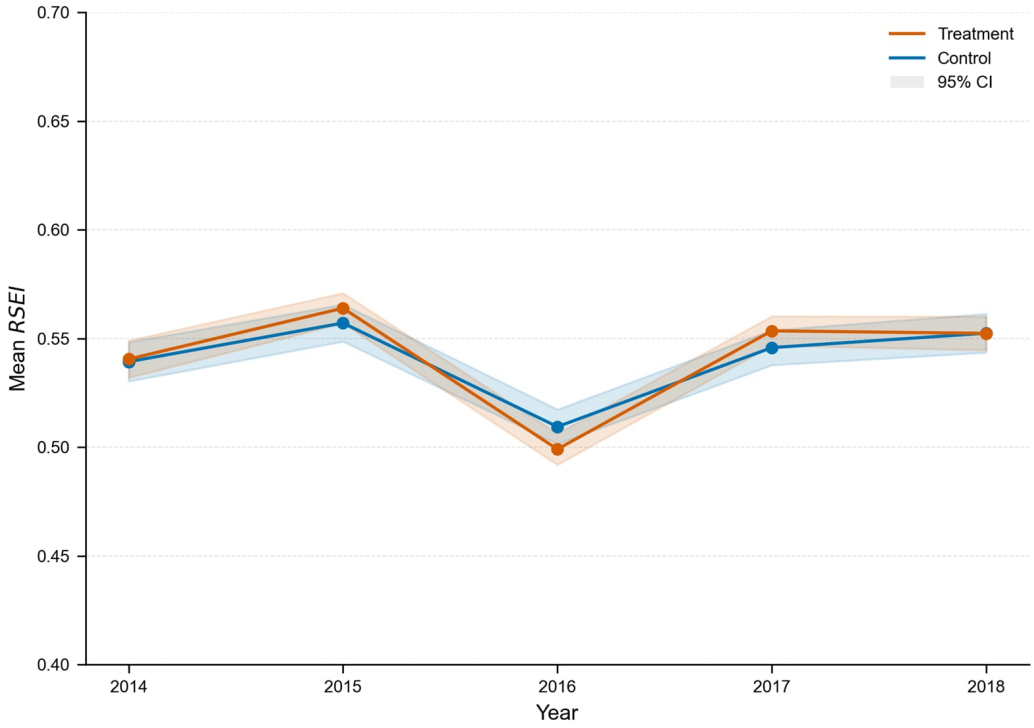

Figure 6. Pre-policy parallel-trends visualization.

Table 4. Pre-policy parallel-trends test results.

Variable	Coefficient	SE	t	*p*	95% CI
Treat × Trend	0.001	0.001	1.42	0.156	[−0.001, 0.003]
Tmean	0.022	0.004	5.37	<0.001	[0.014, 0.031]
Prec	0.007	0.001	6.73	<0.001	[0.005, 0.009]
NLI	0.001	0.000	4.02	<0.001	[0.001, 0.002]
POI density	0.003	0.001	5.17	<0.001	[0.002, 0.004]
ISA	-2.44×10^{-7}	4.28×10^{-8}	−5.69	<0.001	$[-3.28 \times 10^{-7}, -1.60 \times 10^{-7}]$

Note: SE = standard error; CI = confidence interval. Dependent variable = RSEI. The model includes grid-fixed effects and year-fixed effects, and the standard errors are clustered at the grid level. Observations = 4950; grids = 990. RSEI = Remote Sensing Ecological Index; Tmean = annual mean air temperature; Prec = precipitation; NLI = nighttime light index; POI = point of interest; ISA = impervious surface area. Treat × Trend = interaction between the treatment group dummy and linear time trend.

