sustainability

MDPI

Article

Cropland Zoning Based on District and County Scales in the Black Soil Region of Northeastern China

Yong Li [1], Liping Wang [2,3], Yunfei Yu [4], Deqiang Zang [5], Xilong Dai [6] and Shufeng Zheng [1,*]

1 School of Government, Heilongjiang University, Harbin 150080, China; liyong2023@iga.ac.cn
2 State Key Laboratory of Black Soils Conservation and Utilization, Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Changchun 130102, China; wangliping@iga.ac.cn
3 University of Chinese Academy of Sciences, Beijing 100049, China
4 Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Changchun 130102, China; yuyunfei@iga.ac.cn
5 School of Public Administration and Law, Northeast Agricultural University, Harbin 150030, China; s221202036@neau.edu.cn
6 College of Geographical Science, Harbin Normal University, Harbin 150025, China; dxlhsd@stu.hrbnu.edu.cn
* Correspondence: 2009020@hlju.edu.cn; Tel.: +86-150-0467-6791

Abstract: The black soil region of northeastern China, one of the world's major black soil belts, is China's main grain-producing area, producing a quarter of China's commercial grain. However, over-exploitation and unsustainable management practices have led to a steady decline in the quality of arable land. Scientific and reasonable zoning of arable land is the key to ensuring that black soil arable land achieves sustainable development. In this study, the 317 districts and counties under the jurisdiction of Heilongjiang, Jilin, and Liaoning Provinces in the northeast region and the four eastern leagues of the Inner Mongolia Autonomous Region were taken as the study area, and arable land zoning in the northeast black soil region was explored through group analysis. Ten types of indicators were selected according to the four levels of climate, soil, vegetation, and topography of the northeast black soil region, including average precipitation and average temperature for many years at the climate level, organic matter content and soil texture (including clay, silt, and sand) at the soil level, NDVI and EVI indicators at the vegetation level, and DEM and slope indicators at the topographic level. In accordance with the principle of distinguishing differences and summarizing commonalities, nine scenarios of dividing the northeast black soil zones into 2 regions to 10 regions were explored, and these nine zoning scenarios were evaluated in terms of zoning. The results showed that (1) the spatial variability of cropland zoning in the northeast black soil zone based on four indicators, namely climate, soil, vegetation, and topography, was significant; (2) the results of the nine types of zoning based on cropland in the northeast black soil zone showed that intra-zonal zoning was optimal when zoning the northeast black soil zone into six types of zones, which enhanced the variability between the zones and the consistency within the zones; and (3) the assessment of large-scale cropland zoning using the pseudo F-statistic and area-weighted standard deviation methods revealed similarities in their outcomes. The results provide a scientific basis for the subregional protection of arable land in the black soil zone and help to formulate effective policies for different regions.

Keywords: district and county scales; spatial heterogeneity; northeastern black soil region; zoning

Citation: Li, Y.; Wang, L.; Yu, Y.; Zang, D.; Dai, X.; Zheng, S. Cropland Zoning Based on District and County Scales in the Black Soil Region of Northeastern China. *Sustainability* **2024**, *16*, 3341. https://doi.org/10.3390/su16083341

Academic Editor: Kęstutis Venslauskas

Received: 21 March 2024
Revised: 11 April 2024
Accepted: 12 April 2024
Published: 16 April 2024

1. Introduction

The northeast black soil area is one of the major black soil belts in the world, serving as the main production area for grain crops in China and playing a crucial role in preserving national food security [1]. Often referred to as the "giant panda" among arable lands, the commercial grain output from the black soil in northeast China alone accounts for a quarter of the national output [2]. However, due to continuous agricultural activities and irrational land management practices, this region is facing serious ecological challenges, with the black soil experiencing problems of "thinning, leaning, hardening, and acidification", which significantly impact the maintenance of regional and even national food production [3–6]. Owing to the diversity and complexity of black soil in northeast China with respect to climatic characteristics, soil texture, vegetation types, topographic distribution, and other endowments, there are variations in the level of fertility, utilization, and degradation of black soil across the region [7]. To achieve sustainable development of black soil in northeast China, it is imperative to adopt scientific and reasonable methods for effective zoning based on the characteristics of different geographical areas and their respective challenges and to develop black soil protection and utilization techniques tailored to the specific needs of each region. By doing so, we can overcome the obstacles to sustainable agricultural development and ensure that the black soil of northeast China achieves sustainable development on a large scale, in multiple dimensions, and at a deep level.

Currently, significant disparities exist in the scale of zoning. At the global scale, urban eco-zone mapping has been completed [8,9], including further analyzing the feedback mechanisms of eco-zones on human impacts and reflecting new trends in zoning research [10]. These zoning results have enhanced the understanding of geospatial differentiation and provided an important reference for comprehensive geographic zoning. However, these studies predominantly focus on zoning at the global scale, making it challenging to characterize the regional scale and adapt to the needs of optimizing and regulating the changing geographical system. At the field scale, previous zoning management has concentrated on precise management zoning, but rarely at the district and county scales.

Regarding zoning indicators, previous zoning indicators have been relatively one-dimensional. Some studies have utilized wheat yield data to divide farmland into SOM zones [11], while others have conducted zoning through remote sensing monitoring of crop growth and nutritional diagnostic indicators [12] or based on multi-year yield data [13]. Additionally, it has been shown that there is spatial heterogeneity and spatial correlation in SOM at small scales [14]; the analysis confirmed that incorporating topographic factors for spatial interpolation of SOM improves the global prediction accuracy R2 to 0.75. By comparing NDVI with the enhanced vegetation index (EVI), the analysis verified the effectiveness of both NDVI and EVI in assessing vegetation sensitivity and tree diversity recognition [15–17]. The significant role of NDVI in regional detection, based on its characteristics, was clarified, providing strong support for zoning based on NDVI [18,19].

In the zoning model, the coupled coordination model was employed to assess the relationship between black soil protection and regional development [20,21]. Soil nutrient data, including total nitrogen, effective phosphorus, and quick-acting potassium, were extracted using principal component analysis (PCA), and the specific management areas for rice cultivation were delineated through interpolation based on the K-means fuzzy clustering method [22–24]. Utilizing the C-means fuzzy clustering method, the study on management zoning of soil nutrients in oasis cotton fields in Xinjiang confirmed the validity and rationality of the zoning results [25–28].

From the zoning studies discussed above, it is evident that there are two primary issues: the scale of analysis (either too large or too small) and the neglect of district and county scales. Studies at overly large scales frequently encompass multiple land categories, failing to consider cropland independently. Conversely, studies at excessively small scales overly concentrate on specific fields, leading to an imbalance between scale and land categorization. To address this issue, priority should be given to zonal targeting. Global-scale studies do not partition cropland distinctly, whereas field-scale studies, being too limited, focus exclusively on a handful of fields. To achieve a relative balance, a combination of study scales and land types ought to be considered. Moreover, certain zoning results demonstrate fragmentation. Previous grid-based, large-scale zoning produced fragmented outcomes, unlike district- and county-scale studies, which tend to yield more cohesive zoning results. Within the county-dominated management framework in China in particular, this approach is more feasible for promotion and implementation.

Another issue is the lack of reliable evaluation of zoning results at large scales. Many zoning evaluations at the field scale frequently rely on yield; however, yield alone may not offer an accurate and reasonable assessment of zoning effectiveness. The assessment of zoning effectiveness should consider a variety of factors, not solely yield. The effectiveness of zoning is related not only to yield but also a range of other factors, including soil quality, water use efficiency, fertilizer utilization, pest control, and agroecosystem sustainability. Therefore, relying solely on yield to evaluate the effectiveness of zoning may overlook other crucial factors.

The challenges confronting cropland management at the district and county scales are especially pronounced. Firstly, land management at these scales frequently lacks effective long-term planning and scientific guidance, resulting in inefficient utilization of cropland and, in some cases, serious wastage of land resources. Secondly, significant variations in natural conditions, economic development levels, and agricultural production methods exist between districts and counties, necessitating the adaptation of cropland management strategies to local conditions—a task often challenging to accomplish in practice. Furthermore, the absence of effective information exchange and technical support has hindered many districts and counties in implementing arable land protection and sustainable utilization. Finally, with the growing impact of global climate change, there is an urgent need to adapt agricultural production patterns and cropland management strategies in the northeastern black soil zone to these changes.

In the practice of black soil arable land protection and utilization, it is imperative to fully comprehend the "territoriality" of these efforts, thoroughly analyze the spatial differentiation characteristics of black soil arable land protection, and, based on this analysis, implement differentiated protection and utilization measures in subregions and sub-types to circumvent the longstanding issue of homogenization in black soil protection and utilization efforts.

Therefore, this study employs 10 types of indicators across four dimensions—climate, soil, vegetation, and topography—as data sources, utilizing the method of grouping analysis to conduct cropland zoning in the northeast black soil region. This study proposes the following hypotheses: 1. cropland zoning at the district and county scales can be effectively achieved using the multi-indicator approach; and 2. the pseudo F-statistic and area-weighted standard deviation are capable of identifying the optimal management zoning method and developing a comprehensive evaluation methodology for large-scale cropland management zoning.

2. Materials and Methods

2.1. Study Area

The study area, situated in northeastern China between longitudes 135°08′ and 115°52′ E and latitudes 38°72′ and 53°56′ N, encompasses Heilongjiang, Jilin, and Liaoning Provinces, along with the "Four Eastern Leagues" (Chifeng, Tongliao, Hulunbeier Cities, and Xing'an League) in the Inner Mongolia Autonomous Region (Figure 1). This area includes 317 counties (cities, districts, and banners) across the Hulunbeier Grassland, both large and small Xing'anling, the Three Rivers Plain, the Songnen Plain, parts of the Songliao Plain, and the Changbai Mountain region. Its total area of 1,090,000 square kilometers represents approximately 12% of the global black soil zone's total area [29]. Regarding the definition of black soil, the term lacks a uniform global definition. In the World Reference Base (WRB) classification, the majority of black soils correspond to Chernozems, Kastanozems, and Phaeozems [30]. In the United States and Argentina, black soil correspond to the Mollisols Order according to the Soil Taxonomy [31]. In Ukraine, these soil types are included in a group characterized by humus-accumulative soil formation, primarily Chernozems, comparable to the "black earths" of the Russian Federation. Numerous regional variants exist; for instance, in China, what were originally termed "black soils" are now classified as "Isohumisols" according to the Chinese Soil Taxonomy [32]. In China, the term "black soil" typically refers to the soil found in the northeast region characterized by a deep, black topsoil layer rich in organic matter and predominantly located in the northeast plains. The northeast black soil region features a basin-like topography, comprising a central plain surrounded by mountains on three sides. The region's terrain includes mountains, plains, hills, and terraces, with 94.84% of its arable land located in plains and gently sloped areas of 7° or less. Northeast China is the country's primary production area for corn, soybeans, wheat, and rice. Since the early 20th century, marked by large-scale immigration and land resettlement, agricultural development and utilization intensity in the northeast black soil area have progressively increased. This is evidenced by the expanding sown area for crops and heightened inputs of agricultural production materials.

