

Article

High-Resolution Crop Mapping and Suitability Assessment in China's Three Northeastern Provinces (2000–2023): Implications for Optimizing Crop Layout

Xiaoxiao Wang ¹ , Huaifu Zhao ^{1,2,*} , Guanying Zhao ¹, Xuzhou Qu ¹, Congjie Cao ¹, Jiacheng Qian ³, Sheng Fu ¹, Tao Wang ¹ and Huiqin Han ¹

- ¹ School of Land Science and Technology, China University of Geosciences, Beijing 100083, China; 3012220018@email.cugb.edu.cn (X.W.); 2012240067@email.cugb.edu.cn (G.Z.); 3012240031@email.cugb.edu.cn (X.Q.); 3012230019@email.cugb.edu.cn (C.C.); jason_fusheng@163.com (S.F.); 3012230010@email.cugb.edu.cn (T.W.); 3012240030@email.cugb.edu.cn (H.H.)
² Key Laboratory of Land Consolidation, Ministry of Natural Resources, Beijing 100035, China
³ Jiangsu Academy of Agricultural Sciences, Nanjing 210014, China; 20250047@jaas.ac.cn
* Correspondence: zhaohuafu@cugb.edu.cn

Abstract

The three northeastern provinces of China are the country's most important grain-producing region, particularly for maize, soybean, and rice, and form its largest commercial grain base. Over the past two decades, cropping structures in this region have undergone notable shifts driven by both climate change and human activities. Generating long-term, high-resolution maps of multi-crop distribution and evaluating their suitability is essential for understanding cropping dynamics, optimizing land use, and promoting sustainable agriculture. In this study, we integrated multi-source satellite imagery from Landsat and Sentinel-2 to map the distribution of rice, maize, and soybean from 2000 to 2023 using a Random Forest classifier. A crop suitability assessment framework was developed by combining a multi-criteria evaluation model with the MaxEnt model. Reliable training samples were derived by overlaying suitability evaluation results with stable crop growth areas, and environmental variables—including climate, topography, soil, hydrology, and anthropogenic factors—were incorporated into MaxEnt to assess suitability. Furthermore, the spatial consistency between actual cultivation and suitability was evaluated to identify areas of misallocated land use. The results show that: (1) the six classification maps achieved an average overall accuracy of 91.05% and a Kappa coefficient of 0.857; (2) the cultivation area of all three crops expanded, with maize showing the largest increase, followed by soybean and rice, and the dominant conversion being from soybean to maize; (3) suitability areas ranked as soybean (376,692 km²) > maize (329,056 km²) > rice (311,869 km²), with substantial spatial overlap, particularly between maize and soybean, suggesting strong competition; and (4) in 2023, highly suitable zones accounted for 57.39% of rice, 39.69% of maize, and 28.89% of soybean cultivation, indicating a closer alignment between actual distribution and suitability for rice, weaker for maize, and weakest for soybean, whose suitable zones were often displaced by rice and maize. These findings provide insights to guide farmers in optimizing crop allocation and offer a scientific basis for policymakers in designing cultivated land protection strategies in Northeast China.

Keywords: crop classification; decision-making; random forest; Google Earth Engine; Maxent; policy and strategies



Academic Editors: Carlo Leifert and Jayme Garcia Arnal Barbedo

Received: 2 September 2025

Revised: 30 October 2025

Accepted: 6 November 2025

Published: 10 November 2025

Citation: Wang, X.; Zhao, H.; Zhao, G.; Qu, X.; Cao, C.; Qian, J.; Fu, S.; Wang, T.; Han, H. High-Resolution Crop Mapping and Suitability Assessment in China's Three Northeastern Provinces (2000–2023): Implications for Optimizing Crop Layout. *Agronomy* **2025**, *15*, 2587. <https://doi.org/10.3390/agronomy15112587>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The supply of staple food crops plays a key role in maintaining stable and sustainable agricultural systems. Over the past few decades, the global yield of major crops has shown a steady upward trend, reaching 9.9 billion tons in 2023—an increase of 28% since 2010 (<https://openknowledge.fao.org>, accessed on 1 November 2025). This growth has been driven by a combination of technological advancements, intensified farming practices, and the expansion of cultivated areas. However, challenges such as climate change, land degradation, regional conflicts, and trade disputes have made global food security increasingly complex and uncertain [1]. Reliable crop distribution data plays a fundamental role in forecasting yields and guiding sustainable agricultural policies [2–4].

The three northeastern provinces of China serve as a major commercial grain-producing region, primarily cultivating rice, maize, and soybean. In 2024, the grain output of the three northeastern provinces accounted for 20.9% of the national total, with rice, maize, and soybean comprising 98.6% of the region's total sown grain area (<https://www.stats.gov.cn/>, accessed on 1 November 2025). As grain crop acreage continues to expand, the region's role as a cornerstone of national food security has become increasingly reinforced. Simultaneously, agricultural structural adjustments—including the conversion of drylands to paddy fields and maize–soybean rotation practices—have become more prevalent. In 2016, the abolition of the maize purchase policy in major maize-producing areas led to a significant drop in market prices, accelerating the conversion of drylands into paddy fields. That same year, in response to shifts in dietary consumption patterns—characterized by decreasing demand for grains and rising demand for soybeans and potatoes—the Chinese government introduced policies to optimize the cropping structure. However, these changes were implemented largely without adequate assessment of crop suitability, resulting in the irrational expansion and distribution of certain crops. Recent global studies have shown that scientifically and rationally adjusting cropping structures (i.e., changing crop types or their spatial distribution) is an effective way to address these issues [5]. Acquiring long-term, high-resolution information on crop evolution characteristics and assessing crop suitability are prerequisites for achieving this goal [6].