4.5. Baseline PSM-DID Regression Results

Table 5 reports the estimation results of the three models. In all three models, the coefficient β of the core explanatory variable DID was significantly positive. In Model 1 (no controls), β is 0.017 (SE = 0.003, t = 6.79); in Model 2 (predetermined controls), β is 0.016 (SE = 0.003, t = 5.89) with a 95% confidence interval of [0.011, 0.022]; and in Model 3 (contemporaneous controls), the benchmark specification, β is 0.022 (SE = 0.003, t = 7.91) with a 95% confidence interval of [0.016, 0.027]. All three coefficients were significant at

the 1% level. Taking Model 3 as the benchmark, TGEBZ planning increased the RSEI of the treatment grids by approximately 0.022 on average relative to the control grids. Model 2, which excludes controls that may respond to the planning, gives 0.016; the sign and significance of the estimate therefore do not depend on those controls, and the difference between the two coefficients reflects their inclusion. Among the control variables, the coefficient of ISA is significantly negative (-2.73×10^{-7}, $p < 0.001$), so that within a grid and year, a larger impervious area is associated with a lower RSEI. The remaining control variables are insignificant in Model 3.

Table 5. PSM-DID baseline regression results.

Variable	Model 1 (no Controls)	Model 2 (Predetermined Controls; Total Effect)	Model 3 (Contemporaneous Controls; Benchmark)
DID (Treat × Post)	0.017 ***	0.016 ***	0.022 ***
	(0.003)	(0.003)	(0.003)
Tmean		0.010	0.015
		(0.015)	(0.014)
Prec		0.001	0.000
		(0.001)	(0.001)
NLI			0.000
			(0.000)
POI density			−0.001
			(0.001)
ISA			-2.73×10^{-7} ***
			(2.64×10^{-8})
Grid fixed-effects	Yes	Yes	Yes
Year fixed-effects	Yes	Yes	Yes
Observations	9900	9900	9900
Number of grids	990	990	990
Within R^2	0.015	0.016	0.065

Note: Cluster-robust standard errors (grid-level clustering) are in parentheses. *** $p < 0.01$. Dependent variable = RSEI. RSEI = Remote Sensing Ecological Index; Tmean = annual mean air temperature; Prec = precipitation; NLI = nighttime light index; POI = point of interest; ISA = impervious surface area. DID = interaction between the treatment-group dummy and the post-policy dummy; Within R^2 = coefficient of determination computed on the data after both grid and year fixed effects have been removed.

4.6. Robustness Checks

As shown in Table 6 and Figure 7, β is significantly positive (ranging from 0.021 to 0.032) across all policy time window specifications (R0–R4), indicating that the core conclusion does not depend on a particular policy starting point or treatment of the transition years. Meanwhile, β is larger when the post-policy period is defined as 2020–2023 (R1 = 0.032; R4 = 0.027), suggesting a certain release lag of the policy effect. In terms of the time placebo tests (Figure 8), when 2016 is designated as the pseudo policy year, β equals 0.003 ($p = 0.319$), which is insignificant; when 2017 is designated as the pseudo policy year, β equals 0.005 ($p = 0.045$), which is significant at the 5% level. The significant 2017 placebo is consistent with the shape of the pre-policy gap rather than a sustained divergence: the gap between the two groups is +0.001 in 2014, +0.007 in 2015, −0.010 in 2016, +0.008 in 2017 and 0.000 in 2018, so the pseudo-policy year coincides with the 2016 dip and the rebound that follows it, while the end-to-end gap is essentially unchanged. At 0.005, the placebo coefficient is about one fifth of the baseline estimate, and dropping 2016 or 2017 leaves β at 0.020 and 0.024, both significant at the 1% level. Other pre-policy differences cannot be excluded, so parallel trends remain an identifying assumption.

Table 6. Robustness checks: policy-time-window specifications and time-placebo tests.

Specification	Description	β	SE	*p*	95% CI	Obs.
R0	Baseline (post-2019)	0.022 ***	0.003	<0.001	[0.016, 0.027]	9900
R1	Post-2020, keep 2019	0.032 ***	0.003	<0.001	[0.026, 0.037]	9900
R2	Drop 2019, post-2020	0.028 ***	0.003	<0.001	[0.022, 0.034]	8910
R3	Drop 2018, post-2019	0.021 ***	0.003	<0.001	[0.015, 0.027]	8910
R4	Drop 2018–2019, post-2020	0.027 ***	0.003	<0.001	[0.021, 0.034]	7920
P1	Placebo 2016	0.003	0.003	0.319	[−0.003, 0.010]	—
P2	Placebo 2017	0.005	0.002	0.045	[0.000, 0.009]	—

Note: Obs. = number of observations; β = DID coefficient; SE = standard error; CI = confidence interval. *** $p < 0.01$. R0–R4 are alternative real-policy-period definitions; P1–P2 are pre-policy placebo regressions. Standard errors are clustered at the grid level.