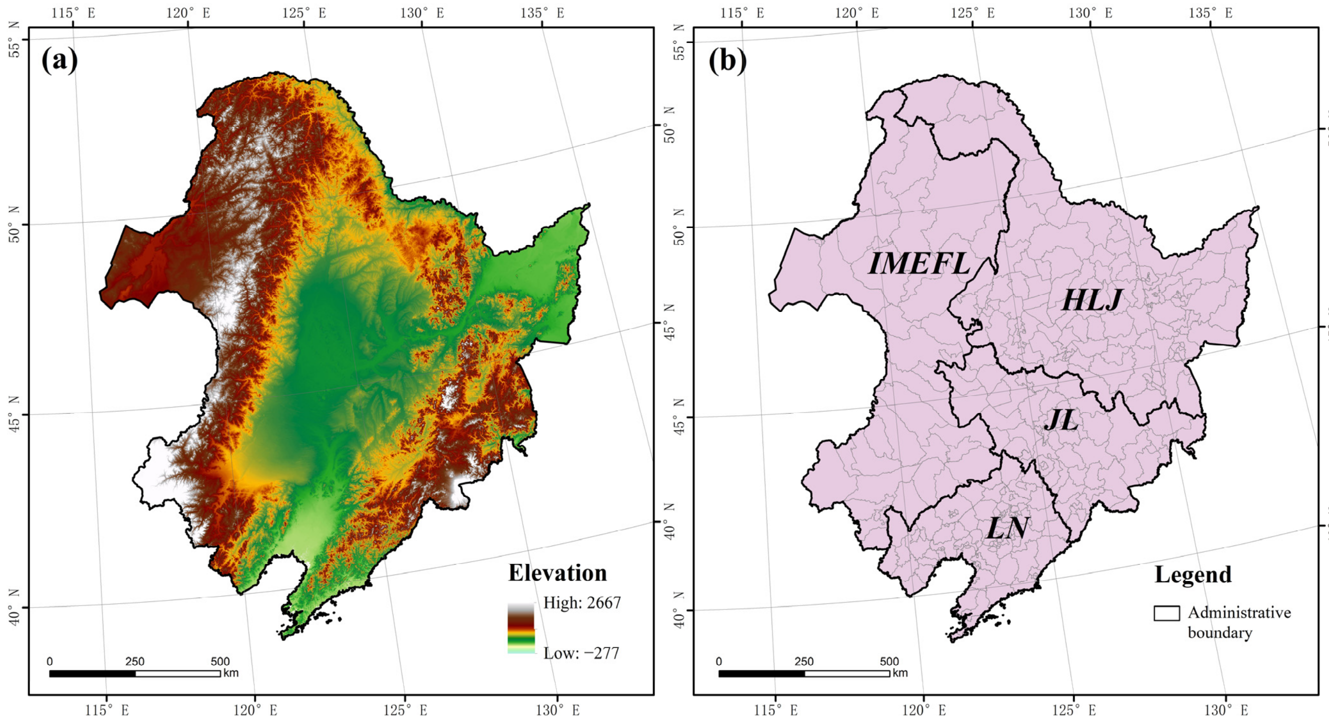

Figure 1. (**a**) Digital elevation model (DEM) of northeast China's black soil region. (**b**) Northeast China's black soil region's districts and counties. (HLJ: Heilongjiang Province; JL: Jilin Province; LN: Liaoning Province; IMEFL: Inner Mongolia East Four Leagues).

The sowing area for crops in the northeast black soil region has continued to expand, resulting in a "grain-dominated" agricultural planting pattern. By 2020, 28.655 million hectares of grain were sown in the northeast, with grain crops accounting for 93.3% of the sown area of crops. Rice, corn, and soybeans constituted 98.9% of the sown area of grain crops. Data from the Third National Land Survey for 2019–2021 reveal that the cultivated land area in the northeast decreased from 37.504 million hectares in 2019 to 37.4165 million hectares at the end of 2021, indicating a reduction of 83.9 million hectares or 0.22% (Table 1) [33]. Consequently, amid decreasing arable land area, spatial zoning in the northeast black soil region can facilitate a rational implementation of policies and measures tailored to different regions, thus addressing the constraints to sustainable agricultural development and ensuring the black soil's wide-scale, multifaceted, and in-depth sustainable development.

Table 1. Change in arable land area (10^2 km^2) in the northeast during the period of 2019–2021.

Province	2019	2021	Change
Liaoning Province	518.21	515.36	−2.85
Jilin Province	749.85	744.98	−4.87
Heilongjiang Province	1719.54	1716.58	−2.96
Inner Mongolia East Four Leagues	762.44	764.73	2.29
Total	3750.04	3741.65	−8.39

2.2. Data Sources

This study involved the collection of 10 factors from satellite remote sensing and spatial datasets to delineate cropland zones in the northeast black soil region while covering certain aspects, such as climate, topography, soil, and vegetation (Table 2). Climatic variables, comprising mean annual precipitation (Pre_mean) and mean annual temperature (Tem_mean), were derived from the aggregation of nearly two decades of annual average data spanning from 2002 to 2022. Topographic variables, such as elevation (Ele) and slope (Slo), were derived using the "surface" function in ArcGIS 10.6 based on a 30 m resolution digital elevation model (DEM). Soil variables encompassed organic matter (SOM_mean), clay (Clay_mean), silt (Silt_mean), and sand (Sand_mean). Vegetation variables comprised the normalized vegetation index (NDVI_mean) and the enhanced vegetation index (EVI_mean). The land use type dataset for the northeast region was sourced from the Resource and Environment Science Data Center of the Chinese Academy of Sciences (https://www.resdc.cn/DOI/DOI.aspx?DOIID=54, accessed on 29 December 2023), featuring a resolution of 30 m × 30 m; vector data for administrative divisions within the northeast were sourced from the National Center for Basic Geographic Information (https://www.webmap.cn/, accessed on 27 December 2023). For streamlined spatial data analysis and processing, all data were converted to a consistent coordinate and projection system (WGS, 1984; UTM Zone 52 N).

Table 2. Covariates for cropland management zoning in the northeastern black soil region.

Factor	Abbreviation	Resolution	Name	Data Source
Climate	Pre_mean	1 km	Precipitation	National Earth System Science Data Center, National Science and Technology Infrastructure of China (http://www.geodata.cn, accessed on 27 December 2023)
	Tem_mean	1 km	Temperature	National Earth System Science Data Center, National Science and Technology Infrastructure of China (http://www.geodata.cn, accessed on 27 December 2023)
Soil	SOM_mean	30 m	Soil Organic Matter	[34] Xiangtian Meng, Yilin Bao, Chong Luo, Xinle Zhang, Huanjun Liu, SOC content of global Mollisols at a 30 m spatial resolution from 1984 to 2021 generated by the novel ML–CNN prediction model, *Remote Sensing of Environment*, Volume 300, 2024, 113911, ISSN 0034-4257, https://doi.org/10.1016/j.rse.2023.113911
	Clay_mean	1 km	Mean clay	The dataset is provided by the Data Center for Resources and Environmental Sciences, Chinese Academy of Sciences (RESDC) (http://www.resdc.cn, accessed on 27 December 2023)
	Silt_mean	1 km	Mean silt	The dataset is provided by the Data Center for Resources and Environmental Sciences, Chinese Academy of Sciences (RESDC) (http://www.resdc.cn, accessed on 29 December 2023)
	Sand_mean	1 km	Mean sand	The dataset is provided by the Data Center for Resources and Environmental Sciences, Chinese Academy of Sciences (RESDC) (http://www.resdc.cn, accessed on 29 December 2023)
Vegetation	NDVI_mean	250 m	Mean normalized difference vegetation index	[35] Xiong, C., Ma, H., Liang, S. et al. Improved global 250 m 8-day NDVI and EVI products from 2000–2021 using the LSTM model. *Sci Data* 10, 800 (2023). https://doi.org/10.1038/s41597-023-02695-x

Table 2. *Cont.*

Factor	Abbreviation	Resolution	Name	Data Source
	EVI_mean	250 m	Mean enhanced vegetation index	[35] Xiong, C., Ma, H., Liang, S. et al. Improved global 250 m 8-day NDVI and EVI products from 2000–2021 using the LSTM model. *Sci Data* 10, 800 (2023). https://doi.org/10.1038/s41597-023-02695-x
Terrain	Ele	30 m	Elevation	Calculated from DEM
	Slo	30 m	Slope	Calculated from DEM

Ten zoning indicators were collected and categorized into four main types: climate, soil, vegetation, and topography. Specific averages from these 10 indicators were calculated for all 317 districts and counties within the study area (Figure 2).

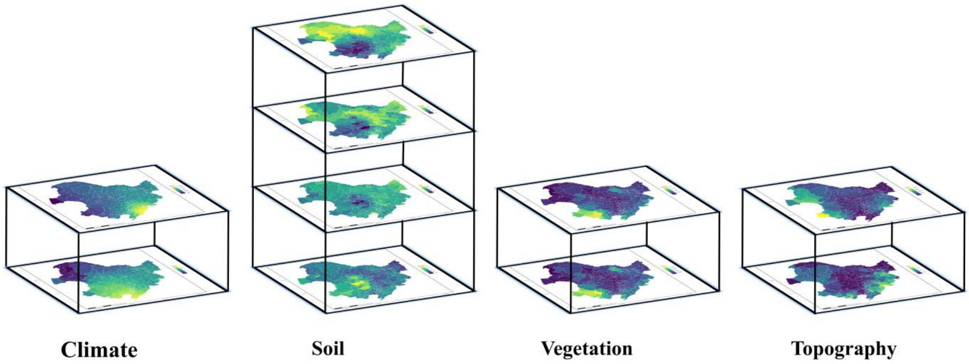

Figure 2. Evaluation of the indexes for cropland zoning in the northeastern black soil region. (Climate: includes temperature and precipitation; Soil: includes organic matter content, sand content, silt content, and clay content; Vegetation: includes NDVI and EVI; Topography: includes elevation and slope).

A. Climate. To characterize the region's climatic conditions, the average annual precipitation and temperature were calculated. Recognizing that data from a single year may not be sufficiently representative, the average values for these variables were derived from the last 20 years, from 2002 to 2022. This involved summing the data across these years and then calculating the mean, thus yielding the average annual precipitation and temperature for each district and county in the northeast black soil region.

B. Soil. Soil properties are principally characterized by organic matter content and soil texture. The organic matter content data for the northeast region were derived from the ML–CNN model, which inverted actual sampling data spanning from 1984 to 2021 [34]. This dataset spans a significant period and is based on highly accurate actual sampling data. Soil texture, a critical soil physicochemical property, refers to the combination of mineral particles of varying sizes and diameters within the soil. Closely related to soil aeration, fertilizer and water retention, and cultivation difficulty, soil texture serves as a vital foundation for formulating soil utilization, management, and improvement measures. Data were compiled from 1:1 million soil type maps and soil profile data obtained from the Second Soil Census, classifying soil texture based on the content of sand, silt, and clay. Data categories include sand, silt, and clay, each reflecting the percentage of content of differently textured particles.

C. Vegetation. For assessing vegetation, the greenness indexes from remote sensing, specifically the normalized difference vegetation index (NDVI) and enhanced vegetation index (EVI), were selected. NDVI and EVI, as vegetation indices, offer a comprehensive and sensitive measure of vegetation growth status. Monitoring and analyzing these indices allows for a timely understanding of vegetation health, growth trends, and coverage, providing a scientific basis for cultivated land zoning in the northeast black soil region.

