sustainability

MDPI

Article

The Layout of Maize Variety Test Sites Based on the Spatiotemporal Classification of the Planting Environment

Xuli Zan [1,2], Zuliang Zhao [1,2], Wei Liu [1,2], Xiaodong Zhang [1,2], Zhe Liu [1,2,*], Shaoming Li [1,2] and Dehai Zhu [1,2]

[1] College of Land Science and Technology, China Agricultural University, Beijing 100083, China
[2] Key Laboratory of Remote Sensing for Agri-Hazards, Ministry of Agriculture and Rural Affairs, Beijing 100083, China
* Correspondence: liuz@cau.edu.cn; Tel.: +86-1381-072-0768

Received: 12 June 2019; Accepted: 30 June 2019; Published: 9 July 2019

Abstract: An appropriate layout of crop multi-environment trial (MET) sites is imperative for evaluating new crop varieties' performance in terms of agronomic traits and stress tolerance, and this information is used to determine the utilization value and suitable promotion region of new varieties. Actually, traditional maize test sites have been selected according to the experience of breeding experts, which leads to the strong subjective and unscientific conclusions regarding sites, as well as test results that are not representative of the target population of environments (TPE). Therefore, in this study, we proposed a new method for MET sites layout. Meteorological data, maize growth period data, and county-level maize planting area data were collected for the spatiotemporal classification of a given maize planting region to analyze change rules in the environmental category of each minimum research unit within the study period. If the occurrence frequency of its final attribution category reaches a certain threshold (50%), this minimum research unit is classified as a typical environment region; otherwise, it is classified as an atypical environment region. Then, the number of test sites in each environmental category is allocated by spatial stratified sampling. At last, we establish the optimal test sites layout and a reliability measurement (test adequacy) methods. The practicability of this method was proved by taking the Three Northeastern Provinces of China as the study area. The result shows that there should be 112 test sites in the study area, the distribution of the test sites is uniform, and the environmental representation is high. Test adequacy analysis of the test sites reveals that most of the environmental categories have a test adequacy that reaches 1 in each test period. The method proposed in this paper provides support for the scientific layout of crop varieties test sites and helps to improve the representative and reliability of variety test results while optimizing resources.

Keywords: maize; Multi-environment trials; test sites layout; spatiotemporal classification; spatial stratified sampling; GIS

1. Introduction

Maize (*Zea mays* L) is one of the main food, feed, and industrial material crops produced all over the world, accounting for about 13% of the world's arable land [1]. The sustained high yield and quality of maize are significant to food security and economic stability [2–4] in many areas. Meanwhile, maize is widely distributed, and a rational layout according to the ecological suitability of maize varieties is an important way to reduce the risk of inadequate promotion and increase yield. In order to adapt maize to different planting environments and meet the needs of food safety and economic development, breeders continuously select or improve new maize varieties. Before being planted on a large-scale in the target population of environments (TPE), new maize varieties need to be tested in multi-environment trial (MET) to estimate their productivity and adaptability, as well as determine the appropriate promotion region [5–9].Therefore, determining how to set up or optimize the test sites of MET is key to breeding and promoting new maize varieties in different ecological environments.

Because of the diversification of maize planting environments and the complexity of farming systems, MET sites should be representative of the local ecological environment and planting forms [10–12]. The number of test environments and the location of test sites affect the accuracy and reliability of the test results [13,14]. At present, there are three kinds of methods used to set up and optimize MET sites for crop varieties. The first one relies on expert experience to select test sites, the second involves analyzing the size of test sites using classical sampling theory, and the third entails calculating the test size and sites layout based on the spatial clustering results of the crop planting environment.

The selection of test sites based on expert experience is to choose suitable sites for breeding tests according to professional knowledge, breeding and cultivation experience of breeders in an ecological zone. For example, Troyer [15] pointed out that 200 tests can be used to study future performance of maize, and the more test points, years and mazie ecological zones covered, the more accurate the results. However, this method has no quantitative standards for sites selection and layout, which leads to test sites layout arbitrarily.

Calculating the tests number using classical sampling theory involves calculating the minimum sample size for the accurate evaluation of important crop phenotypic traits in an ecological zone. The aim is to solve problems of insufficient MET sites and the high risk of inadequate variety promotion. The minimum sample size is mainly calculated by the parameter estimation method of important statistics (variance, mean, etc.) of crop variety traits in MET. For example, Piepho [16] proposed a calculation method for determining the number of tests with the aim of accurately estimating the variance of the tested varieties' phenotypes. However, this method is a measure of the tested varieties' stability and can not distinguish the difference in test sizes between different crops and traits. On this basis, Zhe Liu [17] proposed a method for calculating the number of tests required to accurately estimate and compare the means of the tested varieties' phenotypic in MET. Although this kind of method can solve the problem of an insufficient number of tests, it should be pointed out that (1) it was derived from the parameter estimation method for data with a normal distribution, but most crop phenotypic data do not follow a normal distribution, so this method can not be directly applied; (2) the optimization of a test sites layout is not further explored based on the calculation of minimum tests number.

Determination of test site number and layout based on spatial clustering of the crop planting environment is a method that aims to improve the test sites' representativeness of the planting environment. Meteorological, soil, and other data of an ecological region are used to divide the region into several ecological sub-regions by the spatial clustering method, and on this basis, various sampling strategies are used to lay out test sites. XU Naiyin [18] evaluated the discrimination, representativeness and ideality of a cotton variety test environment in the Yangtze River Basin by classifying the cotton variety's ecological region based on different traits. Peterson [19] used MET data to divide 56 sites from the IWWPN (International Winter Wheat Performance Nurse) into seven regions by hierarchical clustering method and further described the links between winter wheat test sites. Zuliang Zhao et al. [20] used the spatial clustering method to cluster the multi-year indicators of the maize planting environment in Jilin Province, China, and proposed a MET site layout strategy that was highly representative of the planting environment. This kind of method takes into account the actual growth environment of the crop and selects test sites with high environmental representativeness, which are closest to the design concept of MET. However, the spatial clustering method uses the mean value of multi-year environmental indicators, and it only expresses the static characteristics of the planting environment and neglects the dynamic characteristics of a changing environment or extreme climate events.