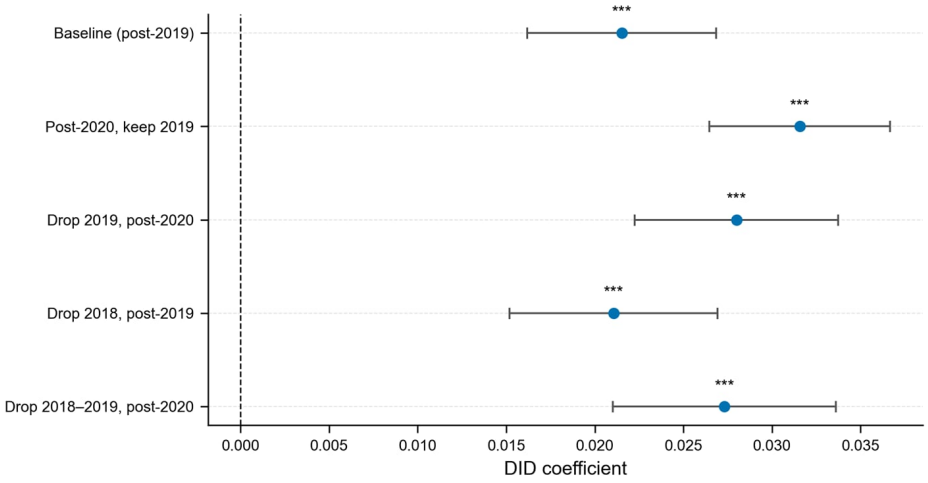

Figure 7. Time-window robustness results. *** $p < 0.01$.

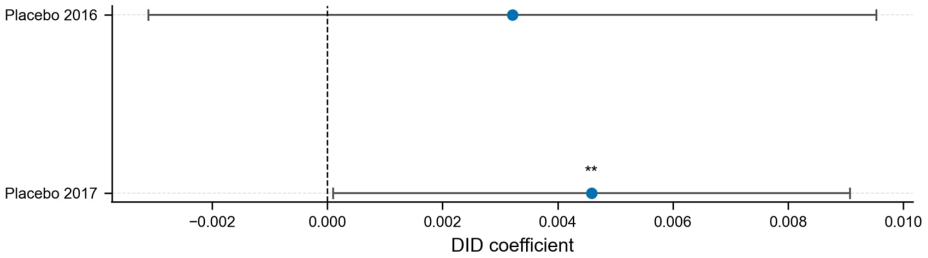

Figure 8. DID coefficients in time placebo tests. ** $p < 0.05$.

As shown in Table 7, after excluding each year from 2014 to 2023 in turn, β remains significantly positive, with coefficients ranging from 0.016 to 0.028. This indicates that the baseline results are robust and not driven by any particular year.

Table 7. Leave-one-year-out robustness check results.

Specification	β	Significance	Specification	β	Significance
Drop 2014	0.016	***	Drop 2019	0.028	***
Drop 2015	0.024	***	Drop 2020	0.022	***
Drop 2016	0.020	***	Drop 2021	0.018	***
Drop 2017	0.024	***	Drop 2022	0.022	***
Drop 2018	0.021	***	Drop 2023	0.018	***

Note: β = DID coefficient; *** $p < 0.01$. Each specification re-runs baseline Model 3, excluding one-year observations.