The newly introduced global 250 m resolution, eight-day interval NDVI and EVI products offer unique advantages for satellite vegetation monitoring. Utilizing the advanced Long Short-Term Memory (LSTM) neural network method effectively minimizes the impact of clouds and other contaminants, yielding more accurate vegetation information. Its high spatial resolution of 250 m renders it ideal for detailed observation and analysis of surface vegetation conditions. When combined with the Savitzky–Golay filter (SG), Global Land and Surface Satellite (GLASS) Leaf Area Index (LAI) fitting, upper envelope methods for model training, and intercomparisons with the MODIS VI product, among other techniques, this product demonstrates low root mean square error, affirming its reliability and superiority in vegetation monitoring [35].

D. Topography. Elevation and slope were chosen as key indicators for cropland management zoning in the northeast black soil region in aiming to delineate management areas more comprehensively and scientifically, thereby enhancing agricultural production efficiency and preserving the ecological environment. Firstly, these topographic factors significantly influence hydrological processes and soil moisture distribution, which are crucial for the rational utilization of water resources and effective land drainage. Secondly, as they are fundamental to land use planning, elevation and slope guide agricultural activities and other land uses across various areas, facilitating sustainable land resource utilization. Moreover, acknowledging the climate and soil condition heterogeneity in areas of varied elevations and slopes, crop adaptability and yield can be enhanced through targeted management strategies. In ecological environmental protection, rational management area division can mitigate soil erosion and optimize local ecosystem protection. The selection of these indicators lays a scientific foundation for cropland management in the northeast black soil region, promoting the synergistic advancement of sustainable agricultural development and ecological preservation.

2.3. Methods

2.3.1. Heterogeneity between Subregions

The Moran index was employed to determine spatial correlation by assessing heterogeneity between partitions [36]. It is highly effective for measuring spatial autocorrelation by quantifying the degree of similarity between an object in geographic space and its neighboring objects [37]. This is particularly pertinent to our research, which concentrates on zoning northeastern black soil cropland based on multiple metrics. The Moran index enables the detection and quantification of spatial clusters or outliers, which is essential for identifying regions of homogeneity or heterogeneity. Compared to other models, like Geary's C or Getis–Ord General G, the Moran index is more sensitive to spatial autocorrelation, providing clearer and more explicit results in assessing the level of heterogeneity between subzones based on the correlation magnitude [38,39]. Furthermore, Moran's index can be applied to both global and local spatial autocorrelation analyses, thus offering a comprehensive view of spatial patterns, which contributes to its widespread use in spatial analysis [40].

- Global spatial autocorrelation.

Global spatial autocorrelation (Global Moran's I) is utilized to identify spatial patterns throughout the entire study area, thus determining whether phenomena are spatially clustered [41].

$$\text{Global Moran's I} = \frac{1}{\sum_{i=1}^{n}\sum_{j=1}^{n} Wij} \cdot \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} Wij\left(Hi-\overline{H}\right)\left(Hj-\overline{H}\right)}{\sum_{i=1}^{n}\left(Hi-\overline{H}\right)^{2}} \quad (1)$$

In the formula, Hi and Hj represent the corresponding indicator values for different districts, i and j, respectively; $\overline{H}$ denotes the average value of the indicator; n is the number of districts in the northeastern black soil region; and Wij represents the spatial weight matrix. The Global Moran's I index range of the index is $[-1, 1]$, and >0 indicates that in the spatial positive correlation, the observed attributes are in a spatially clustered spatial pattern; the closer it is to 1, the stronger the positive correlation. In contrast, <0 indicates that in the spatial negative correlation, the observed attributes are in a spatially discrete spatial pattern, and the closer it is to -1, the stronger the negative correlation; when it is close to 0, the observed attributes do not have spatial autocorrelation, and they are spatially randomly distributed.

- Local spatial autocorrelation.

Local spatial autocorrelation, commonly known as clustering and outlier analysis, utilizes clustering and outlier analysis (Anselin Local Moran's I) to facilitate the identification of outliers characterized by high values surrounded predominantly by low values (HL) and, vice versa, where low values are predominantly surrounded by high values (LH) [34].

$$\text{Anselin Local Moran's I} = \frac{Hi - \overline{H}}{Si^2} \sum\nolimits_{j=1, j \neq i}^{n} Wi, j(Hj - \overline{H}) \tag{2}$$

$$Si^2 = \frac{\sum_{j=1, j \neq i}^{n} \left(Hj - \overline{H}\right)^2}{n-1} \tag{3}$$

In the formula, the total number of counties in the northeast black earth region is n; Hi is the value of the attribute of counties in the northeast black earth region; $\overline{H}$ is the mean value; and Wi, j denotes the spatial weight between districts i and j.

2.3.2. Homogeneity within Partitions

According to the standard deviation of the image element value inside the partition, the area size is considered to give the corresponding weight to judge the homogeneity inside the partition, and the formula is as follows:

$$S = \frac{\sum\limits_{i=1}^{n} a_i s_i}{\sum\limits_{i=1}^{n} a_i} \tag{4}$$

In the formula, s_i is the standard deviation of the range image value of partition i, a_i is the area of partition i, and n is the total number of partitioned partitions in the study area after partitioning. A smaller S value indicates lower spatial heterogeneity within the partition, signifying a high degree of homogeneity.

2.3.3. Assessment of Optimal Number of Groups (Calculation of Pseudo F-Statistics)

Cluster validity is measured using the Calinski–Harabasz pseudo F-statistic, which is a ratio reflecting intra-group similarity and inter-group dissimilarity [42,43]. Larger values of the pseudo F-statistic represent stronger intra-class group ties and more dispersed inter-group distances, i.e., better clustering results. The pseudo F-statistic is calculated as:

$$F = \frac{\left(\frac{R^2}{n_c - 1}\right)}{\left(\frac{1 - R^2}{n - n_c}\right)} \tag{5}$$

$$R^2 = \frac{SST - SSE}{SST} \tag{6}$$

$$SST = \sum_{i=1}^{n_c} \sum_{j=1}^{n_i} \sum_{k=1}^{n_v} \left(V_{ij}^k - \overline{V^k}\right)^2 \tag{7}$$

$$SSE = \sum_{i=1}^{n_c} \sum_{j=1}^{n_i} \sum_{k=1}^{n_v} \left(V_{ij}^k - \overline{V_t^k} \right)^2 \quad (8)$$

where R^2 is a measure of how well the regression model fits the data and indicates the proportion of variance explained, SST is the degree of dissimilarity between groups, as portrayed by the sum of squares between groups, and SSE is the degree of dissimilarity within groups, as portrayed by the sum of squares within groups. n is the number of elements, n_i is the number of elements in group i, n_c is the number of classes (groups), n_v is the number of variables used to make groupings of the elements, V_{ij}^k is the value of the variable for element k of element j in group i, $\overline{V^k}$ is the average of the variables for variable k, and $\overline{V_t^k}$ is the average of the values of variables for k in group i.

2.4. Research Framework

The research framework of this study is shown in Figure 3.

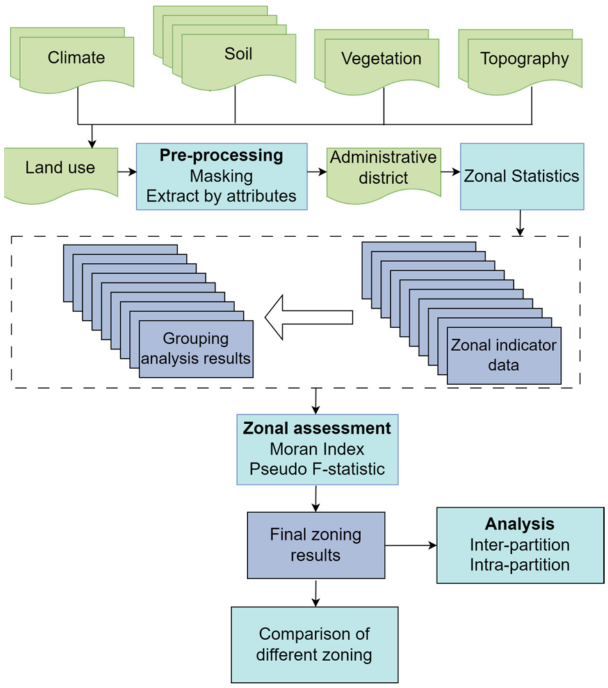

Figure 3. The research framework for cropland zoning in the northeast black soil region.

Initially, we employed 10 categories of indicators across four dimensions—climate, soil, vegetation, and topography—integrated with the cropland extent layer and the district and county administrative division extent layer for masking and statistical partitioning. Subsequently, these processed indicator layers across 10 categories served as inputs to delineate various partitions. Ultimately, we evaluated the zoning outcomes using the pseudo F-statistic and area-weighted standard deviation methods to determine the most suitable zoning configurations.

3. Results

3.1. Descriptive Statistics

Descriptive statistics for 10 elements across four major hierarchical indicators were analyzed in 317 districts and counties within the northeast black soil region (Table 3). These elements include precipitation (Pre), temperature (Tem), soil organic matter content (SOM), normalized vegetation index (NDVI), enhanced vegetation index (EVI), elevation (Ele), slope (Slo), and the contents of clay (Clay), sand (Sand), and silt (Silt).

Table 3. The descriptive statistical characteristics.

Norm	N	Mean	SD	Median	Min	Max	Skewness	Kurtosis
Pre	317	704.12	193.18	676.02	203.33	1262.58	0.61	0.39
Tem	317	5.85	2.88	5.62	–3.02	11.74	–0.30	–0.15
SOM	317	38.22	8.20	39.02	15.95	57.25	–0.29	–0.22
NDVI	317	0.05	0.01	0.05	0.03	0.11	1.83	3.13
EVI	317	0.05	0.01	0.04	0.03	0.08	1.52	2.14
Ele	317	239.53	196.01	185.80	1.86	1229.94	1.71	3.52
Slo	317	3.44	1.83	2.88	1.32	9.54	0.94	0.06
Clay	317	24.82	3.31	24.78	13.60	32.41	–0.25	–0.31
Sand	317	45.67	6.60	46.04	27.72	71.06	0.49	1.43
Silt	317	29.51	4.12	29.51	11.61	45.16	–0.59	3.83

The average precipitation was determined to be 704.12 mm, exhibiting a slight rightward skew, with a maximum value reaching 1262.58 mm. The average temperature was recorded at 5.85 °C, demonstrating a slight leftward skew. The mean soil organic matter content was found to be 38.22 g/kg, indicating a slight leftward skew. For the normalized vegetation index (NDVI) and enhanced vegetation index (EVI), the mean values were 0.05 and 0.04, respectively, both displaying a positive skew, suggesting a relatively concentrated vegetation distribution. The mean elevation was 239.53 m, characterized by a degree of positive skew and a long-tailed distribution. The average slope was found to be 3.44 degrees, exhibiting a slight positive skew. The average contents of clay, sand, and silt were 24.82%, 45.67%, and 29.51%, respectively, with clay content showing a slight leftward skew and sand content a rightward skew.

3.2. Correlation Analysis

Figure 3 presents a scatterplot matrix depicting pairwise relationships among various variables, typically environmental and soil-related parameters. Additionally, LOESS (locally weighted scatterplot smoothing) curves are displayed in each plot, serving as a nonparametric method to accentuate data trends. Furthermore, the matrix features histograms along the diagonal to illustrate the distribution of individual variables and includes correlation coefficients for each variable pair, with an asterisk denoting the significance level of the correlation.