In the face of these challenges, robust data and timely information are essential for enabling faster and more evidence-based analyses to improve the utilization and distribution of food, land, and water resources. With the growing accessibility of open-access data, cloud computing infrastructure, and advanced machine learning techniques, satellite remote sensing has emerged as a crucial data source for land cover and land use mapping. Numerous land cover and cropland distribution products are currently available [7,8]. Nonetheless, the intrinsic diversity and complexity of agricultural systems render crop mapping a challenging endeavor. The absence of fine-resolution crop distribution data continues to constrain efforts to achieve global food security goals [9]. Crop mapping typically involves four main components: remote sensing data, distinguishing features, classification algorithms, and sample quality. In terms of remote sensing data, large-scale crop identification has often relied on MODIS imagery, which, despite its temporal richness, suffers from coarse spatial resolution. In contrast, data from GF-1 WFV, Sentinel-2, and the Landsat series offer higher spatial resolution and are better suited for regional-scale, fine-grained crop mapping [10]. Spectral features (e.g., red and near-infrared bands) and their derived vegetation indices (e.g., NDVI, EVI) are among the most commonly used variables due to their ease of acquisition. For instance, Tian et al. [11], Zhang et al. [12], and Zhong et al. [13] utilized NDVI and EVI to map cropland areas in China, North Korea, and the United States, respectively. Classification approaches include both traditional machine learning techniques—such as decision trees, support vector machines, and ran-

dom forests—as well as deep learning architectures, including convolutional and recurrent neural networks. Among these, the random forest algorithm has become one of the most widely used methods due to its speed, robustness, and ease of parameterization [14,15]. High-quality training samples are critical for accurate crop mapping. Although field surveys provide reliable samples, they are often constrained by high costs and low efficiency. Visual interpretation from high-resolution imagery is another common method, but the spectral similarity among different crops frequently makes interpretation challenging. Prior research has shown that selecting training data from validated datasets can be an effective alternative [16,17]. In recent years, several gridded crop distribution datasets have been developed for China, including the three northeastern provinces. For example, Luo et al. mapped the spatial distribution of rice, maize, and wheat throughout China from 2000 to 2015 at a resolution of 1 km utilizing LAI products [18]. Zhai et al. produced the first 30 m resolution crop maps for the three northeastern provinces in 2015 [19], followed by Xuan et al. [20], who generated 30 m resolution rice, maize, and soybean maps from 2013 to 2021. Liu et al. mapped these three crops at 500 m resolution from 2000 to 2020 [21], revealing long-term trends in staple crop dynamics. In addition, Li et al. developed soybean distribution maps of Heilongjiang Province for 1984–2020 [22], while Yang et al. identified crop distributions in the same region for 2015 and 2016 [23]. These data products have provided valuable methodological and technical support for crop mapping in the three northeastern provinces. However, due to the following limitations, it remains necessary to map the long-term crop distribution in the three northeastern provinces. Existing datasets either focus on single crops, have limited time spans, or lack sufficient spatial resolution. Furthermore, the spatial overlaps among distribution maps of different crop types also prevent the direct integration of multiple single-crop distribution maps for use. Therefore, there is a clear need for a long-term, high-resolution crop distribution dataset for the three northeastern provinces to support agricultural monitoring and ensure regional food security.

Land suitability analysis forms a fundamental basis for the management of both current and future land use. Crop suitability assessment is a critical tool for evaluating the potential of land to support specific crops [24], and it plays a vital role in optimizing crop spatial distribution and planting structure. Extensive research has been conducted on the influencing factors and evaluation methods of crop suitability. Many studies have employed parametric approaches and multi-criteria evaluation (MCE) methods to assess land suitability for specific crops. However, these traditional methods are often subject to high levels of subjectivity, with the determination of weights, index classification, and factor selection heavily dependent on expert judgment. In recent years, machine learning approaches—such as Random Forest and the Maximum Entropy (MaxEnt) model—have been increasingly utilized as objective and reliable alternatives for suitability prediction [25]. The MaxEnt model, based on the maximum entropy principle, estimates species distributions in unsampled areas using only presence data [26]. This method has been widely used for suitability assessments of soybean [27], rice [28], citrus [29], and maize [30]. Although the MaxEnt model had been widely applied, it still exhibited several limitations. The selection of sample points was critical to the accuracy of its outputs. Previous studies primarily obtained sample points from museums, agricultural meteorological stations, or field surveys; however, these data often lacked systematic design or sufficient representativeness [31]. In summary, the multi-criteria decision-making model integrates various factors or indicators and incorporates broad and authoritative expert knowledge, demonstrating great potential for achieving highly accurate spatial decision-making. However, since the suitable growth environments differ significantly among crops, applying this method requires the construction of multiple evaluation index systems, which are both time-consuming and costly. By contrast, machine learning approaches such as the Max-

imum Entropy model (MaxEnt) can effectively simplify the complex steps of traditional assessment methods, including indicator selection and weight assignment. Nevertheless, their predictive performance is highly dependent on the quality of the sample points, and if the samples are not appropriately selected, the accuracy and reliability of the results may be compromised. Using a single evaluation method often results in suboptimal performance in aspects such as weight determination and control of sample quality. Therefore, integrating multiple evaluation approaches can serve as an effective strategy to improve the accuracy of crop-specific suitability assessments [32].