To further rule out the possibility that the true DID coefficient reflects a chance occurrence, the treatment group labels were randomly shuffled, and the DID coefficients were re-estimated 300 times. As shown in Table 8 and Figure 9, the placebo DID coefficients from 300 random permutations are symmetrically distributed around zero (mean = 0.000, SD = 0.002, range = −0.007 to 0.006), whereas the true DID coefficient of 0.022 lies beyond the extreme-right tail of this distribution. The empirical two-tailed *p*-value is 0.003, which is below 0.05, indicating that the true policy effect significantly differs from the distribution of placebo coefficients under random assignment. This further supports the authenticity and robustness of the policy's effect.

Table 8. Randomization-placebo-test summary statistics.

Statistic	Value
True DID coefficient (Model 3)	0.022
Mean placebo coefficient	0.000
SD of placebo coefficients	0.002
Min placebo coefficient	−0.007
Max placebo coefficient	0.006
5th percentile of placebo coefficients	−0.004
95th percentile of placebo coefficients	0.004
Number of permutations	300
Empirical two-tailed *p*-value	0.003

Note: SD = standard deviation. Placebo coefficients were obtained from K = 300 random reassignments of treatment labels, with the number of treated grids held constant. The empirical two-tailed *p*-value was computed following Equation (24).

Additional details are provided regarding the timing of the policy implementation. Preparatory and pilot construction began in parts of the zone in 2018. As the plan was approved and fully enacted in 2019, the baseline treats 2018 as a pre-policy year. Under the full-control specification, the estimate remains 0.021 whether 2018 is retained as pre-policy or excluded as a transition year (SE = 0.003 in both), and falls to 0.017 when 2018 is coded as post-policy (SE = 0.003); all three are significant at the 1% level. Therefore, the main conclusion does not depend on how year 2018 is classified.

Further checks address the sensitivity of the estimate to spatial dependence and the definition of the treatment group. Conley standard errors that allow for correlation between grids within 5–20 km raise the standard error of β from 0.003 to 0.005–0.006, leaving the estimate significant at the 1% level (t = 3.9–4.6). Excluding control grids within 1.5, 3 and 5 km of a treated grid raises β from 0.022 to 0.023, 0.024 and 0.025, all significant at the 1% level, and the direction of this change is consistent with positive spillovers attenuating rather than inflating the estimate. Coverage is high across the matched sample—the median is 1.00 and the mean 0.88—so the alternative definitions differ mainly in the 61 treated grids whose coverage is below 0.5. Requiring at least 50%, 75% or 90% of a grid to lie inside the zone reduces the sample to 929, 911 and 892 grids and leaves β within 0.0215–0.0222 (SE = 0.0029–0.0030, significant at the 1%); a continuous specification that interacts the coverage ratio with the post-policy dummy yields β = 0.020 (SE = 0.003). The estimate therefore does not depend on how much of a grid lies inside the boundary.