In summary, in order to realize the purpose of MET, the test sites should represent the local planting environment to the greatest extent possible. The classification of a crop planting environment can help us fully cognize the environment, realize the minimum number of test sites (saving resources) while ensuring the test efficiency, which is an essential part of the test site layout. However, the common clustering method for MET site layouts still uses the mean values of multi-year indicators to describe the environment, and this approach ignores changes in environmental characteristics. Maize varieties screened at such test sites are likely to fail to adapt to changing growth conditions, posing planting risks.

Therefore, we proposed a maize variety test site layout method that is based on the spatiotemporal classification of the maize planting environment. In this study, we (1) conducted spatial–temporal classification by using the multi-year indicators of the maize planting environment to analyze the change rules and trends of different environmental categories; (2) determined the minimum number of test sites for each environmental category through spatial stratified sampling method; and (3) studied the layout of the maize variety test sites and developed a method for measuring the reliability of the test sites. The Three Northeastern Provinces of China was taken as an example to verify the practicability of this method.

2. Materials and Methods

2.1. Data Acquisition and Preprocessing

Data for maize growth periods between 1993 and 2013 were obtained from the Crop Growth and Farmland Soil Moisture dataset provided by the National Meteorological Center of China, which records the date of crop development and other data rely on the agricultural meteorological stations. Meteorological data from 1993 to 2013 were obtained from National Meteorological Center of China, including daily average temperatures, daily average rainfall and sunshine duration. The maize planting area data came from county-level economic statistics data of the Ministry of Agriculture. The elevation data came from the SRTM (Shuttle Radar Topography Mission) 90m resolution DEM (Digital Elevation Model) data, and the spatial analysis function in ArcGIS was used to process the DEM data to obtain the slope data. The Three Northeastern Provinces of China and the meteorological stations are shown in Figure 1.

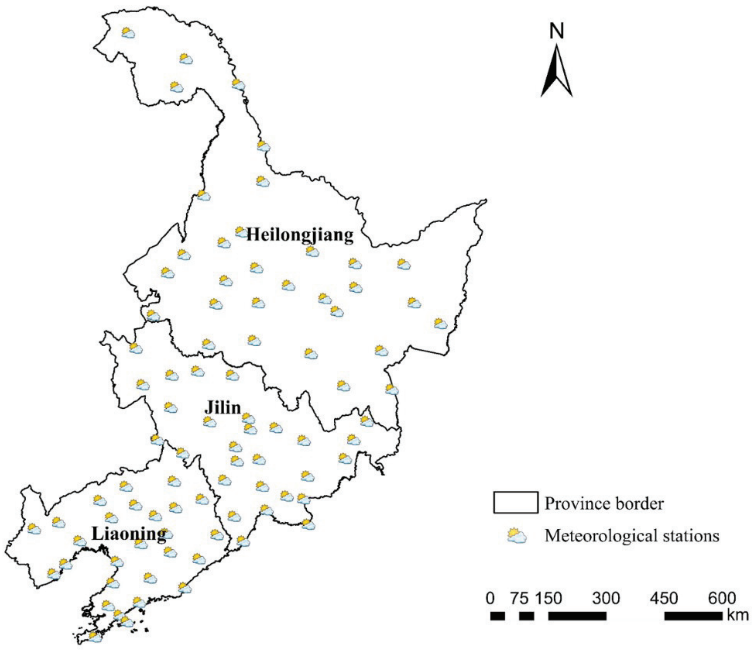

Figure 1. The Three Northeastern Provinces of China and Meteorological stations.

For this research, we established a dataset of multi-year environmental indicators, including accumulated temperature, accumulated precipitation, accumulated sunshine hours, elevation, slope and maize planting area. In order to obtain the raster data covering the whole study area, the accumulated temperature, accumulated precipitation, and accumulated sunshine hours of the meteorological stations were interpolated by spatial interpolation method [21,22]. The above indicators were processed into 10 km × 10 km grids, and the indicators were calculated according to Equations (1)–(3).

1. accT (accumulated temperature) is the cumulative value of the daily average temperature from maize emergence to maturity and the formula is as follow:

$$accT = \sum_{i=1}^{n} t_i \tag{1}$$

where n is the number of days in the maize growth period (from emergence to maturity), and t_i is the daily average temperature of the ith day that is greater than or equal to 10 °C.

2. accP (accumulated precipitation) is the cumulative value of the daily average precipitation from maize emergence to maturity and the formula is as follow:

$$accR = \sum_{i=1}^{n} p_i \tag{2}$$

where p_i denotes the daily average precipitation of the ith day, and n has the same meaning as in Equation (1).

3. accS (accumulated sunshine) hours is the cumulative value of the daily sunshine hours from maize emergence to maturity and the formula is as follow:

$$accS = \sum_{i=1}^{n} s_i \tag{3}$$

where s_i denotes the daily sunshine hours of the ith day, and n has the same meaning as in Equation (1).

2.2. Spatiotemporal Classification of Maize Planting Environment

The spatiotemporal classification of the maize planting environment accounts for the changes in the environment between years. This classification can well represent the development trend of planting environmental changes and is more instructive for current as well as future agricultural activities. The method used here involved the following steps: (1) the number of maize planting environment cluster was determined by R^2 and semi-R^2 statistics;(2) the K-means clustering method was used to cluster the multi-year environmental indicators dataset in the maize growth period (from maize emergence to maturity), and the location information of the study units was considered during the clustering process to ensure that the clustering results had spatial continuity [23]; (3) spatial continuity adjustments were made according to spatial adjustment rules (detailed rules were proposed by Zuliang Zhao et al. [20]); (4) category attribution analysis, the environmental category of each geographical grid (minimum research unit) was counted over the years. For each geographical grid, the most frequent category was selected as its final attribution category, and the occurrence probability of the final attribution category P was defined as the geographical grid's attribution degree to the final attribution category.

Assuming that the sample size of n is divided into k categories, R^2 and semi-R^2 are calculated according to Equations (4) and (5), respectively.

$$R_k^2 = 1 - \frac{\sum_{t=1}^{k} W_t}{T} \tag{4}$$

$$semi - R_k^2 = R_{k+1}^2 - R_k^2 \tag{5}$$

where W_t is the sum of squared deviations in category t, and T is the sum of squares of total deviations for n samples. R^2 is the ratio of sum of squared deviations between interclass to the sum of squares of total deviations for all samples, and it increases as the number of clusters increases. The semi-R^2 statistic is the decreasing of the sum of squared deviations of interclass after merging a new category. The objective of clustering is to have small intraclass differences and as many differences between classes as possible. Therefore, we assume that the number of categories is optimal when both the R^2 and semi-R^2 statistics are relatively large.The spatial structure flow of the spatiotemporal classification is shown in Figure 2.