Figure 4 displays pairwise relationships among diverse environmental and soil parameters. Each subfigure features a scatterplot with an overlaid LOESS curve (locally weighted scatterplot smoothing) to underscore potential nonlinear relationships in the data. Histograms along the diagonal display the distribution of each variable, with significant skewness and kurtosis in parameters like NDVI and EVI indicating deviations from normal distribution. Correlation coefficients (Corr) for each variable pair are presented below each plot, with asterisks indicating correlation significance levels ($p < 0.05$, $p < 0.01$, $p < 0.001$). Initial analysis uncovered several significant correlations. For instance, temperature (Tem) exhibited a moderately negative correlation (Corr: −0.65) with soil organic matter (SOM), suggesting that soil organic matter may decrease as temperature increases, likely due to heightened microbial decomposition rates. A strong positive correlation (Corr: 0.97) was observed between NDVI and EVI, as both are vegetation indices. Elevation (Ele) demonstrated a slight negative correlation (Corr: −0.29 **) with temperature, aligning with the adiabatic cooling effect at higher elevations. Notably, significant correlations were observed among soil textural components, including Clay_mean, Silt_mean, and Sand_mean. The percentage of Sand_mean was negatively correlated with Pulverized_mean (Corr: −0.91 ***), consistent with soil texture classification.

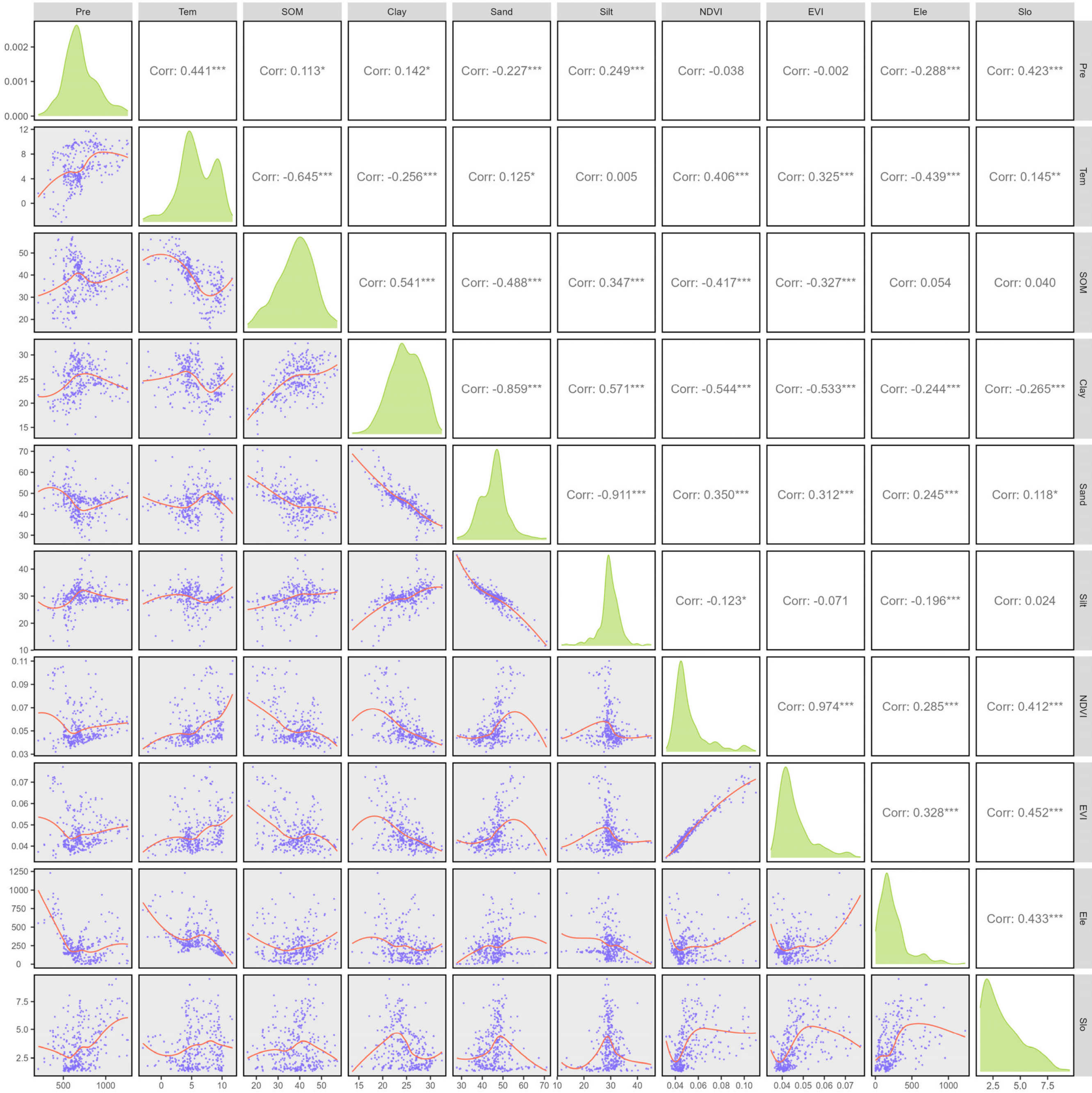

Figure 4. Correlation of indicators and locally weighted linear regression linear results (precipitation (Pre), temperature (Tem), soil organic matter (SOM), clay (Clay), sand (Sand), silt (Silt), normalized difference vegetation index (NDVI), enhanced vegetation index (EVI), elevation (Ele), and slope (Slo); (* for $p < 0.05$, ** for $p < 0.01$, *** for $p < 0.001$)).

3.3. Analysis of Spatial Heterogeneity at the Unidimensional Indicator Level

The northeast black soil region exhibits significant spatial heterogeneity across various single-dimensional indicators (Figure 5). Regarding average rainfall, there is a trend of higher rainfall in the southeast and lower rainfall in the northwest, with the southeast experiencing more abundant rainfall, positively impacting agricultural production. The pattern of higher temperatures in the south and lower temperatures in the north suggests that the southern region is more conducive to a variety of crops. However, the trend of lower organic matter content in the south and higher content in the north may indicate greater soil fertility in the northern region. Variations in soil particle size revealed that clay particles predominantly occurred in the central, northwestern, and northeastern parts of the northeast plain, sand particles were primarily in the southwest, and silt particles exhibited a distribution of low in the center and high in the periphery, providing a crucial foundation for soil zoning in the northeast black soil region. Temperature predominantly exhibited a trend of being higher in the south and lower in the north. Regarding vegetation indices, NDVI and EVI were higher in the southwest and lower in the northwest, northeast, and center, reflecting the region's varied vegetation conditions. This correlation with factors like soil properties, temperature, and rainfall provides comprehensive vegetation information for cropland management. Additionally, elevation and slope were lower in the central, southern, and northeastern areas, whereas the Daxinganling, Xiaoxinganling, and Changbai Mountains exhibited higher values. This information is valuable for designing agricultural land and developing agricultural management strategies.

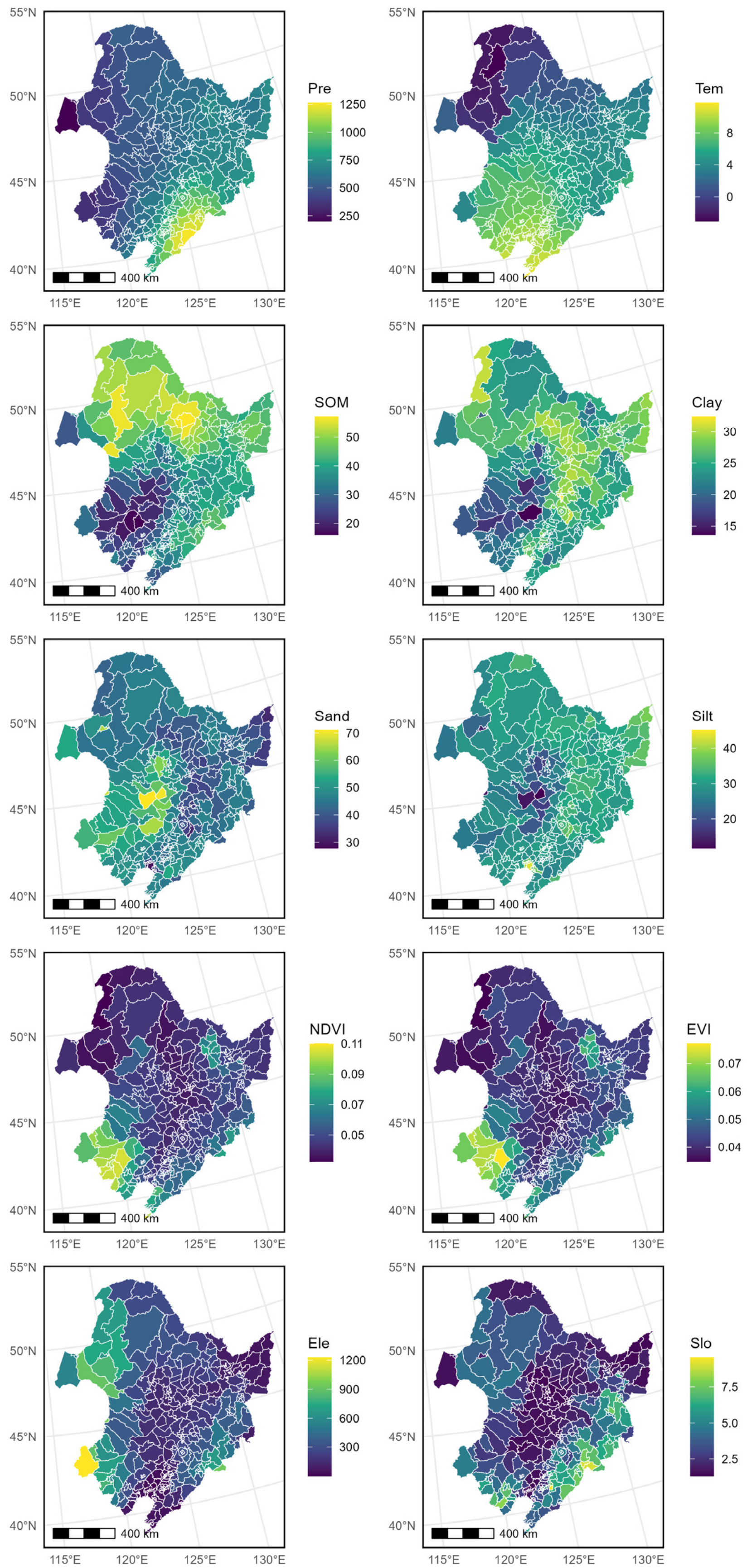

Figure 5. Indicator data used for cropland zoning in the northeastern black soil region.

3.4. Zoning of Cropland Management Types in the Northeastern Black Soil Zone

The spatial distribution of arable land in northeast China's black soil region is characterized by spatial clustering. Using indicators of 10 elements across four levels—climate, soil, vegetation, and topography—the black soil region of northeast China was divided into six zones by applying partitioning principles and methods. This division allows for a systematic depiction of the spatial pattern of arable land in the region, providing a scientific basis for the comprehensive protection and utilization of the black soil area in a regional context.