In addition to the aforementioned studies that independently focused on either land use and land cover (LULC) change analysis or land suitability evaluation, recent research has begun to integrate LULC dynamics with suitability assessment. These efforts have laid a valuable foundation for the framework proposed in this study [33–35]. However, most existing studies still suffer from several limitations, including broad and unspecific research targets, limited temporal coverage, and a predominant reliance on qualitative approaches for suitability evaluation. In summary, we aim to propose an integrated methodology that combines long-term crop distribution mapping with a comprehensive, objective, and efficient framework for crop suitability assessment. For crop mapping, the Random Forest algorithm was employed to delineate the spatial distribution of three major staple crops—rice, maize, and soybean—across a 23-year period, aiming to address how and why crop distribution patterns have changed and to provide robust data support for agricultural management. For suitability evaluation, we integrate the strengths of a multi-criteria decision-making model and the MaxEnt model to construct a comprehensive framework for assessing crop suitability across the three northeastern provinces of China. This approach aims to identify the optimal spatial configuration of crops and to promote both spatial optimization of agricultural land use and regional food security. In this framework, the multi-criteria decision-making model leverages extensive expert knowledge to ensure the reliability of the sample points, while the MaxEnt model provides a scientific and objective evaluation of the suitability of different crops. This integrated framework not only overcomes the heavy reliance on expert knowledge inherent in the multi-criteria decision-making model but also effectively ensures the quality of sample points required by the MaxEnt model.

This study addresses three key research questions: (1) What are the spatiotemporal evolution patterns of the three primary grain crops—rice, maize, and soybean—in the three northeastern provinces from 2000 to 2023? What are the conversion dynamics among these crops? (2) Where are the suitable growing areas for each crop, considering key environmental and anthropogenic factors, and to what extent do these zones spatially overlap, potentially leading to land use conflicts? (3) To what extent do current crop planting areas align with their respective suitability zones, considering the regions of mismatched land use and potential areas for future agricultural optimization?

2. Materials and Methods

2.1. Study Area

The three northeastern provinces, comprising Heilongjiang, Jilin, and Liaoning provinces (118°11'E to 135°05'E, 38°43'N to 53°33'N; Figure 1), exhibit cold continental (D) climates with dry winters and warm to hot summers. Cropping systems are dominated by single-season cultivation, with major crops including rice, summer maize, spring wheat, and soybean. In 2023, rice, maize, and soybean accounted for 24.32%, 67.5%, and 7.09%, respectively, of the region's total grain production. In recent decades, significant shifts in cropping structure have been observed, driven by climate change and agricultural policies. As a result, the three northeastern provinces have emerged as a representative region

for studying long-term changes in cropping patterns. This study focuses on analyzing the spatiotemporal evolution of crop planting and evaluating crop suitability, providing essential insights for optimizing crop distribution and promoting sustainable agricultural development.

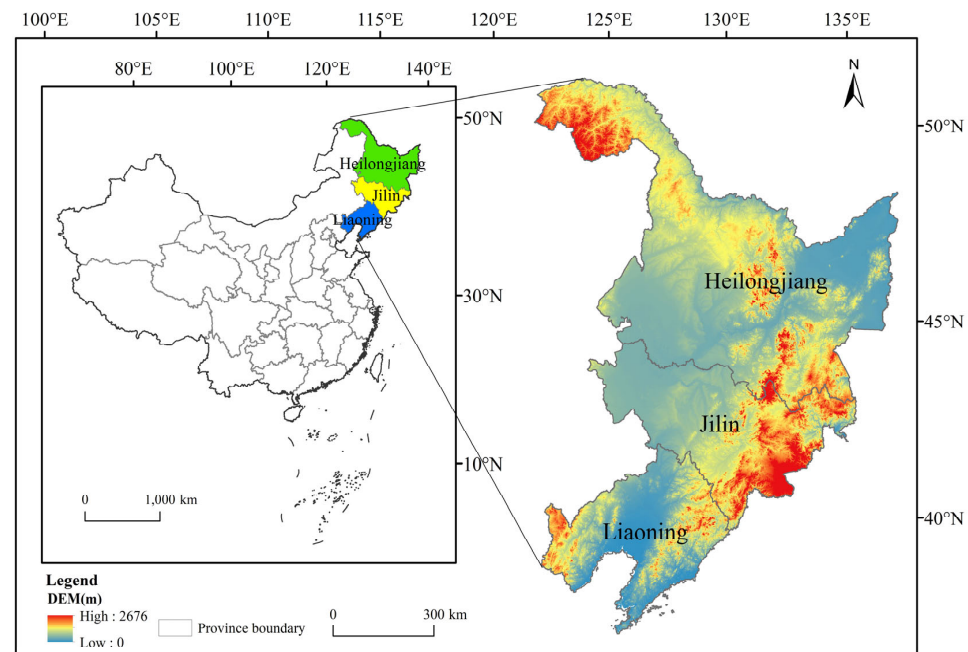


Figure 1. The overview of the study area.