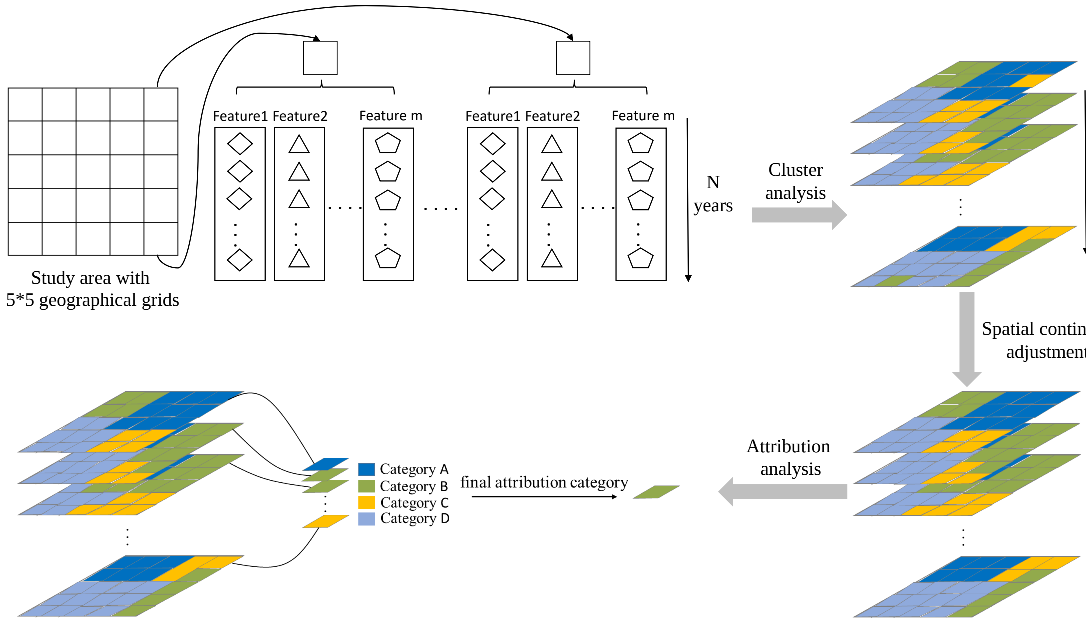

Figure 2. The flowchart of the spatiotemporal classification.

2.3. Calculating the Number of Test Sites

The spatiotemporal classification of the planting environment requires that test sites can fully realize the environmental categories that have appeared in the study area over the years and best reflect the environmental characteristics during the study period. These requirements are necessary to fully test new maize varieties and reduce the risk of crop promotion. According to the attribution degree P, each environmental category was divided into a typical environment region (attribution degree P greater than 50%, representing a region with stable environmental characteristics) and an atypical environment region (attribution degree P less than 50%, representing a region with large fluctuations in environmental characteristics). The atypical environment regions of all environmental categories composed the MIX layer.

The principle of the test sites layout is that each site can represent the main environmental category's characteristics as much as possible. So the number of test sites in each environmental category should be determined by the degree of the environmental category's differentiation [24]. Planting environment factors such as accumulated temperature, altitude, precipitation, etc. [25,26], have obvious spatial correlations, therefore, the classical sampling theory based on the independence of sampling units is not suitable. The spatial sampling method is suitable for objects with spatial correlation [27–30], in the case of fully cognition of each environmental category, the number of test sites required can be determined through the spatial stratification sampling model [31–33].

$$\begin{cases} X = \dfrac{\left(\sum_{h=1}^{L} W_h S_h \sqrt{C_h}\right) \sum_{h=1}^{L} \left(\frac{W_h S_h}{\sqrt{C_h}}\right)}{y + \frac{1}{n} \sum_{h=1}^{L} W_h {S_h}^2} \\ W_h = \frac{n_h}{n} \end{cases} \tag{6}$$

where X is the total number of tests, n_h is the number of geographical grids in environmental category h, n is the total number of geographical grids in the study area, W_h is the weight of environmental category h, S_h is the standard deviation of category h's environmental indicators, C_h is the cost of a single sample in category h, and y is the sampling accuracy (the standard deviation of the overall mean estimate), which becomes larger as the number of samples X increases. The number of samples was determined by calculating the derivative of X and setting the error change threshold.

The unit sampling costs of each environmental category were considered to be the same, and the Neyman optimal allocation principle was used to allocate the sample size of each category [34].

$$X_h = \frac{\frac{W_h S_h}{\sqrt{C_h}}}{\sum_{h=1}^{L} \frac{W_h S_h}{\sqrt{C_h}}} \times X = \frac{n_h S_h}{\sum_{h=1}^{L} n_h S_h} \times X \tag{7}$$

where X_h is the number of tests required in environmental category h, and the other parameters have the same meaning as Equation (6).

There were several indicators for planting environmental classification. The dimensions of each indicator were different, so we used the indicators' deviation matrix as the standard deviation of each environmental category [35].

$$S_h = \sqrt{\frac{1}{n_h - 1} \sum_{j=1}^{n_h} (X_{hj} - \bar{X}_h)'(X_{hj} - \bar{X}_h)} \tag{8}$$

where n_h is the number of geographical grids in category h, X_{hj} is the indicators multidimensional vector of the jth sample in h, and $\bar{x}$ is the indicators' multidimensional vector of category h's gravity center.

According to the ratio of geographical grids' number in typical and atypical environment region in each environmental category to allocate the number of test sites.

$$\begin{cases} \alpha_h = \frac{\mu_h}{n_h} \times \sigma_h \\ \tau = \sum_{h=1}^{L} \sigma_h - \sum_{h=1}^{L} \alpha_h \end{cases} \tag{9}$$

where α_h is the number of test sites that should be selected in the typical environment region of category h, μ_h and n_h are the number of grids of typical environment region and total region in category h, respectively. σ_h is the number of test sites that should be selected in environmental category h. The test cycle of maize is three years, so $\sigma_h = \frac{X_h}{3}$. τ is the number of test sites in the MIX layer, and L is the total number of categories.