The grouping analysis tool within the ArcGIS 10.6 Cluster Distribution Mapping Toolset enables the grouping of element attributes under spatial constraints. This analysis employs a connectivity map (minimum spanning tree) to identify natural groupings of administrative counties in the northeast black earth region, thus ensuring their spatial connectivity as much as possible. The spatial constraint method used in this analysis is the Delaunay triangulation (DT) method, whereby elements in the same group share at least one natural neighborhood with another element in the group [44,45]. The optimal number of groups for subgroup analysis is determined by the pseudo F-statistic, which measures subgroup effectiveness. The range of pseudo F-statistic values for groups between 2 and 10 is [58.7704, 64.1336], with the highest value corresponding to six groups (Figure 6). Consequently, the cultivated land in the northeast black soil region was divided into six types of zones. In Figure 7, the weighted standard deviation obtained from the calculation when the number of groups is six is overall smaller, which also indicates that the partitioning effect is optimal when the grouping is six. Figure 8 displays the results of the grouping analysis. Based on the characteristics of the types of zones in the northeast black soil region (Table 4), the groups were named the southwest, central–south, central–west, northwest, southeast, and northeast zones, respectively.

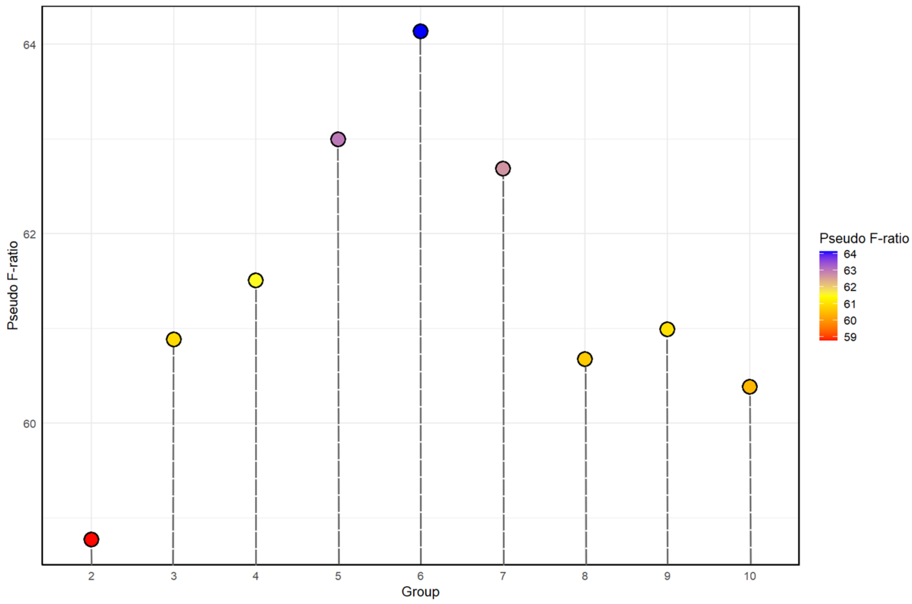

Figure 6. Changes in the trend of pseudo F-statistics with groups of 2–10.

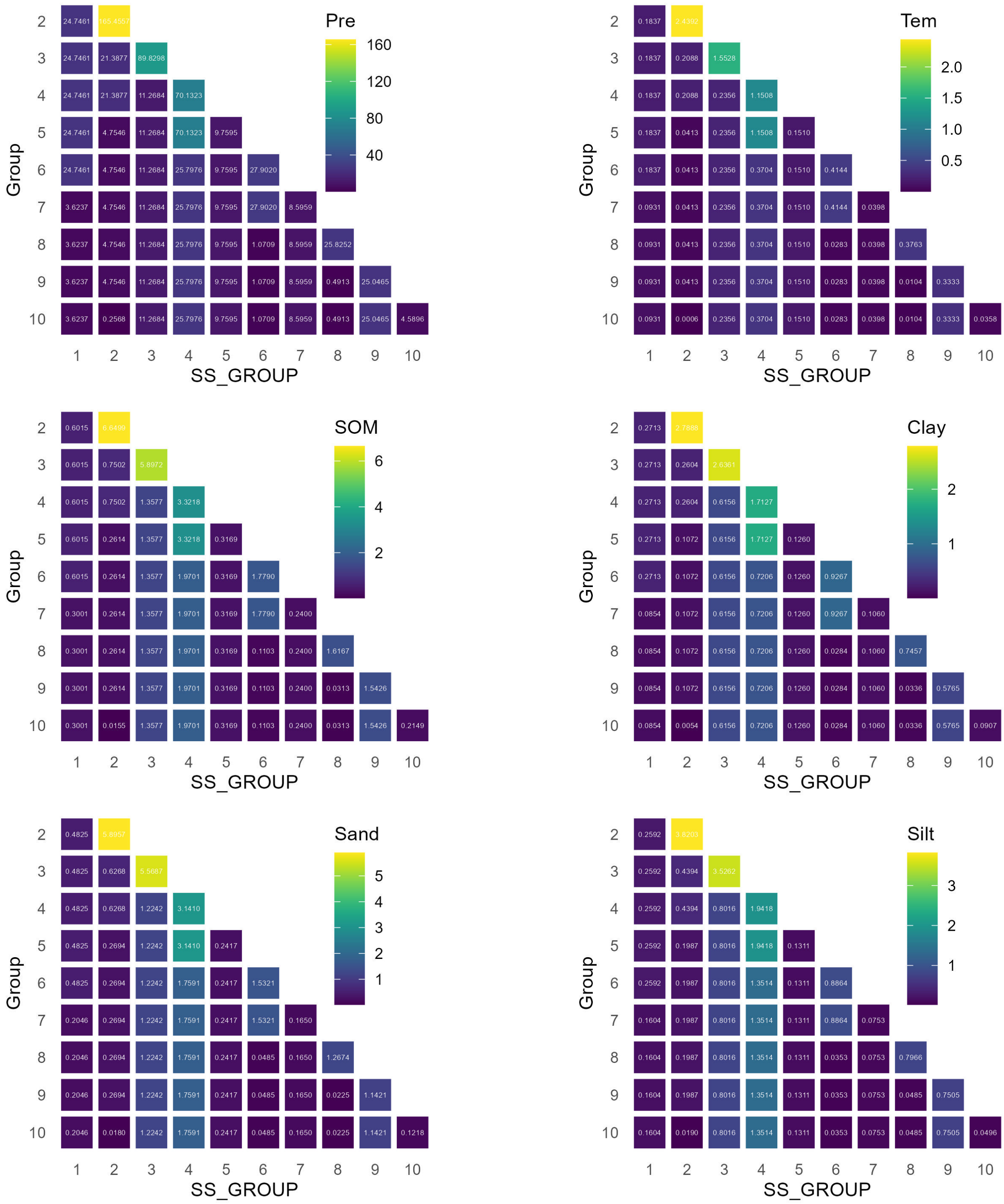

Figure 7. *Cont.*

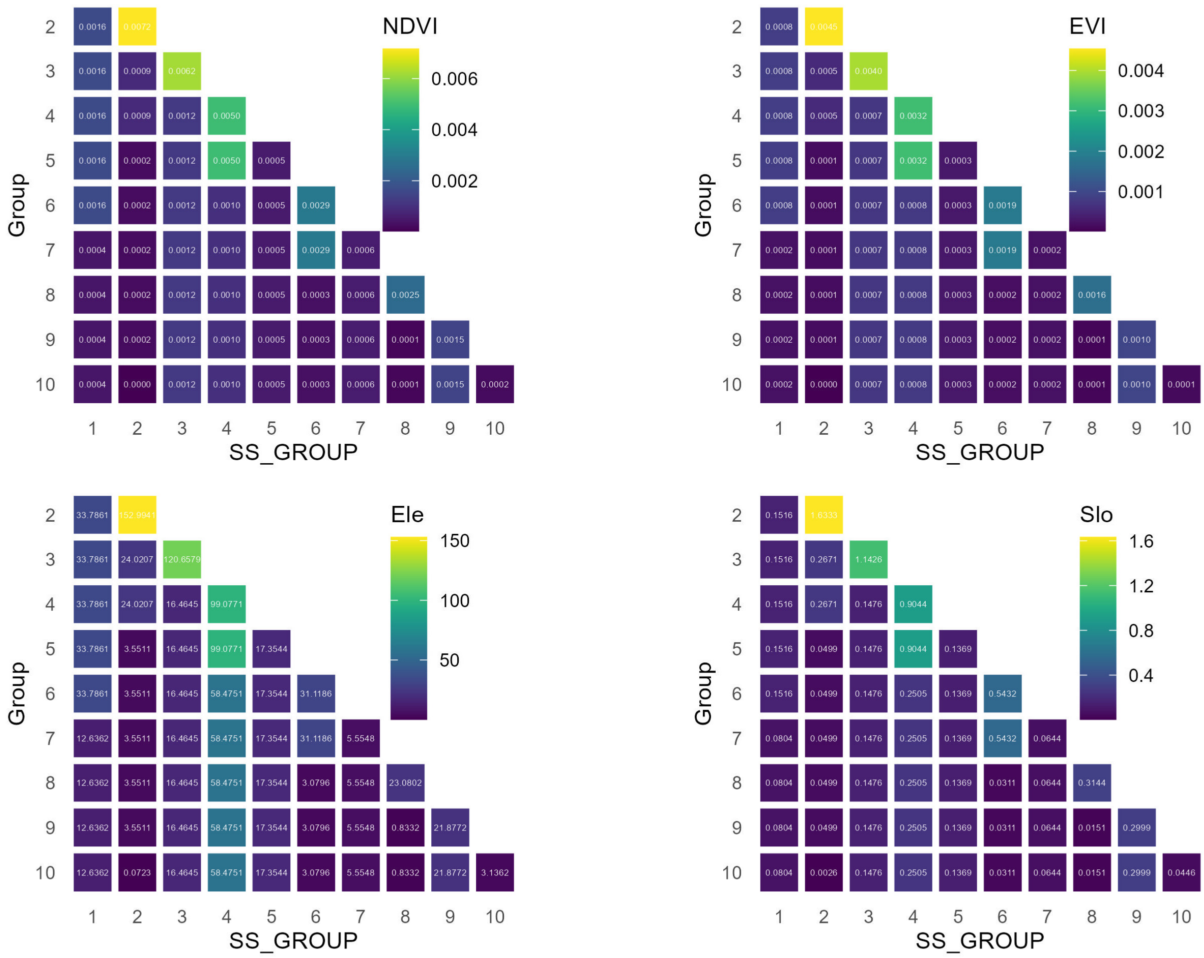

Figure 7. Trends in weighted standard deviation for grouping into 2–10 groups.