2.2. Workflow of This Study

The research framework and main steps of this study consist of three parts (Figure 2). (1) Crop mapping: This study utilized the Random Forest (RF) method to identify rice, maize, and soybean in the three northeastern provinces from 2000 to 2023 (Figure 2). All data processing was conducted on the Google Earth Engine (GEE) platform (<https://earthengine.google.com/>, accessed on 7 November 2025). First, remote sensing imagery was preprocessed, which involved the detection and elimination of clouds and shadows. Next, multi-source training samples were collected, consisting of both field survey data and automatically extracted samples from existing data products. Finally, multiple feature variables were integrated to characterize the spectral and temporal signatures of different crop types. After applying a cropland mask for major crops, crop classification was performed using the RF algorithm. To ensure consistency in the classification results, all crop identification outputs were ultimately exported at a 30 m resolution. (2) Crop suitability assessment: The core idea of this section was to integrate a multi-criteria decision-making framework with the MaxEnt model to establish a crop suitability assessment framework (Figure 3). This framework was designed to overcome the subjectivity of traditional approaches and the limitations of sample data quality in machine learning methods. Farmland suitability was first assessed using the multi-criteria decision-making model to identify highly suitable areas. These areas were then overlaid with stable crop-growing regions to generate high-quality sample points. Finally, the MaxEnt model was applied to produce suitability distribution maps for the target crops. (3) Implications for crop spatial optimization: The consistency between the current crop distribution and the modeled suitability distribution was analyzed, and targeted policy recommendations were proposed based on the results to guide the optimization of crop spatial layout.

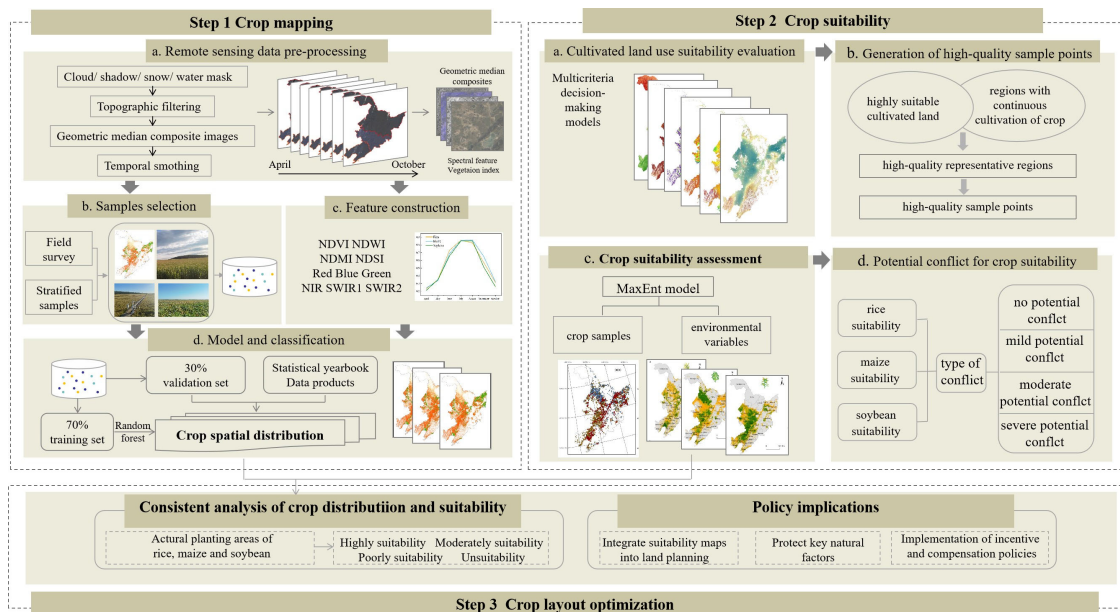


Figure 2. Methodology flowchart of this study.