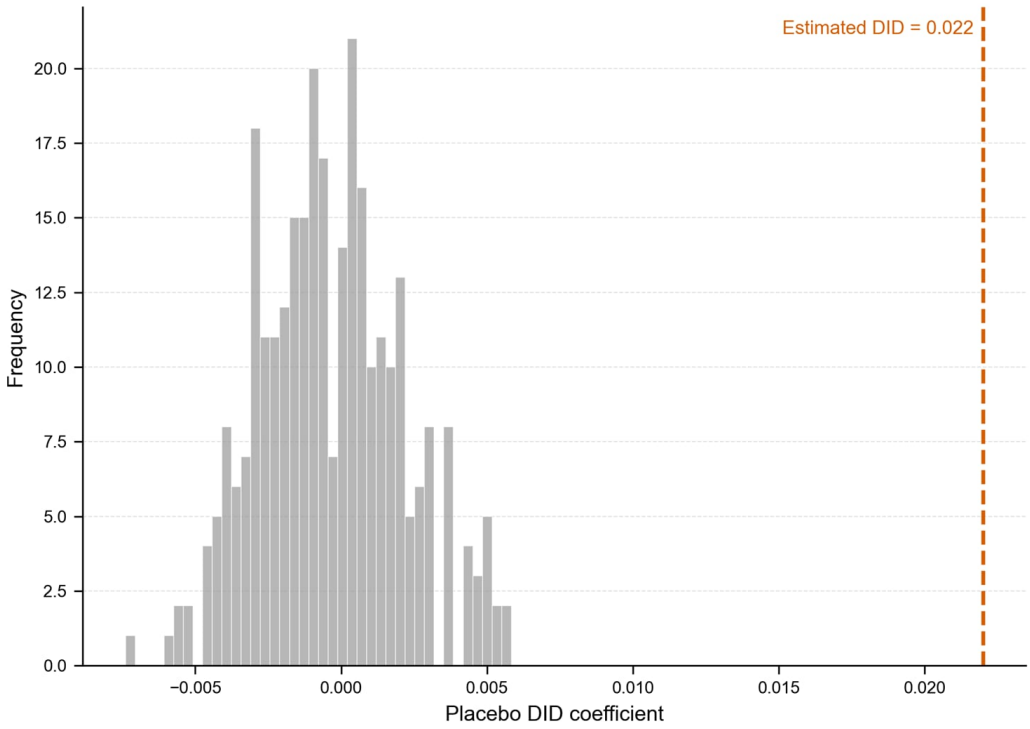

Figure 9. Randomization-placebo permutation distribution.

5. Discussion

Based on the RSEI-based ecological quality measurement and PSM-DID causal inference, this study provides evidence that TGEBZ planning significantly improves ecological quality within the zone, and this conclusion remains valid under multiple robustness checks, thereby corroborating the ecological effectiveness of greenbelts as spatial regulation instruments. Unlike existing studies that focus on the role of greenbelts in curbing urban sprawl [2], this study extends the perspective to track the temporal changes in regional ecological quality, directly addressing how greenbelt planning improves regional ecological conditions. In contrast to evaluations built on a single vegetation indicator such as NDVI [11], the RSEI integrates greenness, wetness, dryness, and heat, avoiding the one-sided depiction of ecological states inherent in single indices. Here, the four components did not respond alike: the gain is carried mainly by a reduction in the heat component, whereas greenness, wetness, and dryness changed by amounts that cannot be separated from zero; the treated grids have therefore become relatively cooler rather than greener or wetter. Theoretically, this study converts the greenbelt from a normative planning ideal into a testable causal proposition, examining whether continuous green space separating built-up areas genuinely preserves ecological function.

From a methodological standpoint, this study shares the PSM-DID causal inference framework with policy evaluation studies, such as those on low-carbon and smart cities [18,19,21], but extends it from policy evaluation to greenbelt planning centered on ecological space regulation. Moreover, identification was conducted at the 1 km grid scale, which is finer than that of previous studies using administrative units as the unit of analysis. Finally, this study applies the RSEI framework [12,13], extending it from the spatiotemporal description of ecological indices to causal policy evaluation and thereby bridging, to a certain extent, the gap between remote sensing measures of ecological quality and the identification of policy effects. Two assumptions underlie the estimate, namely comparability of matched grids and parallel trends; both are supported only indirectly by the balance and placebo tests, so residual confounding cannot be excluded.

Regarding effect timing, a certain release lag of the policy effect was observed, with estimates being substantially larger when the post-policy period was defined as 2020–2023 than when 2019 was taken as the policy starting point. This finding resonates with the results of Wang et al. concerning the time lag of low-carbon city pilot policy effects [19]. The underlying reason is that ecological processes, such as vegetation growth and water body recovery, are inherently gradual, and planning implementation and construction disturbances in the early stage may partially offset some ecological dividends, with ecological benefits emerging only in subsequent years. In terms of robustness, the pre-policy parallel-trends test did not reject a common linear pre-trend, all standardized mean differences after matching were below 0.1, and the randomization test was significant.

In terms of policy implications, the positive ecological benefits of TGEBZ planning indicate that spatial regulation through the construction of metropolitan green buffers is a promising governance path for addressing ecological degradation in megacities. Tianjin should continue to advance TGEBZ construction, maintain the stability of ecological space regulation, and prevent the rebound of development intensity from weakening ecological improvement effects. At the same time, Tianjin should continue to strengthen the regulation of construction land boundaries and control surface hardening to consolidate the existing ecological improvement achievements. Given the release lag of ecological effects, a regularized and intelligent ecological performance evaluation system should be established based on long-term RSEI monitoring technology to continuously track the ecological quality dynamics of the TGEBZ and surrounding areas and provide a scientific basis for ecological restoration, planning optimization, and policy adjustment.