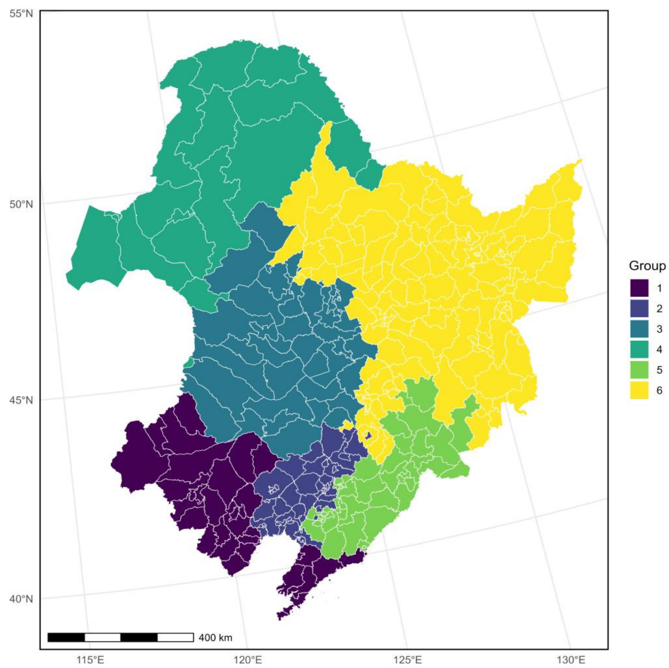

Figure 8. Zoning results of arable land in the northeast black soil region. (1. The southwestern region. 2. The central–south region. 3. The central–western region. 4. The northwest region. 5. The southeast region. 6. The northeast region).

Table 4. Statistical results of group analysis.

Category	Group	Number of Counties	Area (km^2)	Percentage (%)	Moran Index	Expectation	Variance	Moran Statistic	Moran_p_Value
Overall	-	317	1,241,301.42	100	0.7706	−0.0032	0.0013	21.5056	8.2×10^{-63}
Southwest	1	38	135,147.43	10.89	0.7142	−0.0286	0.0176	5.6052	1.6×10^{-5}
Central–South	2	47	56,032.92	4.51	0.6611	−0.0227	0.0105	6.6847	1.2×10^{-8}
Central–West	3	42	230,094.94	18.54	0.6141	−0.0244	0.0106	6.2140	9.8×10^{-6}
Northwest	4	18	305,053.38	24.58	0.3097	−0.0625	0.0298	2.1301	1.0×10^{-1}
Southeast	5	44	98,981.41	7.97	0.5148	−0.0233	0.0112	5.0715	1.7×10^{-3}
Northeast	6	128	415,991.34	33.51	0.6605	−0.0079	0.0034	11.3884	5.2×10^{-16}

The spdep package in R was utilized to analyze the cultivated land in the northeast black soil region, examining the spatial autocorrelation of both the entire study area and its subgroups based on 10 indicator attributes across climate, soil, vegetation, and topographical levels. Table 4 presents the number, area, and share of each subgroup area, along with key statistical indicators related to spatial autocorrelation, such as Moran's I, expected value, variance, Moran's statistic, and p-value. The overall Moran's I index of 0.7706 and the p-value of 8.2×10^{-63} for the northeastern black soil region indicate strong positive spatial autocorrelation, meaning similar observations are more likely to cluster spatially. The Moran's statistic of 21.5056 further supports the significance of the overall spatial autocorrelation. The Moran's index for each subgroup region ranged from 0.3079 to 0.7142, indicating that similar values are spatially adjacent to each other in patterns of clustering or dispersion, with most groups exhibiting strong positive spatial autocorrelation. In particular, the northwest district's Moran index of 0.3079 suggests a more dispersed spatial distribution and weaker correlation with neighboring values. This region encompasses 18 county-level administrative districts, the fewest among the sub-districts, yet spans an area of 305,053.38 km^2, accounting for 24.58% of the total, leading to a relatively dispersed spatial pattern among districts and counties.

The southwestern region is primarily located in the southern part of the Daxinganling Mountains, encompassing the Hunsandak and Horqin sands, Nuruer Chengshan in eastern Inner Mongolia, and a portion of the southern Liaodong Peninsula. It covers 38 administrative districts and counties, spanning an area of 135,147.43 km^2, which constitutes 10.89% of the total area of the northeast black soil region. Apart from a small portion of the southern Liaodong Peninsula, the region is predominantly semi-arid, experiencing severe cold in winter and warm, cool summers. Precipitation ranges from a maximum of 1250.45 mm to a minimum of 320.15 mm, with an average annual precipitation of 621.12 mm. This area is characterized by agriculture and animal husbandry, with dry land being predominant, and the main crops include corn, wheat, soybeans, and potatoes.

The central–south region, primarily located on the west side of the Liaohe Plain and east of the Horqin Sands and Nuruer Chengshan Mountains, encompasses 47 administrative districts and counties. This region spans an area of 56,032.92 km^2, the smallest portion of the northeast black soil region, representing 4.51% of the total area, yet it includes a significant number of administrative districts. The region enjoys a mild climate and abundant sunlight, with annual precipitation ranging from 594.02 to 933.07 mm. It boasts the highest average annual temperature among the six types of districts at 9.01 °C. Characterized by its flat terrain and extensive arable land, the region has abundant groundwater resources and favorable natural conditions. Rice, corn, and soybeans are the predominant crops.

The central–western region is situated in the central portion of the Daxinganling Mountains, the Songnen Plain, the Liaohe Plain, and the area north of the Horqin Sandy Land. It encompasses 42 administrative counties, spanning an area of 230,094.94 km^2, which accounts for 18.53% of the total area of northeast China. This region experiences a temperate, continental, semi-humid, and semi-arid monsoon climate, characterized by mild conditions and annual precipitation ranging between 429.50 mm and 727.44 mm. It is a typical black soil concentration area, with the northern Songnen Plain (above 45° N latitude) featuring a broad expanse of medium- and thick-layered black soil. Most of the region's cultivated land is flat, with low average elevation and fertile soil, providing an excellent foundation for agricultural production. Corn, rice, and soybeans are the primary crops, with the largest dryland area among the six types of zones.

The northwest region is situated in the northern section of the Daxingan Mountains, encompassing the Hulunbeier Plateau, the northwestern part of Heilongjiang Province, and the northeastern part of the Inner Mongolia Autonomous Region. This region includes 18 administrative districts, spanning 305,053.38 km^2, which represents 24.58% of the northeast black soil region. Characterized by mountainous terrain, sparsely populated areas, and high average elevation, it has the smallest proportion of reclaimed arable land among the six district types. With its continental monsoon climate, cool and humid conditions, and high latitude, it records the lowest average annual temperature among the six sub-districts, at −0.44 °C. The region receives an average precipitation of 424.21 mm, with minimal overall variation. In areas where dark brown soil prevails, arable land primarily consists of slopes ranging from 2° to 6°, and dry fields are common. The main crops include corn, soybeans, and oilseed rape.

The southeast region is situated in the southern Changbai Mountains, the eastern Liaohe Plain, and the northern Liaodong Peninsula. This region encompasses 44 administrative counties, covering an area of 98,981.41 km^2, which represents 7.97% of the northeast black earth region. Characterized by its mountainous and hilly terrain, this region hosts the headwaters of the Songhua, Yalu, and Tumen Rivers. The climate is mild and humid, with frequent cloud cover and limited sunshine in the mountainous areas. Annual precipitation ranges from 795.73 mm to 1262.58 mm, averaging 1027.44 mm, the highest among the six regions. With topography that is higher in the northeast and lower in the southwest, the region predominantly features sloping arable land, with corn and rice as the main crops.

The northeast region predominantly spans the Songhua River Basin, encompassing the Xiao Xing'anling and Changbai Mountains, as well as the Sanjiang Plain. This region includes 128 administrative districts and spans 415,991.34 km^2, making up 33.51% of the northeast black soil region. It features complex topographic variations and diverse landforms. Characterized by high elevations in the north and south and lower elevations in the center, it boasts the largest proportion and area of reclaimed arable land among the six zones. The region experiences a temperate humid continental monsoon climate with long, cold winters and short, warm summers. Annual precipitation varies from 516.33 mm to 1006.45 mm. Despite abundant precipitation, the region's average temperature is 4.18 °C, leading to low and cold temperatures. It has the widest area of rice cultivation among the six zones.

4. Discussion

4.1. Scientific and Comprehensive Indicators Are Key to Measuring Regional Zoning

Various scholars have approached regional zoning with distinct focuses, including zoning based on soil nutrient status, remote sensing monitoring of crop growth, nutritional diagnostic indicators, or multi-year yield data [46]. Zoning based on single-dimension indicators may yield results with insufficient accuracy. Therefore, we developed an indicator system encompassing four dimensions—climate, soil, vegetation, and topography—and integrated 10 indicators across these dimensions with equal weighting. This approach effectively circumvents the issue wherein single-dimension indicators fail to fully capture the comprehensive characteristics of each regional attribute for cropland zoning in the northeast black soil region. In our indicator system, we emphasized the climate, soil, vegetation, and topographical characteristics of the northeast black soil region, prioritizing the zoning of natural attributes more than previous studies.

4.2. Information on the Sanjiang Plain and Songnen Plain in the Same Subregion

Previous zoning typically divided the Sanjiang Plain and Songnen Plain into two separate zones, focusing heavily on topography and geomorphology [7,47]. However, our findings suggest a distinct categorization of the two zones, indicating that our analytical method emphasizes climate, soil, vegetation, and other factors over solely topographic and geomorphic considerations. The Sanjiang Plain and Songnen Plain share numerous similarities in topography, climate, soil, and vegetation. Both plains are situated at the same latitude. For individual indicators, the Sanjiang Plain and Songnen Plain show strong concordance in climate (precipitation and air temperature), soil (organic matter, silt, and clay), and vegetation (NDVI and EVI), as depicted in Figure 4. This implies that beyond differences in elevation and slope, other indicators significantly shape this area beyond the topographic factors of elevation and slope.

Furthermore, as shown in Figure 9, when the number of partitions is either 9 or 10, the Xiao Xing'anling mountain range, situated between the Sanjiang Plain and Songnen Plain, forms a distinct partition on its own. Meanwhile, the Sanjiang Plain and Songnen Plain remain contiguous, except for this small division, further validating the rationale for grouping the Sanjiang Plain and Songnen Plain within the same partition.