2.4. Construction of the Sampling Probability Raster

When setting up a maize test site, we not only consider the environmental representativeness of the test site but also the local cultivation habits, soil conditions, etc. In view of this, this study set the environmental category's representativeness and the county-level maize planting area of the geographical grids as indicators for maize test sites layout. In order to fully recognize all environmental categories, test sites should cover as many environmental categories as possible and best represent the environmental characteristics. The environmental representativeness of the geographical grids in typical environment regions was expressed by the distance from the center of the final category. The environmental representativeness of atypical environment regions was expressed by the changing frequency of the geographical grid attribution category. Combined with the county-level maize planting area, we calculated the sampling probability raster.

Typical environment ragion:

$$\begin{cases} \omega_{jh} = \omega_{jh} + \begin{cases} 1 - \frac{d_{jhk}}{max(d_{jhk}} & \mathrm{j}_k \in h \\ 0 & \mathrm{j}_k \notin h \end{cases} \\ p_{jh} = \frac{\omega_{jh}}{max(\omega_{jh})} \times 100\% \\ prb_{jh} = \begin{cases} (p_{jh} \times w1 + \frac{area_j}{max(area_j)} \times w2) \times 100\% & area_j \neq 0 \\ 0 & area_j = 0 \end{cases} \end{cases} \quad (10)$$

where ω_{jh} is the cumulative representation value of the geographical grid j for the final attribution category h; p_{jh} represents the environmental representation of the geographical grid j for the final attribution category h; d_{jhk} is the distance between the environmental characteristic value of the geographical grid j and the final attribution category h in kth year, and if the geographical grid does not belong to h in kth year, then ω_{jh} plus 0. prb_{jh} is the probability that the geographical grid j is drawn as a test site in category h; $area_j$ is the maize planting area of the geographical grid j, and if $area_j$ is 0, then geographical grid j is not a suitable test site. $w1$ and $w2$ are the weights of the environmental representation and maize planting area, respectively.

Atypical environment region:

$$\begin{cases} p_{jm} = \frac{n_j}{max(n_j)} \times 100\% \\ prb_{jm} = \begin{cases} (p_{jm} \times w1 + \frac{area_j}{max(area_j)} \times w2) \times 100\% & area_j \neq 0 \\ 0 & area_j = 0 \end{cases} \end{cases} \quad (11)$$

where p_{jm} is the representation of the geographical grid j to the MIX layer, and n_j is the frequency of category changes of geographical grid j over the years. prb_{jm} is the probability that geographical grid j is drawn as a test site in the MIX layer. $area_j$, $w1$ and $w2$ have the same meaning as in Equation (10).

In practical applications, it is often expected that the sample sites have a uniform spatial distribution and a low spatial autocorrelation coefficient. Spatially balanced sampling was proved to be an effective and feasible method, which emphasizes the random equal probability and the spatial equilibrium distribution of the sample sites and greatly reduces the occurrence of nonresponsive sample units by filtering the inclusion probability [36]. In this study, we used the spatially balanced sampling method to determine the specific number of test sites layout.

2.5. Reliability Evaluation of Results

When selecting test sites, the construction of sampling probability raster has taken into account the test site's environmental representativeness to the final attribution category or the frequency of change in the test environment's characteristics. However, whether test sites are sufficiently reflective of each environmental category in every test period (the maize test period is usually 3 years in China) is uncertain. In order to judge whether the test sites layout is suitable, this study put forward the "test adequacy" index:

$$\theta_{hc} = \frac{\sum \sigma_{hc}}{X_h} \quad (12)$$

where θ_{hc} is the test adequacy of environmental category h in test period c, $\sum \sigma_{hc}$ is the cumulative number of tests that category h experienced in test period c, and X_h is the number of tests required to fully recognize environmental category h, which is calculated by Equation (7).

If $\theta_{hc} \geq 1$, let θ_{hc} be 1, indicating that the number of tests in the test period c can meet the requirement of fully recognizing environmental category h; $\theta_{hc} < 1$ means that the tests for h are not sufficient in test period c. If all environmental categories can be fully tested in test period c, the variety test results within this test period are considered to be reliable. Otherwise, the test results are considered to be unreliable.

3. Results

3.1. The Result of Spatiotemporal Classification

We set the number of categories as 2–9 and calculate the corresponding R^2 and semi-R^2 statistics to determine the final number of categories, which was 7. Through the spatial continuity adjustments and category attribution analysis of the clustering results for twenty years, the spatiotemporal classification result of the maize planting environment in the three northeastern provinces was obtained (Figure 3). Of the three digits in the legend, the first digit represents the final attribution category of the geographical grid, the second and third digits indicate the geographical grid's attribution degree to the final attribution category. For example, "252" indicates that the final geographical category of the geographical grid is environmental category 2, and the degree of attribution is 52%; "700" indicates that the final geographical category of the geographical grid is environmental category 7, and the degree of attribution is "100%".

Figure 3 reveals that many geographical grids have a low degree of attribution (the paler colors in the same color series) to their final attribution category, and it can be known that many areas have large fluctuations in environmental characteristics. If the mean value of multi-year environmental characteristics is used directly for clustering analysis, the annual change in environmental characteristics will be ignored.