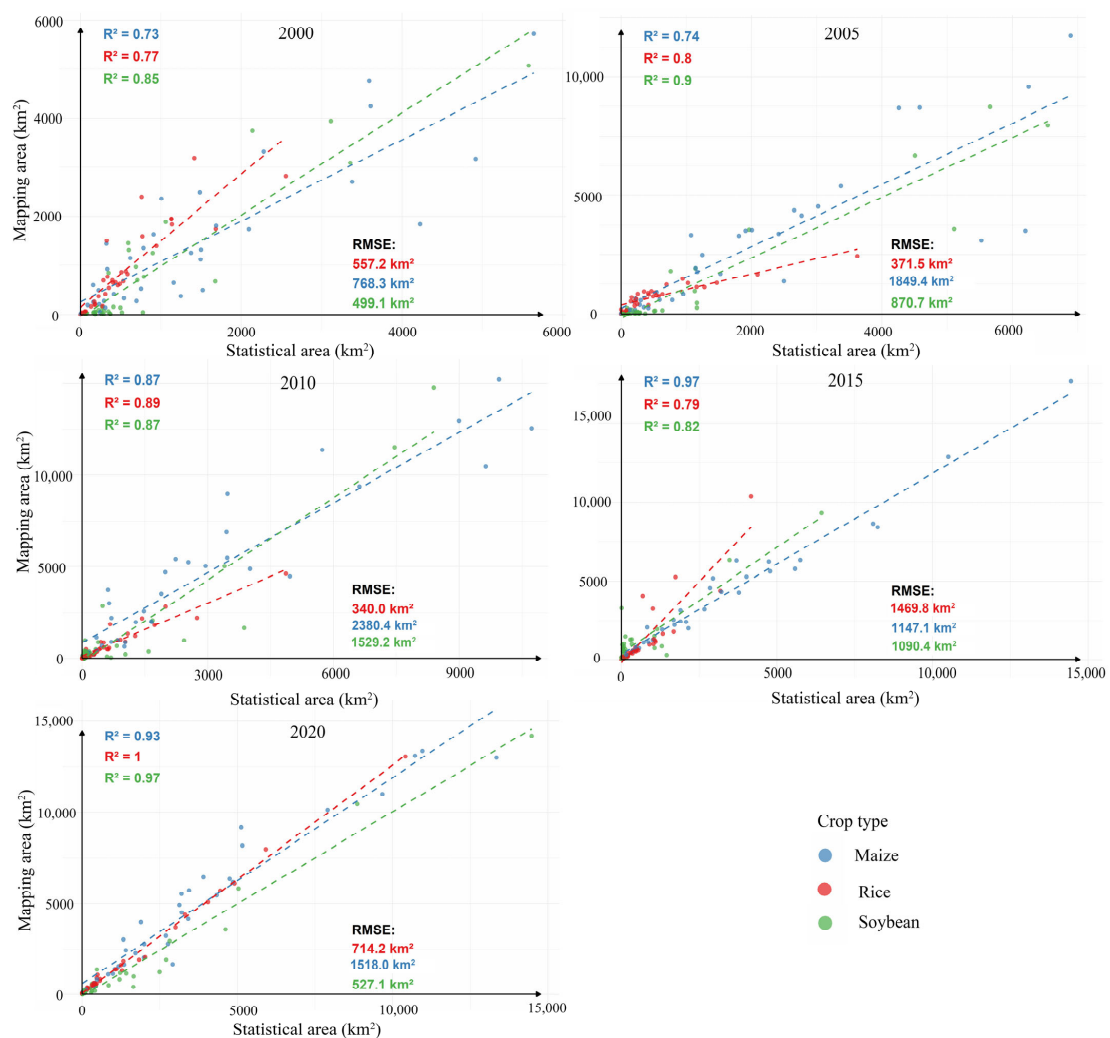


Figure 3. Comparison of crop mapping area with statistical data.

2.3. Crop Mapping

2.3.1. Image Processing

The growing season for major grain crops in the three northeastern provinces typically spans from April to October [21]. To capture crop dynamics during the growing seasons from 2000 to 2023, the quality of multi-source satellite data was evaluated, and the optimal imagery for each target year was selected. A multi-source satellite time-series analysis framework was subsequently established. Specifically, Landsat 7 imagery (2000), Landsat 5 imagery (2005 and 2010), Landsat 8 imagery (2015 and 2020), and Sentinel-2 imagery (2023) were acquired from Google Earth Engine (GEE) using the datasets LANDSAT/LE07/C02/T1_L2, LANDSAT/LT05/C02/T1_L2, LANDSAT/LC08/C02/T1_L2, and COPERNICUS/S2_SR_HARMONIZED, respectively. Cloud and cloud-shadow pixels were eliminated using the QA_PIXEL band for Landsat or the S2Masks module for Sentinel-2. In addition, we integrated ALOS AW3D30 digital elevation model (DEM) data to calculate elevation and slope characteristics and established a terrain mask with elevation <2000 m and slope <30° to exclude non-cultivated areas, as these factors to some extent influence farmers' crop-type choices [36]. In addition, we applied the JRC Global Surface Water dataset to mask permanent water bodies, ensuring that rivers, lakes, and reservoirs would not be misclassified as crops. Table S1 presents the number of images for each period. To address sensor-specific differences in temporal resolution, compositing strategies were adapted accordingly. Landsat 5, 7 and 8 images were composited using 30-day geometric median windows, while Sentinel-2 data were composited using 15-day intervals. Following this, a moving average filter was applied using a ± 30 -day window for Landsat composites and a ± 15 -day window for Sentinel-2 composites. This approach effectively filled residual cloud gaps and smoothed crop phenology profiles. As a result, a continuous, cloud-free time-series dataset for the growing seasons from 2000 to 2023 was generated.

2.3.2. Sample Point Selection

Since rice, maize, and soybean are the predominant grain crops in the study area, sample points for these three crops were collected using two primary approaches (Table 1). First, field surveys were conducted from July to December 2023 to collect ground-truth crop samples for that year. Second, the overlapping areas identified by multiple existing crop classification products were considered reliable sources for sample extraction [17]. For the years 2000, 2005, 2010, 2015, and 2020, sample points were mainly derived from consensus regions among multiple published datasets (Table 1). For rice, distribution maps were extracted from the datasets of Luo [18], Shen [37], and Xuan [20]. The intersecting areas across the same year were identified and used as rice sample sources. Similarly, for maize, overlapping areas from the datasets of Luo [18], Peng [38], and Fu [20] were used. For soybean, common regions identified in the datasets of Li [23] and Liu [39] served as the basis for sample extraction. Stratified sampling was then applied to the consensus regions for each crop and year to obtain representative training samples. Finally, all samples were manually verified using high-resolution satellite imagery throughout the crop growing season to ensure data quality (Supplementary Figure S1).