Globally, numerous contemporary megacities, particularly in America and Asia, encounter shared challenges, such as spatial fragmentation and expansive contiguous development, with less attention paid to adequate ecological buffer zones as an important spatial intervention [29,30]. The TGEBZ case study indicates that proactively delineating greenbelts and constructing ecological buffer corridors between urban clusters may help contain disorderly urban expansion and balance urban ecological patterns, and that ecological quality inside the zone improves measurably after the plan is implemented. When undertaking ecological planning, similar cities can draw on the whole-process governance model of "spatial delineation–rigid control–dynamic monitoring–performance evaluation," abandon the governance misconceptions that emphasize construction over maintenance and favor short planning cycles, fully respect the laws of ecosystem evolution, value the long-term cumulative effects of ecological policies, and quantify policy effectiveness through rigorous empirical methods. Thus, orderly urban spatial development and continuous improvement in ecological quality can be achieved simultaneously.

Finally, several limitations of this study point to directions for future work. First, the RSEI calculation relies on multi-scene Landsat imagery and can be disturbed by residual cloud contamination and seasonal-phase differences. Because the imagery covers only May–October and the scene numbers vary by year, part of the interannual variation in RSEI may reflect phenological differences rather than ecological change, although within-year averaging reduces this effect. Follow-up work can further unify image seasonal-selection windows and strengthen outlier-filtering to raise monitoring consistency and precision. Second, the TGEBZ is a long-term ecological governance program, whereas this study covers only five years before and after the policy intervention. Future research may extend the time-series length to track the long-term sustainability of policy-induced ecological gains. Third, the RSEI describes the surface ecological condition captured by four remotely sensed components and does not measure biodiversity, habitat quality, or ecosystem services; therefore, the findings are read as evidence of ecological condition in that specific sense; extending the index to those attributes is left for future work.

6. Conclusions

Using the Tianjin Green Ecological Barrier Zone (TGEBZ) as a case study, this study developed a panel dataset at a 1 km grid level for the period of 2014–2023. The RSEI was used to assess ecological quality, and the PSM-DID model was applied to evaluate the impact of greenbelt planning on regional ecological quality.

Four remote-sensing components were compressed into an annual RSEI by principal component analysis, and 716 grids inside the zone and 1367 outside were matched one-to-one with a 0.05 caliper, retaining 495 pairs, or 990 grids and 9900 grid-year observations, with all standardized mean differences below 0.077. Pre-policy trends did not differ between the two groups ($\theta = 0.001$, $p = 0.156$), and the effect was estimated using a two-way fixed-effects difference-in-differences model with grid-clustered standard errors.

From 2014 to 2023, ecological quality inside the TGEBZ improved markedly: the RSEI rose from 0.5231 to 0.5894 (+12.66%), while the grids outside the zone increased by only 2.26%. The lowest value was recorded in 2016, followed by a gradual recovery and a peak in 2021, and the index remained relatively high from 2020 to 2023. The planning is estimated to have raised the RSEI inside the zone by 0.022 ($p < 0.01$, 95% CI [0.016, 0.027]); restricting the controls to predetermined climate variables gives 0.016. The estimate remains robust across a battery of robustness checks.

These findings indicate that spatial regulation through greenbelt planning is an effective instrument for improving ecological quality in a large metropolitan region, and they call for institutionalizing long-term ecological monitoring to track policy performance. The study also offers a transferable RSEI-PSM-DID paradigm for evaluating ecological governance policies in comparable cities.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/land15122289/s1, Table S1: Annual means of the scene-level PC1 variance share and loadings of the RSEI, 2014–2023.