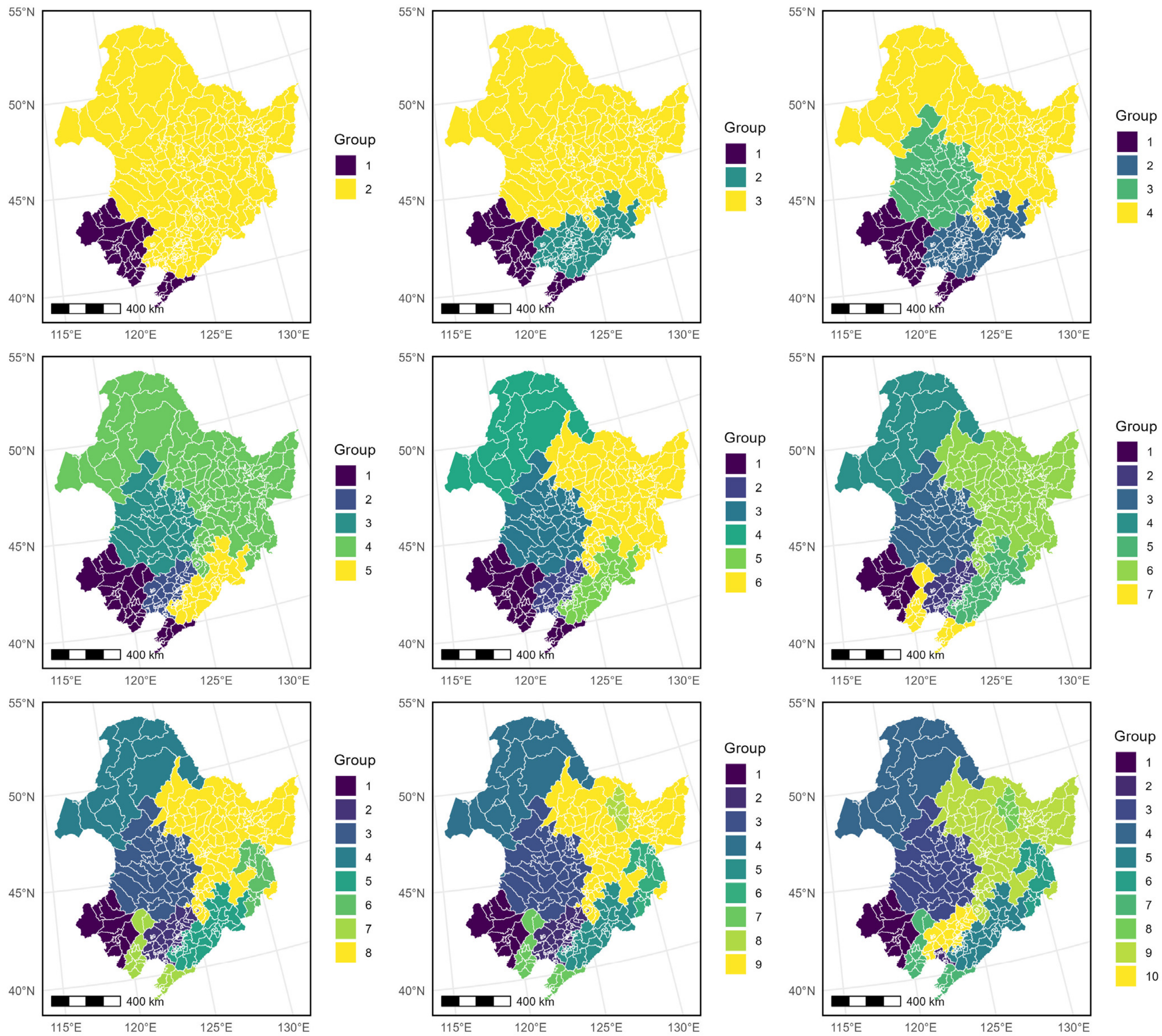

Figure 9. Zoning results for groupings from 2 to 10.

4.3. Evaluation of the Results of Zoning

Comprehensive evaluation is lacking in assessing the results of large-scale zoning. Currently, many zoning evaluations primarily rely on base yields, but yield-based assessments do not fully capture the actual effects of zoning. Therefore, the evaluation of zoning effectiveness should consider multiple factors, not solely yield [13]. The effectiveness of zoning is related not only to yield but also to a range of other factors, including soil quality, water use efficiency, fertilizer utilization, pest control, and agroecosystem sustainability. Therefore, relying solely on yield to evaluate the effectiveness of zoning may overlook other crucial factors. Additionally, the impact of certain variables on yield varies minimally, implying that the model will still perform well even if some variables are excluded from the analysis [48]. If using a large number of variables to analyze an agricultural region is not feasible, a smaller dataset may suffice to yield reliable results.

Some zonal evaluations involve calculating coefficients of variation based on the variability of soil nutrients within a zone relative to the variability between zones [49]. Others calculate the coefficient of variation or Moran's index based solely on yield or NDVI, focusing on a single dimension of soil nutrients or yield [50]. This type of evaluation typically assesses small scales, like field scale, and requires updated evaluation for larger scales, such as the northeast black soil region. Therefore, this paper calculated the Moran index for 10 indicators, employing the direct method of calculating the Moran index for each indicator separately and then averaging these indices.

In calculating the multivariate Moran's index, directly summing the individual Moran's indices, while straightforward, may fail to capture the interactions between different indicators. Ideally, multiple indicators should be synthesized into a composite index through the standardized calculation of their weighted average. The challenge lies in determining the weights of each indicator reasonably. Additionally, it is possible to directly calculate the multivariate Moran index, requiring the construction of a multivariate spatial autocorrelation model. The challenge here is the involvement of more complex statistical techniques and calculation processes. A prevalent approach is to downscale indicators and employ principal component analysis (PCA) to extract principal components that account for most of the data variability, followed by calculating the Moran index for these components [51,52]. However, the 10 indicators are ultimately consolidated into a single indicator for calculating the Moran index (also achievable by calculating the bivariate Moran index). If more than two indicators are extracted, it reverts to the previous challenge of computing the multivariate Moran index [53–55]. Considering the aforementioned concerns, along with the relatively independent nature of these indicators, this paper ultimately opts for the most straightforward approach to realize the multivariate Moran index calculation at this stage. Effective evaluation of large-scale, multi-indicator zoning outcomes is achievable through the application of the multivariate Moran index calculation method. In other analogous regions, this identical method may be employed for zoning assessments. Furthermore, the multivariate Moran index calculation method extends to the zoning assessment of diverse land types beyond merely arable land.

4.4. Shortcoming of the Study and Future Research Direction

In this paper, considering the spatial distribution characteristics of the northeast black soil region, we divided it into six regions, southwest, central–south, central–west, northwest, southeast, and northeast, using the grouping analysis tool while ensuring spatial connectivity as much as possible. This division can offer valuable references for the formulation of policies related to cultivated land protection and management in the northeast black soil region. However, similarly to other studies, this research has uncertainties and limitations that can be addressed in future work.

Firstly, as most of the input data for the study area are based on larger scales, with the majority of indicators having a resolution of 30 m × 30 m and a few at 1 km × 1 km, these need to be re-sampled and processed. The accuracy of the cropland zoning results in the northeast black soil zone is highly dependent on the resolution of the input data, which may introduce uncertainties and limitations to the zoning. Secondly, the current zoning indicators for the northeast black soil zone are based on 10 types of indicators across four levels: climate, soil, vegetation, and topography. These only consider natural attribute factors for zoning, while human attribute factors are not accounted for. The integration of qualitative and quantitative methods presents a challenge in the zoning of arable land in the northeast black soil region. This study primarily focuses on quantitative indicators, and there are deficiencies in the selection of qualitative indicators. Thirdly, the northeast black soil region encompasses 317 districts and counties. Validating zoning results at such a large scale cannot rely on a single dimension. In practice, due to the spatial differentiation within the northeast black soil region's cropland zoning, it is feasible to continue using the grouping analysis method for further zoning and to delineate regions at a smaller scale. This paper focuses solely on the spatial differentiation of cropland areas within the large-scale scope of the northeast black soil region and addresses cropland zoning, reflecting the relative consistency of cropland distribution. However, other land types besides cropland require further analysis.

5. Conclusions

This study employed spatial distribution data of climate, soil, vegetation, and topographical attributes across 317 districts and counties within the northeast black soil region's cultivated land areas to delineate six regions: southwest, central–south, central–west, northwest, southeast, and northeast. The findings revealed significant spatial variations in cropland zoning within the northeast black soil region based on indicators spanning the four dimensions of climate, soil, vegetation, and topography. The analysis of nine zoning types indicated that the optimal zoning effect was achieved when the northeast black soil region was divided into six zones, enhancing both inter-zonal variability and intra-zonal consistency. Furthermore, large-scale cropland zoning was evaluated using the pseudo F-statistic and area-weighted standard deviation methods, revealing similarities in the results. Regarding research methodology, this study achieves effective evaluation of large-scale, multi-indicator zoning outcomes through the application of the multivariate Moran index calculation method. This approach not only broadens the scope of arable land zoning evaluation across various regions but also facilitates zoning assessment for land types beyond arable land. Furthermore, this study thoroughly examines the consistency and variability patterns of arable land across various spatial scales in the northeast black soil area, offering a scientific foundation for arable land zoning protection in northeast China's black soil region. Subsequent efforts can further subdivide the area into six zoning districts, identify predominant factors, and enact differentiated management strategies to enhance arable land management effectiveness.