Table 1. Number and source of sample points.

Year	Crop	Counts	Sources
2000	Rice	1478	[18,37]
	Maize	1385	[18,38]
	Soybean	1238	[23,39]

Table 1. *Cont.*

Year	Crop	Counts	Sources
2005	Rice	1467	[18,37]
	Maize	1480	[18,38]
	Soybean	1301	[23,39]
2010	Rice	904	[18,37]
	Maize	2107	[18,38]
	Soybean	1004	[23,39]
2015	Rice	900	[18,37]
	Maize	2100	[18,38]
	Soybean	1000	[23,39]
2020	Rice	1492	[20,37]
	Maize	1489	[20,38]
	Soybean	1483	[23,39]
2023	Rice	914	Field survey
	Maize	1797	
	Soybean	913	

2.3.3. Feature Construction

A unified set of spectral bands was extracted, including visible (Blue, Green, Red), near-infrared (NIR), and shortwave infrared (SWIR1, SWIR2) reflectance. For Sentinel-2 imagery, additional red-edge bands (B5, B6, B7) were incorporated to improve crop type discrimination. In addition, several spectral indices—such as the Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI), Normalized Difference Moisture Index (NDMI), and Normalized Difference Snow Index (NDSI)—were calculated (Table 2) to build a comprehensive feature set for classification. These indices have been shown to be effective in distinguishing rice, maize, and soybean in previous studies [40–43].

Table 2. Spectral indices for crop classification.

Spectral Indices	Formula	References
Normalized Difference Vegetation Index (NDVI)	$(\text{NIR} - \text{Red}) / (\text{NIR} + \text{Red})$	[44]
Normalized Difference Water Index (NDWI)	$(\text{Green} - \text{NIR}) / (\text{Green} + \text{NIR})$	[24]
Normalized Difference Moisture Index (NDMI)	$(\text{NIR} - \text{SWIR1}) / (\text{NIR} + \text{SWIR1})$	[45]
Normalized Difference Snow Index (NDSI)	$(\text{Green} - \text{SWIR1}) / (\text{Green} + \text{SWIR1})$	[44]

2.3.4. Extraction of Composite High-Confidence Cropland Mask for Major Crops

To accurately identify staple crops under complex land-cover conditions, a high-confidence crop mask was first constructed. This approach was designed to focus classification efforts on the most probable cultivated areas, thereby preventing confusion between crops and non-agricultural features such as forests and water bodies. Considering the potential uncertainties and scale differences in individual remote sensing products, the final mask was generated by overlaying five high-precision single-crop datasets with two land use/land cover (LULC) datasets. This multi-source data fusion strategy aimed to enhance the overall confidence of the mask through cross-validation [46]. Specifically, the five crop-specific datasets for rice, maize, and soybean were first combined, and any pixel labeled as a crop in at least one dataset was retained as a major crop pixel. Next, two LULC datasets were overlaid, and only pixels consistently labeled as cropland in both datasets were retained as cropland pixels. Finally, the cropland layer derived in the second step was used as a mask to extract major crop pixels from the first step, resulting in the final maximum masked area for major crops. All datasets used in this process were resampled

to a spatial resolution of 30 m. Details of the five crop datasets and two LULC datasets are provided in Table 3.

Table 3. Data sources for crop-masked areas.

Name	Resolution	Sources
1 km-grid crop harvesting area dataset for three main crops of China from 2000 to 2015	1 km	[18]
China Crop Dataset-Rice	30 m	[37]
China Crop Dataset-Maize	30 m	[39]
30 m soybean dataset for China from 2000 to 2022	30 m	[38]
Soybean Area of Heilongjiang from 1984 to 2020		[23]
Land use and cover monitoring in China	30 m/1 km	http://www.resdc.cn , accessed on 7 November 2025
China land cover dataset	30 m	[47]

2.3.5. Modeling and Classification

The RF algorithm, a prevalent ensemble learning technique utilizing bagging, was used for crop classification. By aggregating multiple decision trees, RF reduces generalization error and mitigates overfitting, thereby enhancing classification accuracy and robustness. Numerous studies have successfully applied the RF algorithm on the Google Earth Engine (GEE) platform for land cover classification and crop mapping, demonstrating superior accuracy compared to other classification approaches [48,49]. Although increasing the number of decision trees yields only marginal improvements in accuracy, it results in a linear rise in computational cost. We tested different numbers of decision trees, including 10, 50, 100, 150, and 200, following previous research [43,50]. Considering both classification accuracy and computational efficiency, 100 trees were ultimately selected for crop classification. For the remaining parameters, the commonly used default configuration of the RF classifier on the GEE platform was adopted (Table S2). Specifically, the maximum number of features considered at each split was set to the square root of the total number of features, the tree depth was left unrestricted, and the minimum number of samples per leaf node was set to one [16,51]. Since RF requires a substantial quantity of samples to train the classifier effectively, all sample points were randomly divided into two subsets: 70% for training and 30% for accuracy assessment and model validation, which is a commonly used approach in related research [36]. After the initial classification was completed, the results were manually refined using visual interpretation, referencing both satellite imagery and existing classification products to correct potential misclassifications.