Author Contributions: Conceptualization, Y.W.; methodology, Y.W., C.H. and Z.Z.; software, C.H.; validation, Y.W. and C.H.; formal analysis, C.H.; investigation, Y.W.; resources, Z.Z.; data curation, C.H.; writing—original draft preparation, Y.W. and C.H.; writing—review and editing, Y.W. and Z.Z.; visualization, C.H.; supervision, Z.Z.; project administration, Z.Z.; funding acquisition, Z.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ministry of Natural Resources, China (No. 2023ZRB-SHZ075), Natural Science Foundation of China (No. 52678103) and Open Funds of Tianjin Key Laboratory of Smart City Planning (No. GHKF-202402).

Data Availability Statement: Data will be available on request due to restrictions (any request requires approval from the project funder).

Acknowledgments: The authors appreciate the anonymous reviewers for their constructive comments on the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. United Nations. *World Urbanization Prospects 2025: Summary of Results*; UN DESA/POP/2025/TR/NO. 12; United Nations: New York, NY, USA, 2025. Available online: https://desapublications.un.org/publications/world-urbanization-prospects-2025-summary-results (accessed on 9 November 2026).
2. Pourtaherian, P.; Jaeger, J.A.G. How effective are greenbelts at mitigating urban sprawl? A comparative study of 60 European cities. *Landsc. Urban Plan.* **2022**, *227*, 104532. [CrossRef]