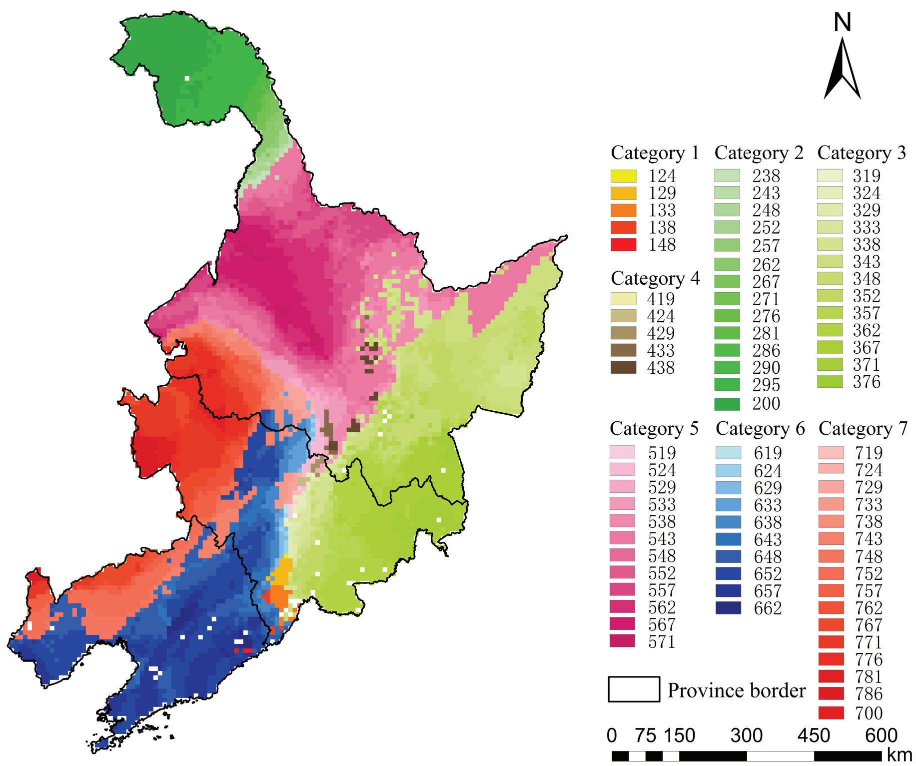

Figure 3. The spatiotemporal classification of the maize planting environment in the Three Northeastern Provinces of China. An environmental category is expressed by color series, with the geographical grids' attribution degree to the final attribution category indicated by pale to deep color gradients, which indicate small to large attribution degrees.

3.2. Test Sites Layout

According to the geographical grids' attribution degree to the final attribution category, the result of spatiotemporal classification was divided into typical environment regions corresponding to each environmental category and the MIX layer (Figure 4). Environmental categories 1 and 4 only covered small areas, and the attribution degree of these geographical grids was less than 50%. So, environmental categories 1, 4, and geographical grids whose attribution degree in other environmental categories were less than 50% composed the MIX layer.

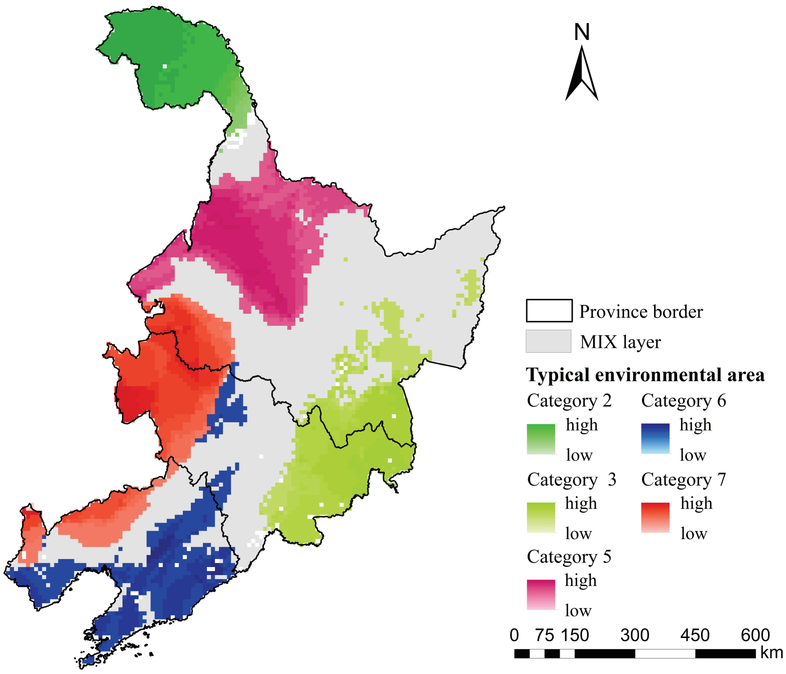

Figure 4. The spatial stratification of environmental categories. An environmental category is expressed by color series, with the geographical grids' attribution degree to the final attribution category indicated by pale to deep color gradients, which indicate small to large attribution degrees.

We calculated the number of test sites in each typical environment region and the MIX layer (Table 1) by using Equations (6)–(9) (the sampling accuracy change threshold was 0.5). The latitude of environmental category 2 is too high, and the annual activity accumulated temperature could not meet the requirement for maize growth; therefore, environmental category 2 was not considered when selecting test sites. Since the maize test period is 3 years in China, the actual number of test sites is 1/3 of the tests required for fully cognition, and the number of test sites for typical environment regions and the MIX layer was divided according to the proportion of geographical grids.

Table 1. Number of Test sites Based on Spatiotemporal Classification of Planting Environment.

Environmental Categories	Testing Number for Fully Cognition	Total Number of Test Sites	Number of Test Sites for Typical Environment Regions	Number of Test Sites for the MIX Layer
category 1	25	8	0	8
category 3	84	28	20	8
category 4	62	21	0	21
category 5	56	19	11	8
category 6	53	18	14	4
category 7	52	18	15	3
Total	332	112	60	52

After calculating the environmental representativeness of the geographical grids and setting the weight of environmental representativeness and the weight of the maize planting area both to 0.5, the sampling probability raster was built (Figure 5). In the sampling probability raster, the deeper the color of the geographical grid, the higher the probability of its selection as the test site in that particular color series (i.e., the same environmental category). It can be seen that the probability at the junctions of environmental categories is higher than that in other regions.

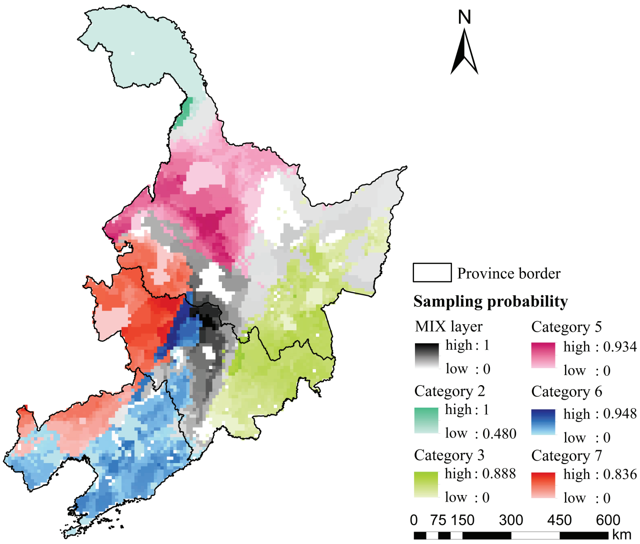

Figure 5. Sampling probability in each environmental stratification. The sampling probability was built on the basis of each environmental stratification. The deeper the color of the geographical grid, the higher the probability of its selection as the test site in that particular color series.

On the basis of the priority of a specific location defined by the sampling probability raster, a set of sampling points with balanced spatial distribution was obtained (Figure 6). The test sites were evenly distributed throughout the study area, and there was a large number of test sites at the junctions of environmental categories. It can be observed from the actual planting area data that the junctions of environmental categories are generally maize-intensive regions, such as Longjiang County in Qiqihar City, Lindian County in Daqing City, Wuchang County in Harbin City, and so on. So, establishing test sites at these junctions is consistent with the actual planting situation.