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

Funding: This work was supported by the Heilongjiang Province Ecological Environment Protection Scientific Research Project (HST2022ST005).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author. The data are not publicly available due to the project requirements.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Liu, B.; Zhang, G.; Xie, Y.; Shen, B.; Gu, Z.; Ding, Y. Delineating the black soil region and typical black soil region of northeastern China. *Chin. Sci. Bull.* **2021**, *66*, 96–106. [CrossRef]
2. Li, B.; Liu, Z.; Huang, F.; Yang, X.; Liu, Z.; Wan, W.; Wang, J.; Xu, Y.; Li, Z.; Ren, T. Ensuring national food security by strengthening high-productivity black soil granary in Northeast China. *Bull. Chin. Acad. Sci. (Chin. Version)* **2021**, *36*, 1184–1193.
3. Zhang, J.; Sun, B.; Zhu, J.; Wang, J.; Pan, X.; Gao, T. Black soil protection and utilization based on harmonization of mountain-river-forest-farmland-lake-grassland-sandy land ecosystems and strategic construction of ecological barrier. *Bull. Chin. Acad. Sci. (Chin. Version)* **2021**, *36*, 1155–1164.
4. Han, X.; Zou, W.; Yang, F. Main achievements, challenges, and recommendations of black soil conservation and utilization in China. *Bull. Chin. Acad. Sci. (Chin. Version)* **2021**, *36*, 1194–1202.
5. Han, X.; Zou, W. Research perspectives and footprint of utilization and protection of black soil in northeast China. *Acta Pedofil. Sin.* **2021**, *58*, 1341–1358.
6. Wang, J.; Xu, X.; Pei, J.; Li, S. Current situations of black soil quality and facing opportunities and challenges in northeast China. *Chin. J. Soil Sci.* **2021**, *52*, 695–701.
7. Xu, Y.; Pei, J.; Li, S.; Zou, H.; Wang, J.; Zhang, J. Main characteristics and utilization countermeasures for black soils in different regions of Northeast China. *Chin. J. Soil Sci.* **2023**, *54*, 495–504.
8. Schneider, A.; Friedl, M.A.; Potere, D. Mapping global urban areas using MODIS 500-m data: New methods and datasets based on 'urban ecoregions'. *Remote Sens. Environ.* **2010**, *114*, 1733–1746. [CrossRef]
9. Demuzere, M.; Kittner, J.; Martilli, A.; Mills, G.; Moede, C.; Stewart, I.D.; Van Vliet, J.; Bechtel, B. A global map of Local Climate Zones to support earth system modelling and urban scale environmental science. *Earth Syst. Sci. Data Discuss.* **2022**, *14*, 3835–3873. [CrossRef]
10. Bailey, R.G.; Bailey, R.G. Ecoregions and climate change. In *Ecoregions: The Ecosystem Geography of the Oceans and Continents*; Springer: New York, NY, USA, 2014; pp. 95–103.
11. Dadrasi, A.; Chaichi, M.; Nehbandani, A.; Soltani, E.; Nemati, A.; Salmani, F.; Heydari, M.; Yousefi, A.R. Global insight into understanding wheat yield and production through Agro-Ecological Zoning. *Sci. Rep.* **2023**, *13*, 15898. [CrossRef]
12. Walton, C.R.; Zak, D.; Audet, J.; Petersen, R.J.; Lange, J.; Oehmke, C.; Wichtmann, W.; Kreyling, J.; Grygoruk, M.; Jabłońska, E. Wetland buffer zones for nitrogen and phosphorus retention: Impacts of soil type, hydrology and vegetation. *Sci. Total Environ.* **2020**, *727*, 138709. [CrossRef] [PubMed]
13. Zymaroieva, A.; Zhukov, O.; Fedonyuk, T.; Pinkin, A. Application of Geographically Weighted Principal Components Analysis Based on Soybean Yield Spatial Variation for Agro-Ecological Zoning of the Territory. *Agron. Res.* **2019**, *17*, 2460–2473. [CrossRef]
14. Mirzaee, S.; Ghorbani-Dashtaki, S.; Mohammadi, J.; Asadi, H.; Asadzadeh, F. Spatial variability of soil organic matter using remote sensing data. *Catena* **2016**, *145*, 118–127. [CrossRef]
15. Eskandari Damaneh, H.; Jafari, M.; Eskandari Damaneh, H.; Behnia, M.; Khoorani, A.; Tiefenbacher, J.P. Testing possible scenario-based responses of vegetation under expected climatic changes in Khuzestan Province. *Air Soil Water Res.* **2021**, *14*, 117862. [CrossRef]
16. Brown, M.; McCarty, J. Is remote sensing useful for finding and monitoring urban farms? *Appl. Geogr.* **2017**, *80*, 23–33. [CrossRef]
17. Djerida, A. Investigating Groundwater Zones Relationship with Rainfall, Temperature and Modis-Derived NDVI and EVI. In Proceedings of the IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium, Kuala Lumpur, Malaysia, 17–22 July 2022; pp. 6440–6443.
18. Nse, O.U.; Okolie, C.J.; Nse, V.O. Dynamics of land cover, land surface temperature and NDVI in Uyo City, Nigeria. *Sci. Afr.* **2020**, *10*, e00599. [CrossRef]
19. Istanbuly, M.H.; Thabeet, A. Studying Changes in Normalized Difference Vegetation Index (NDVI) for Cover in the Area of Aleppo Mountain (Afrin) in Period between (2010–2016). *J. Univ. Babylon Pure Appl. Sci.* **2020**, *28*, 33–47.
20. Wang, M.; Liu, X.; Liu, Z.; Wang, F.; Li, X.; Hou, G.; Zhao, S. Evaluation and driving force analysis of cultivated land quality in black soil region of Northeast China. *Chin. Geogr. Sci.* **2023**, *33*, 601–615. [CrossRef]
21. Zhao, G.; Deng, Z.; Liu, C. Assessment of the Coupling Degree between Agricultural Modernization and the Coordinated Development of Black Soil Protection and Utilization: A Case Study of Heilongjiang Province. *Land* **2024**, *13*, 288. [CrossRef]
22. Davatgar, N.; Neishabouri, M.; Sepaskhah, A. Delineation of site specific nutrient management zones for a paddy cultivated area based on soil fertility using fuzzy clustering. *Geoderma* **2012**, *173*, 111–118. [CrossRef]
23. Zeraatpisheh, M.; Bakhshandeh, E.; Emadi, M.; Li, T.; Xu, M. Integration of PCA and fuzzy clustering for delineation of soil management zones and cost-efficiency analysis in a citrus plantation. *Sustainability* **2020**, *12*, 5809. [CrossRef]
24. Li, X.; Meng, X.; Ji, X.; Zhou, J.; Pan, C.; Gao, N. Zoning technology for the management of ecological and clean small-watersheds via k-means clustering and entropy-weighted TOPSIS: A case study in Beijing. *J. Clean. Prod.* **2023**, *397*, 136449. [CrossRef]
25. Rahbar, A.; Vadiati, M.; Talkhabi, M.; Nadiri, A.A.; Nakhaei, M.; Rahimian, M. A hydrogeochemical analysis of groundwater using hierarchical clustering analysis and fuzzy C-mean clustering methods in Arak plain, Iran. *Environ. Earth Sci.* **2020**, *79*, 342. [CrossRef]
26. Bhagwan, V.P.V.; Theerthala, A.A.; Devi, U.D.M.U.; Neelima, T.; Chary, D.S. Delineation and Evaluation of Management Zones for Site Specific Nutrient Management in Maize Tracts of Northern Telangana using Geostatistical and Fuzzy C Mean Cluster approach. *Res. Sq.* 2023, *preprint*.

27. Zhang, F.; Wang, H.; Zhao, X.; Jiang, Q. Investigation on Zoning Management of Saline Soil in Cotton Fields in Alar Reclamation Area, Xinjiang. *Agriculture* **2023**, *14*, 3. [CrossRef]
28. Wang, M.; Wang, H.; Feng, Y.; He, Y.; Han, Z.; Zhang, B. Investigating Urban Underground Space Suitability Evaluation Using Fuzzy C-Mean Clustering Algorithm—A Case Study of Huancui District, Weihai City. *Appl. Sci.* **2022**, *12*, 12113. [CrossRef]
29. Ao, M.; Zhang, X.; Guan, Y. Research and practice of conservation tillage in black soil region of northeast China. *Bull. Chin. Acad. Sci. (Chin. Version)* **2021**, *36*, 1203–1215.
30. FAO. *World Reference Base for Soil Resources 2014, Update 2015*; FAO: Rome, Italy, 2015.
31. Taxonomy, S. *Keys to Soil Taxonomy*; United States Department of Agriculture Natural Resources Conservation Service: Washington, DC, USA, 2014.
32. FAO. *Global Status of Black Soils*; FAO: Rome, Italy, 2022.
33. Zhou, Y.; Li, X.; Liu, Y. Cultivated land protection and rational use in China. *Land Use Policy* **2021**, *106*, 105454. [CrossRef]
34. Meng, X.; Bao, Y.; Luo, C.; Zhang, X.; Liu, H. SOC content of global Mollisols at a 30 m spatial resolution from 1984 to 2021 generated by the novel ML-CNN prediction model. *Remote Sens. Environ.* **2024**, *300*, 113911. [CrossRef]
35. Xiong, C.; Ma, H.; Liang, S.; He, T.; Zhang, Y.; Zhang, G.; Xu, J. Improved global 250 m 8-day NDVI and EVI products from 2000–2021 using the LSTM model. *Sci. Data* **2023**, *10*, 800. [CrossRef]
36. Chen, Y. Spatial autocorrelation equation based on Moran's index. *Sci. Rep.* **2023**, *13*, 19296. [CrossRef] [PubMed]
37. Chen, Y. New approaches for calculating Moran's index of spatial autocorrelation. *PLoS ONE* **2013**, *8*, e68336. [CrossRef] [PubMed]
38. Griffith, D.A. Interpreting Moran Eigenvector Maps with the Getis-Ord Gi* Statistic. *Prof. Geogr.* **2021**, *73*, 447–463. [CrossRef]
39. Lin, J. Comparison of Moran's I and Geary's C in Multivariate Spatial Pattern Analysis. *Geogr. Anal.* **2023**, *55*, 685–702. [CrossRef]
40. Anselin, L. The Moran scatterplot as an ESDA tool to assess local instability in spatial association. In *Spatial Analytical Perspectives on GIS*; Routledge: Abingdon, UK, 2019; pp. 111–126.
41. Liu, W.; Ma, L.; Smanov, Z.; Samarkhanov, K.; Abuduwaili, J. Clarifying Soil Texture and Salinity Using Local Spatial Statistics (Getis-Ord Gi* and Moran's I) in Kazakh–Uzbekistan Border Area, Central Asia. *Agronomy* **2022**, *12*, 332. [CrossRef]
42. Sinaga, S.J.; Satyahadewi, N.; Perdana, H. Determining the Optimum Number of Clusters in Hierarchical Clustering Using Pseudo-F. *Euler J. Ilm. Mat. Sains Dan Teknol.* **2023**, *11*, 372–382. [CrossRef]
43. Hochin, T.; Hayashi, Y.; Nomiya, H.; Chowdhury, M. Quasi-optimality under pseudo F statistic in clustering data. *J. Contrib.* **2018**, *7*, 320–324. [CrossRef]
44. Miky, Y.; Kamel, A.; Alshouny, A. A combined contour lines iteration algorithm and Delaunay triangulation for terrain modeling enhancement. *Geo-Spat. Inf. Sci.* **2023**, *26*, 558–576. [CrossRef]
45. Gökgöz, T.; Hacar, M. Comparison of two methods for multiresolution terrain modelling in GIS. *Geocarto Int.* **2020**, *35*, 1360–1372. [CrossRef]
46. Liu, H.; Bao, Y.; Xu, M.; Zhang, X.; Meng, X.; Pan, Y.; Yang, H.; Xie, Y. Comparison of precision management zoning methods in black soil area based on SOM and NDVI. *Trans. CSAE* **2019**, *35*, 177–183.
47. Fan, J.; Zhou, K.; Sheng, K.; Guo, R.; Chen, D.; Wang, Y.; Liu, H.; Wang, Z.; Sun, Y.; Zhang, J. Territorial function differentiation and its comprehensive regionalization in China. *Sci. China Earth Sci.* **2023**, *66*, 247–270. [CrossRef]
48. Vianna, L.F.d.N.; Massignan, A.M.; Pandolfo, C.; Dortzbach, D. Evaluating environmental factors, geographic scale and methods for viticultural zoning in the high-altitude region of Santa Catarina, Brazil. *Remote Sens. Appl. Soc. Environ.* **2019**, *13*, 158–170. [CrossRef]
49. Taye, M.; Simane, B.; Zaitchik, B.F.; Selassie, Y.G.; Setegn, S. Rainfall variability across the agro-climatic zones of a tropical highland: The case of the jema watershed, northwestern ethiopia. *Environments* **2019**, *6*, 118. [CrossRef]
50. Ma, S.-N.; Yin, Y.; Yu, Z.-Y.; Meng, L.-H.; Liu, Q.; Zhang, X.-L. UAV remote sensing accurate management zoning with terrain factor. *J. Jilin Agric. Univ.* **2021**, *43*, 205–212.
51. Parhizkar, T.; Rafieipour, E.; Parhizkar, A. Evaluation and improvement of energy consumption prediction models using principal component analysis based feature reduction. *J. Clean. Prod.* **2021**, *279*, 123866. [CrossRef]
52. Chen, J.; Qu, M.; Zhang, J.; Xie, E.; Huang, B.; Zhao, Y. Soil fertility quality assessment based on geographically weighted principal component analysis (GWPCA) in large-scale areas. *Catena* **2021**, *201*, 105197. [CrossRef]
53. Zhang, L.; Hou, Q.; Duan, Y.; Liu, W. Spatial Correlation between Water Resources and Rural Settlements in the Yanhe Watershed Based on Bivariate Spatial Autocorrelation Methods. *Land* **2023**, *12*, 1719. [CrossRef]
54. Ficetola, G.F.; Lunghi, E.; Manenti, R. Microhabitat analyses support relationships between niche breadth and range size when spatial autocorrelation is strong. *Ecography* **2020**, *43*, 724–734. [CrossRef]
55. Darikwa, T.B.; Manda, S.O. Measuring Bivariate Spatial Clustering in Disease Risks. In *Modern Biostatistical Methods for Evidence-Based Global Health Research*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 235–260.