3. Mosharafian, S.; Jaeger, J.A.G. Proposing targets and limits to urban sprawl: How likely are current greenbelt scenarios for Montreal to achieve proposed reference values by 2070? *Environ. Manag.* **2026**, *76*, 45. [CrossRef]
4. Han, A.T.; Daniels, T.L.; Kim, C. Managing urban growth in the wake of climate change: Revisiting greenbelt policy in the US. *Land Use Policy* **2022**, *112*, 105867. [CrossRef]
5. Ma, M.; Jin, Y. What if Beijing had enforced the 1st or 2nd greenbelt?—Analyses from an economic perspective. *Landsc. Urban Plan.* **2019**, *182*, 79–91. [CrossRef]
6. Yuan, Y. Analysis on urban greenbelt plan implementation: An empirical study on the outer ring road greenbelt in Shanghai. *City Plan. Rev.* **2019**, *43*, 58–66. [CrossRef]
7. Kim, J.I.; Hyun, J.Y.; Lee, S.G. The effects of releasing greenbelt restrictions on land development in the case of medium-sized cities in Korea. *Sustainability* **2019**, *11*, 630. [CrossRef]
8. Han, H.; Huang, C.; Ahn, K.-H.; Shu, X.; Lin, L.; Qiu, D. The effects of greenbelt policies on land development: Evidence from the deregulation of the greenbelt in the Seoul metropolitan area. *Sustainability* **2017**, *9*, 1259. [CrossRef]
9. Gong, Y.; Zhou, L.; Sun, D.; Chen, Z. Measurement and analysis of fragmentation and connectivity of green belts in Chinese megacities from a resilience perspective: A case study of Beijing, Xi'an and Chengdu. *Sci. Geogr. Sin.* **2023**, *43*, 1195–1205. [CrossRef]
10. Hwang, J.; Lee, K.; Yoo, Y.; Kim, Y.; Jeon, S.W.; Choi, Y. Greenbelt's impact on urban forest habitat: Insights from morphological analysis in quantifying conservation effects. *Urban For. Urban Green.* **2024**, *97*, 128360. [CrossRef]
11. Gong, F.-Y.; Wang, C. Evaluating the performance of the greenbelt policy in Beijing using multi-source long-term satellite observations from 2000 to 2020. *Remote Sens.* **2023**, *15*, 4766. [CrossRef]
12. Xu, H. A remote sensing index for assessment of regional ecological changes. *China Environ. Sci.* **2013**, *33*, 889–897.
13. Xu, H. A remote sensing urban ecological index and its application. *Acta Ecol. Sin.* **2013**, *33*, 7853–7862. [CrossRef]
14. Xiao, X.; Peng, X. Spatiotemporal changes in Chinese reserve vegetation ecological quality from 2002 to 2022. *Pol. J. Environ. Stud.* **2025**, *34*, 8339–8350. [CrossRef]
15. Hu, N.; Wang, G.; Ma, Z.; Zhao, M.; Yuan, Y.; Zhang, T.; Chen, Q.; Meng, J.; Wang, J. Multi-scenario simulation of land use dynamics and ecological risk: A case study of the Liaohe estuary national wetland reserve using PLUS-Markov and PSR models. *Front. Environ. Sci.* **2024**, *12*, 1361970. [CrossRef]
16. Chang, Y.; Zou, T.; Yoshino, K.; Luo, S.; Zhou, S. Ecological policy benefit valuation based on public feedback: Forest ecosystem services in Wuyishan nature reserve, China. *Sci. Total Environ.* **2019**, *673*, 622–630. [CrossRef]
17. Fan, C.; Gai, Z.; Shen, S.; Sun, S. An implementation evaluation framework of ecological spatial planning based on multi-dimensional data: A case study in China. *Urban For. Urban Green.* **2021**, *63*, 127222. [CrossRef]
18. Fu, Y.; He, C.; Luo, L. Does the low-carbon city policy make a difference? Empirical evidence of the pilot scheme in China with DEA and PSM-DID. *Ecol. Indic.* **2021**, *122*, 107238. [CrossRef]
19. Wang, L.; Shao, J.; Ma, Y. Does China's low-carbon city pilot policy improve energy efficiency? *Energy* **2023**, *283*, 129048. [CrossRef]
20. Luo, C.; Qiang, W.; Lee, H.F. Does the low-carbon city pilot policy work in China? A company-level analysis based on the PSM-DID model. *J. Environ. Manag.* **2023**, *337*, 117725. [CrossRef]
21. Chen, Y.; Chen, S.; Miao, J. Does smart city pilot improve urban green economic efficiency: Accelerator or inhibitor. *Environ. Impact Assess. Rev.* **2024**, *104*, 107328. [CrossRef]
22. Wang, Q.; Liu, Y. Beautifying urban environment: Smart city construction and sustainable pollution control in China. *J. Environ. Manag.* **2024**, *371*, 123262. [CrossRef]
23. 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*, 3907–3925. [CrossRef]
24. Chen, Z.; Liao, L.; Wang, C.; Shi, K.; Yu, B. The 1992-2024 global NPP-VIIRS-like nighttime light annual data from deep learning super-resolution reconstruction. *J. Remote Sens.* **2026**, *6*, 0874. [CrossRef]
25. Peng, S. *1-km Monthly Mean Temperature Dataset for China (1901–2025) [Dataset]*; National Tibetan Plateau/Third Pole Environment Data Center: Beijing, China, 2019. Available online: https://cstr.cn/18406.11.Meteoro.tpdc.270961 (accessed on 18 November 2026). [CrossRef]
26. Peng, S. *1-km Monthly Precipitation Dataset for China (1901–2025) [Dataset]*; National Tibetan Plateau/Third Pole Environment Data Center: Beijing, China, 2020. Available online: https://www.tpdc.ac.cn/zh-hans/data/faae7605-a0f2-4d18-b28f-5cee413766a2 (accessed on 18 November 2026).
27. Zhai, Y.; Roy, D.P.; Martins, V.S.; Zhang, H.K.; Yan, L.; Li, Z. Conterminous United States Landsat-8 top of atmosphere and surface reflectance tasseled cap transformation coefficients. *Remote Sens. Environ.* **2022**, *274*, 112992. [CrossRef]
28. Xu, H.; Ren, M.; Yang, L. Evaluating the consistency of surface brightness, greenness, and wetness observations between Landsat-8 OLI and Landsat-9 OLI2 through underfly images. *Int. J. Appl. Earth Obs. Geoinf.* **2023**, *124*, 103546. [CrossRef]

29. Rukmana, D.; Adharina, N.D. The urgency of planning megacities in the Global South in the next decade. *Town Plan. Rev.* **2025**, *96*, 614–625. [CrossRef]
30. Li, X.; Zhang, J.; Huang, J.; Lin, W.; Wu, S.; Ma, M. To preserve green buffer under polarization and diffusion effects of a fast-developing megalopolis. *Land* **2022**, *11*, 724. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.