3.3. Test Adequacy Analysis of Test Sites

The cumulative value of the number of test sites for each environmental category within a test period was calculated, and the test adequacy for each environmental category was calculated by using Equation (12) (Table 2). Analysis of Table 2 shows that the test adequacy for each test period is consistent with the changes in the spatial distribution of the environmental categories. Except for environmental categories 1 and 4, the test adequacy for the environmental categories in each test period mostly reached 1.

The test adequacy of environmental categories 1 and 4 gradually increased, reflecting an increasing trend in the spatial distribution of these two categories. In order to better integrate the characteristics of the environmental development to guide crop breeding and promotion, more test attention is required for environmental categories with large changes in environmental characteristics or an obvious increase in spatial distribution. Five test periods are selected presented as examples: the test adequacy was deemed to be good in test periods 2007–2009, 2008–2010, 2009–2011 and 2010–2012, so the test results according to this test sites layout in these periods have high reliability. The test period 2011–2013 has low test adequacy for environmental categories 5 and 7, indicating that the environmental categories were not fully recognized during this test period, and the test results have low reliability.

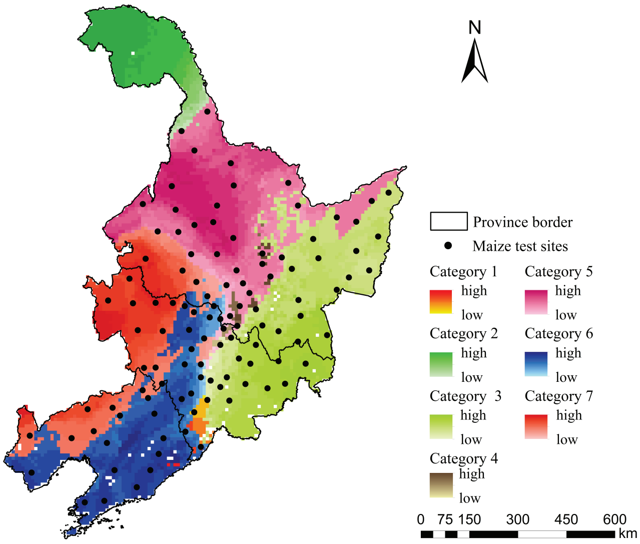

Figure 6. Maize test sites layout based on spatiotemporal classification of planting environment. An environmental category is expressed by color series, with the geographical grids' attribution degree to the final attribution category indicated by pale to deep color gradients, which indicate small to large attribution degrees.

Table 2. Test adequacy for each environmental category in every test period.

Test Period	θ_{1c}	θ_{3c}	θ_{4c}	θ_{5c}	θ_{6c}	θ_{7c}
1993–1995	0	1	0.59	1	1	1
1994–1996	0	1	0.59	1	1	1
1995–1997	0	1	0.19	1	1	1
1996–1998	0.08	1	0.07	1	1	1
1997–1999	0.08	1	0.07	1	1	1
1998–2000	0.08	1	0.07	1	1	1
1999–2001	0	1	0	1	0.54	1
2000–2002	0	1	0	1	0.46	1
2001–2003	0	1	0	1	1	1
2002–2004	0	1	0	1	1	1
2003–2005	0	1	0.04	1	1	1
2004–2006	1	1	1	1	1	1
2005–2007	1	1	1	1	1	1
2006–2008	1	0.59	1	1	1	1
2007–2009	1	0.52	1	1	1	1
2008–2010	1	0.52	1	1	1	1
2009–2011	1	1	1	0.76	1	1
2010–2012	1	1	1	0.76	1	0.82
2011–2013	1	1	1	0.15	1	0.26

4. Discussion

In this study, the spatiotemporal classification of maize planting environment was realized by combining the spatial distribution and temporal variation trends of planting environment characteristics. Then, the minimum number of test sites was calculated by a spatial stratified sampling method, and the layout methods of maize tests sites for typical and atypical environment regions were established. Taking the Three Northeastern Provinces of China as an example, the obtained maize test sites are mainly distributed at junctions between environmental categories. The reason for this phenomenon is that the environmental characteristics change frequently at these junctions. From the perspective of constructing the sampling probability raster, if the MIX layer is not established and the sampling probability raster is directly constructed on the basis of spatiotemporal classification of the maize planting environment, the environmental categories' junctions cannot be selected as test sites because of the low environmental representativeness of the final attribution category. Therefore, in this study, the establishment of the MIX layer as a layer of spatial sampling is of great significance.

In previous studies, most of which have used spatial classification of planting environment to lay out test sites [20,24], planting environment classification methods have been designed to identify the environmental characteristics that have remained relatively stable throughout the years. We referred to Zuliang Zhao's [20] test sites layout method based on the maize planting environment spatial classification to analyze the Three Northeastern Provinces of China. The calculation method, parameter setting of the number of test sites, and the indicators of the sampling probability grid were consistent with this study. A comparison map of the sites layout is presented in Figure 7. Firstly, the number of test sites obtained by the two methods is different: the number of test sites based on spatial classification and spatiotemporal classification is 82 and 112, respectively. The reason for this divergence is that the latter method considers the spatial and temporal environmental variability, while the former only considers the spatial environmental variability, resulting in the deviation of environmental characteristics of the latter is greater than the former. Secondly, the dense areas of the test sites layouts are different. The test sites based on spatial classification are slightly denser in the central and eastern parts of the Three Northeastern Provinces of China than other regions. The test sites based on spatiotemporal classification are mainly concentrated in the middle of the Three Northeastern Provinces of China, mainly because of the frequent changes in environmental characteristics in the central region and the obvious differences in temporal and spatial differentiation. Regions in which these environmental categories change frequently are also maize intensive; therefore, test sites need to be placed in these regions as well.