2.3.6. Accuracy Assessment

In this study, a confusion matrix was constructed for the validation samples, comparing the true classes with the predicted classes, and overall accuracy (OA), Kappa coefficient, Producer accuracy (PA), and User accuracy (UA) were calculated. OA refers to the proportion of correctly classified samples, with values closer to 1 indicating better overall model performance. The Kappa coefficient measures the classification accuracy, with higher values indicating higher precision. PA represents the probability that a certain crop in the model is correctly classified, while UA represents the probability that a sample predicted as a certain crop actually belongs to that crop class [52].

2.3.7. Data Comparison

The mapped areas of the three major crops were compared with official statistical records and existing crop classification products. This study evaluates model performance using two metrics: the coefficient of determination (R^2) and the root mean square error (RMSE). R^2 quantifies the correlation between the crop statistical area and the model-predicted area, thereby assessing the model's goodness of fit. RMSE measures the magni-

tude of the error between the crop statistical yearbook area and the model-predicted area, reflecting the accuracy of the model's predictions. The calculations of R^2 and RMSE were based on comparisons between the mapped crop areas and the statistical yearbook data for 36 prefecture-level cities in each year. We performed the calculations using RStudio software (Version 2023.03.1). Official crop area statistics were obtained from the Heilongjiang Statistical Yearbook, Jilin Statistical Yearbook, Liaoning Statistical Yearbook, and the China Statistical Yearbook. The mapped crop areas for the 36 cities were generated on the Google Earth Engine (GEE) platform. Using the regional statistics function, the pixel areas of each crop category were aggregated according to prefecture-level administrative boundaries.

2.4. Crop Suitability Assessment

2.4.1. Multicriteria Suitability Evaluation Model for Farmland Suitability Assessment

The assessment of farmland suitability using a multi-criteria decision-making model generally involved four steps: (i) defining the evaluation objects, (ii) selecting evaluation indicators and establishing their classification criteria, (iii) determining indicator weights, and (iv) performing weighted overlay of all indicators followed by suitability classification.

The evaluation objects were the distribution areas of rice, maize, and soybean in 2023. Based on a literature review and expert scoring, representative indicators were selected and classified across six dimensions: climate, topography, soil, hydrology, locational conditions, and spatial morphology (Table 4). The importance of each indicator was then determined through a combination of the Analytic Hierarchy Process (AHP) and the entropy weighting method (EWM) [53]. The AHP relied on expert knowledge and subjective judgment, whereas the EWM emphasized the dispersion characteristics of the data and provided greater objectivity. By combining the two approaches, subjective bias arising from expert judgment was reduced, while objective bias caused by data quality was minimized, thereby making the final decision outcomes more robust and reliable. More details on the AHP are provided in Supplementary materials (Table S3–S9) [54–65]. The specific data sources and main parameters for each indicator are provided in the Supplementary materials (Table S10).

To ensure the scientific validity of the combined weights, this study utilized the Lagrange multiplier method to derive the integrated weight of both AHP and EWM (Table 5) [53].

$$w_i = \frac{\sqrt{\gamma_i * \delta_i}}{\sum_{i=1}^n \gamma_i * \delta_i} \quad (1)$$

where w_i represents the combined weight, γ_i represents the AHP weight result, and δ_i represents the EWM weight result.

Then, the farmland suitability score was calculated using a weighted overlay of multiple indicators, as expressed by the following equation:

$$ES = \sum_{i=1}^n x_i * w_i \quad (2)$$

where ES represents the evaluation score, x_i represents the score of each indicator, and w_i represents the comprehensive weight of each indicator.

Finally, the farmland suitability was classified into three levels—highly suitable, moderately suitable, and marginally suitable—based on the natural breaks method.

Table 4. Evaluation indicator system for cultivated land use suitability.

First Level Indicators	Secondary Indicators	Highly Suitable (100)	Very Suitable (80)	Moderately Suitable (60)	Less Suitable (40)	Highly Unsuitable (20)
Climate	≥10 °C accumulated temperature (+)	≥4500	3500–4500	2500–3500	1800–2500	<1800
	annual average precipitation (+>)	≥800	700–800	600–700	450–600	<450
Topography	elevation (−)	<200	200–500	500–800	800–1000	≥1000
	slope (−)	≤2	2–8	8–15	15–25	≥25
	soil layer thickness (+)	≥150	100–150	60–100	30–60	<30
	soil organic matter (+)	≥40	30–40	20–30	10–20	<10
Soil		loam, silty loam, sandy clay loam, silty clay loam, clay loam, and sandy loam		clay, sandy clay, and silty clay	-	sandy soil and loamy sand
	soil texture (+)					
	pH (+)	6–7.9	5.5–6 or 7.9–8.5	5–5.5 or 8.5–9	4.5–5 or 9–9.5	<4.5 or >9.5
	soil moisture (+)	>0.39	0.28–0.39	-	0.18–0.28	<0.18
	potential soil erosion (+)	<4.25	4.25–10.98	10.98–21.31	21.31–38.21	<38.21
Hydrology	distance to water sources (+)	<2	2–5	5–10	10–20	≥20
	groundwater depth (−)	0–10	10–30	30–100	100–200	≥200
Location conditions	distance to roads (−)	≤2	2–5	5–10	10–20	≥20
	distance to rural settlements (−)	≤2	2–5	5–10	10–20	≥20
Spatial form	farmland aggregation (+)	>93.81	89.71–93.81	85.47–89.71	80.77–85.47	<80.77

Note: The potential soil erosion degree was calculated using the Invest model, while agricultural land agglomeration was determined using Fragstats v4.2.1 software.