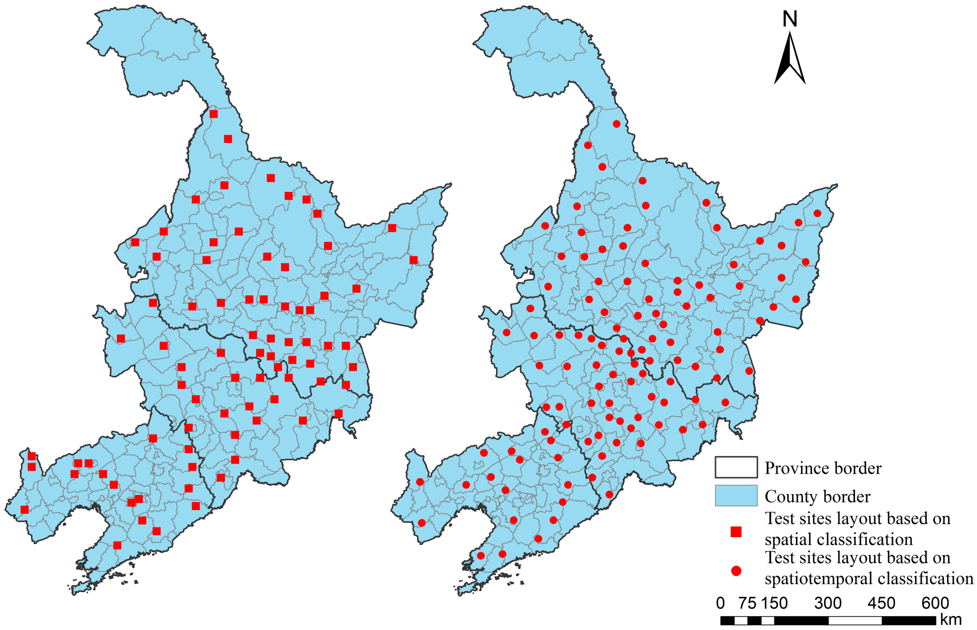

Figure 7. Comparison of test sites layouts based on spatial and spatiotemporal classification of planting environment.

5. Conclusions

This study aims to solve problems such as poor distribution and representativeness of current MET sites of maize varieties, the inability to adapt to the changes in planting environments, and the high risk associated with the promotion of a new maize variety. In this paper, using the Three Northeastern Provinces of China as the study area, we propose a maize variety test sites layout method that is based on the spatiotemporal classification of planting environment. (1) Compared with spatial classification based on mean values of multi-year indicators, the spatiotemporal classification of maize planting environment can express the spatial distribution of planting environment and the trend of annual change. From the perspective of planting environment cognition, this classification method is more suitable for MET at present and even in the future. (2) By distinguishing between typical and atypical environment regions in the planting environment, test sites layout strategies were designed. The test sites were allocated to regions with stable environmental characteristics and those with large fluctuations in environmental characteristics. This approach reduces the risk of new varieties being promoted in regions with fluctuations in environmental characteristics. Combined with the actual planting area of maize, the validity of the test sites layout was verified. (3) The test adequacy index can effectively evaluate the reliability of the test results for a test period. In order to reduce the risk associated with varieties promotion, the varieties testing process needs to cover all environmental categories of TPE, and the test adequacy index can be used to detect whether a crop variety meets the requirement of fully cognition in various planting environments.

This method can also be used to optimize the layout of MET sites in other maize ecological zones or other crops, helps to improve the test efficiency while optimizing resources, or to estimate the deviations in existing test sites layouts being used to evaluate the performance of a new crop variety. The results of this test system can also be used to precisely position new varieties and provide a basis for differentiating cultivation strategies, since the performance of varieties under different environmental categories can be known. In future research, more environmental indicators, such as soil type and organic matter content, can be added to better describe the planting environments. It is also possible to study the test site layout of crop resistance, for example, test sites of lodging resistance.

Author Contributions: X.Z. (Xuli Zan), X.Z. (Xiaodong Zhang) and Z.L.; methodology, S.L. and D.Z.; formal analysis, X.Z. (Xuli Zan), X.Z. (Xiaodong Zhang) and Z.L.; investigation, Z.Z. and W.L.; data curation, X.Z. (Xuli Zan); writing—original draft preparation, All authors contributed to writing—review and editing.

Funding: This research was funded by National Natural Science Foundation of China grant number 41771104 and National Key Research and Development Program of China grant number 2017YFD0300301.

Acknowledgments: We thank the journal's editors and anonymous reviewers for their kind comments and valuable suggestions to improve the quality of this paper.

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

Abbreviations

The following abbreviations are used in this manuscript:

TPE	Target population of environments
MET	Multi-environment trial
DEM	Digital Elevation Model
accT	accumulated temperature
accR	accumulated precipitation
accS	accumulated sunshine

References

1. Zhong, L.; Gong, P.; Biging, G.S. Efficient corn and soybean mapping with temporal extendability: A multi-year experiment using Landsat imagery. *Remote Sens. Environ.* **2014**, *140*, 1–13. [CrossRef]
2. Jingrui, D.; Lizhu, E. Scientific and Technological Innovation of Maize Breeding in China. *J. Maize Sci.* **2010**, *1* , 1–5.
3. Zhao, J.; Wang, R.; Shi, J.; Wang, X. Present Situation and Prospect of Maize at Home and Abroad. *Crops* **2008**, *5* , 5–9.
4. Zhang, Y.E.; Zhu, Q.; Wang, H. Analysis of Consumption Requirement and Production Development Trend of Maize in China. *J. Anhui Agric. Sci.* **2009**, *21* , 10159–10161.
5. Wang, Y.; Li, J.; Chen, B.; Gao, S. Comprehensive evaluation of trial sites for early-mature sorghum cultivar in the spring-planting regions. *J. China Agric. Univ.* **2010**, *15*, 1–7.

6. Gauch, H.; Zobel, R. Accuracy and selection success in yield trial analyses. *Theor. Appl. Genet.* **1989**, *77*, 473–481. [CrossRef]
7. Liu, Z.; Yang, J.; Li, S.; Wang, H.; Li, L.; Zhang, X.; Zhu, D. Optimal method of transforming observables into relative values for multi-environment trials in maize. *Trans. Chin. Soc. Agric. Eng.* **2011**, *27*, 205–209.
8. Putto, C.; Patanothai, A.; Jogloy, S.; Pannangpetch, K.; Boote, K.; Hoogenboom, G. Determination of efficient test sites for evaluation of peanut breeding lines using the CSM-CROPGRO-peanut model. *Field Crops Res.* **2009**, *110*, 272–281. [CrossRef]
9. Löffler, C.M.; Wei, J.; Fast, T.; Gogerty, J.; Langton, S.; Bergman, M.; Merrill, B.; Cooper, M. Classification of maize environments using crop simulation and geographic information systems. *Crop Sci.* **2005**, *45*, 1708–1716. [CrossRef]
10. Yan, W.; Kang, M.S.; Ma, B.; Woods, S.; Cornelius, P.L. GGE biplot vs. AMMI analysis of genotype-by-environment data. *Crop Sci.* **2007**, *47*, 643–653. [CrossRef]
11. Piepho, H.P. Methods for comparing the yield stability of cropping systems. *J. Agron. Crop Sci.* **1998**, *180*, 193–213. [CrossRef]
12. Cooper, M.; DeLacy, I.; Eisemann, R. Recent advances in the study of genotype× environment interactions and their application to plant breeding. In Proceedings of the Australian Plant Breeding Conference, Focused Plant Improvement: Towards Responsible and Sustainable Agriculture, Gold Coast, Australia, 18–23 April 1993; pp. 116–131.
13. Li, T.; Ali, J.; Marcaida, M., III; Angeles, O.; Franje, N.J.; Revilleza, J.E.; Manalo, E.; Redoña, E.; Xu, J.; Li, Z. Combining limited multiple environment trials data with crop modeling to identify widely adaptable rice varieties. *PLoS ONE* **2016**, *11*, e0164456. [CrossRef] [PubMed]
14. Liu, Z.; Li, S.; Zhang, X.; Li, L.; Ma, Q.; An, D.; Zhu, D. Environmental sample size estimation based on variety means estimation and means comparison for multi-environment trial. In Proceedigs of the IEEE 2013 Second International Conference on Agro-Geoinformatics (Agro-Geoinformatics), Fairfax, VA, USA, 12–16 August 2013; pp. 460–465.
15. Troyer, A.F. Phenotypic selection and evaluation of maize inbreds for adaptedness. *Plant Breed. Rev.* **2007**, *28*, 101.
16. Piepho, H.P.; McCulloch, C.E. Transformations in mixed models: Application to risk analysis for a multienvironment trial. *J. Agric. Biol. Environ. Stat.* **2004**, *9*, 123–137. [CrossRef]
17. Liu, Z. Research and Application of Multi-Environment Trial for Phenotying of Crop Varieties. Ph.D. Thesis, China Agricultural University, Beijing, China, 2012.
18. Xu, N.Y. Cotton Mega-Environment Investigation and Test Environment Evaluation Based on GGE Model. Ph.D. Thesis, Nanjing Agricultural University, Nanjing, China, 2012.
19. Peterson, C.; Pfeiffer, W. International winter wheat evaluation: relationships among test sites based on cultivar performance. *Crop Sci.* **1989**, *29*, 276–282. [CrossRef]
20. Zhao, Z.; Zhe, L.; Zhang, X.; Zan, X.; Yao, X.; Wang, S.; Ye, S.; Li, S.; Zhu, D. Spatial Layout of Multi-Environment Test Sites: A Case Study of Maize in Jilin Province. *Sustainability* **2018**, *10*, 1424. [CrossRef]
21. Zhu, L.; Huang, J. Comparison of spatial interpolation method for precipitation of mountain areas in county scale. *Trans. Chin. Soc. Agric. Eng.* **2007**, *23*, 80–85.
22. Song, L.; Tian, Y.; Lun, W.U.; Zhang, H. On Comparison of Spatial Interpolation Methods of Daily Rainfall Data:A Case Study of Shenzhen. *Geo-Inf. Sci.* **2008**, *10* , 566–572.
23. Liu, Z.; Liu, W.; Zan, X.; Feng, W.; Li, S.; Zhang, X. Temporal and Spatial Planting Regionalization Description of Spring Maize in Northeast China Based on Several Years Environmental Characteristics. *Trans. Chin. Soc. Agric. Mach.* **2017**, *6* , 1–9.
24. Wang, S. Spatial distribution of multi-environment trial station for maize varieties—A case Study of Jilin province. Ph.D. Thesis, China Agricultural University, Beijing, China, 2015.
25. Wang, L.; Xiong, W.; Wen, X.; Feng, L. Effect of climatic factors such as temperature, precipitation on maize production in China. *Trans. Chin. Soc. Agric. Eng.* **2014**, *30*, 138–146.
26. Tao, F.; Yokozawa, M.; Xu, Y.; Hayashi, Y.; Zhang, Z. Climate changes and trends in phenology and yields of field crops in China, 1981–2000. *Agric. For. Meteorol.* **2006**, *138*, 82–92. [CrossRef]
27. Zhidong, C.; Jinfeng, W.; Lianfa, L.; Chengsheng, J. Str ata Efficiency and Optimization str ategy of Str atified Sampling on Spatial Population. *Prog. Geogr.* **2008**, *27*, 152–160.

28. Wang, J.F.; Stein, A.; Gao, B.B.; Ge, Y. A review of spatial sampling. *Spat. Stat.* **2012**, *2*, 1–14. [CrossRef]
29. Bao, H.; Yun, X.; Zhang, D.; Li, X.; Liu, X. Accuracy improvement in yield estimation of large-scale grassland by stratified sampling—A case study of lowland meadow. *Acta Agrestia Sin.* **2010**, *18*, 327–332.
30. Wu, L.; Bocquet, M. Optimal redistribution of the background ozone monitoring stations over France. *Atmos. Environ.* **2011**, *45*, 772–783. [CrossRef]
31. Tang, S. Research on Monitoring Stations Layout and Optimization Method of County Cultivated Land Quality Level. Ph.D. Thesis, China Agricultural University, Beijing, China, 2014.
32. Wang, J.F.; Jiang, C.S.; Li, L.F.; Hu, M.G. *Spatial Sampling and Statistical Inference*; Science Press: Beijing, China, 2009.
33. Liu, A.; Wu, Y. A study on simple size allocation in stratified random sampling. *J. Shandong Univ. Financ. Econ.* **2007**, *90*, 49–53.
34. Neyman, J. On the two different aspects of the representative method: the method of stratified sampling and the method of purposive selection. *J. R. Stat. Soc.* **1934**, *97*, 558–625. [CrossRef]
35. Xiang, X.; Jin, X.; Du, X.; Sun, W.; Zhou, Y. Analysis of farmland consolidation implementation status in China based on Ward hierarchical clustering. *Trans. Chin. Soc. Agric. Eng.* **2015**, *31*, 257–265.
36. Stevens, D.L., Jr.; Olsen, A.R. Spatially balanced sampling of natural resources. *J. Am. Stat. Assoc.* **2004**, *99*, 262–278. [CrossRef]