Table 5. Weights of cultivated land use suitability indicators.

First Level Indicators	Secondary Indicators	Weight Calculated by the AHP Method	Weight Calculated by the EWM	Composite Weight
Climate	≥10 °C accumulated temperature (+)	0.2347	0.0480	0.1244
	annual average precipitation (−)	0.0587	0.0871	0.0838
Topography	elevation (−)	0.032	0.0483	0.0460
	slope (−)	0.0641	0.0405	0.0597
	soil layer thickness (+)	0.1239	0.0578	0.0991
	soil organic matter (+)	0.039	0.2803	0.1225
Soil	soil texture (+)	0.0702	0.0187	0.0424
	pH (+)	0.0267	0.0814	0.0546
	soil moisture (+)	0.0168	0.0978	0.0475
	potential soil erosion (−)	0.0168	0.0337	0.0279
Hydrology	distance to water sources (−)	0.0836	0.0489	0.0749
	groundwater depth (−)	0.0836	0.0257	0.0543
	distance to roads (−)	0.0481	0.0404	0.0517
Location conditions	distance to rural settlements (−)	0.0481	0.0600	0.0629
Spatial form	farmland aggregation (+)	0.0536	0.0316	0.0482

2.4.2. MaxEnt Model for Crop Suitability Assessment

MaxEnt version 3.4.3 (<https://github.com/mrmaxent/Maxent/releases/tag/v3.4.3>, accessed on 7 November 2025) was employed to construct the crop suitability distribution model. This model, grounded in ecological niche theory, estimates the most probable distribution of a target species by identifying the maximum entropy distribution under constraints derived from the environmental characteristics of known presence locations. It has been widely used to predict habitat suitability within a given study area [27,66]. The modeling process involved the collection of crop occurrence records and environmental variables, parameter tuning, and accuracy assessment.

(1) Crop Distribution Records

Geographic distribution data for the three major crops were extracted from the interpreted crop classification results. First, crop distribution maps from 2010, 2015, 2020, and 2023 were spatially overlaid to identify planting dynamics for each field parcel. Regions exhibiting continuous cultivation of rice, maize, or soybean were then selected, as these stable parcels were considered representative of actual crop occurrences. To improve the quality of the samples and enhance their spatial representativeness, we subsequently combined the results of the multi-criteria suitability evaluation to extract highly suitable farmland areas. Using the ‘Extract by Mask’ tool in ArcGIS 10.8 software, we screened the stable plots generated in the previous step to obtain highly suitable and stable representative areas. Then, ArcGIS was used to generate random points (with a minimum allowed distance of 1 km) for each crop within these areas. The planting records were converted into CSV format, and samples falling outside the defined environmental variable ranges were filtered out.

(2) Environmental variables

Drawing on previous studies and the physiological characteristics of the three major crops, a total of 36 environmental variables potentially affecting crop suitability were selected. These included 20 climatic variables, 8 soil variables, 2 topographic variables, 2 hydrological variables, and 4 anthropogenic variables (Table 6). Among them, 19 bioclimatic variables were sourced from the WorldClim database (<https://www.worldclim.org/>, accessed on 7 November 2025), characterizing monthly, seasonal, and annual variations in temperature and precipitation. The accumulated temperature above 10 °C, representing thermal conditions critical for crop growth, was sourced from the Resource and Environmental Science Data Center of the Chinese Academy of Sciences (<https://www.resdc.cn/>, accessed on 7 November 2025). Soil properties were derived from the high-resolution national soil grid dataset provided by the National Earth System Science Data Center [67]. A 30 m resolution Digital Elevation Model (DEM) from NASA was used to calculate slope. River data and groundwater table depth were obtained from OpenStreetMap (<https://www.openstreetmap.org/>, accessed on 7 November 2025) and the work of Fan et al. [65], respectively. Road vectors and rural settlement data were collected from OpenStreetMap and the Amap API (<https://developer.amap.com/tools/picker>, accessed on 7 November 2025). Euclidean distance tools were employed to calculate the distance to roads, rivers, and rural settlements. Spatial population distribution and GDP data were also retrieved from the Resource and Environmental Science Data Center of the Chinese Academy of Sciences. All environmental variables were resampled to 1 km × 1 km raster layers and reprojected to the WGS_1984_Albers coordinate system. Variables were exported in ASCII format. To reduce the risk of overfitting in the MaxEnt model caused by multicollinearity and spatial autocorrelation among variables, Spearman correlation analysis was performed using Origin software (version 2024) [68]. Pairs of variables with correlation coefficients greater than 0.8 (see Figure S2) were identified. For each correlated pair, the variable with the lower contribution to the initial MaxEnt model (using all variables) was

removed. As a result, 21 environmental variables were retained for the rice and maize suitability models, and 20 variables were retained for the soybean model.

Table 6. Environmental variables selected for different crops in the MaxEnt model.

Type	Variables	Define (Unit)	Rice	Maize	Soybean
Climate	Bio_1	Annual mean temperature (°C)			✓
	Bio_2	Annual mean diurnal range (°C)	✓	✓	✓
	Bio_3	Isothermality (%)	✓	✓	✓
	Bio_4	Temperature seasonality (°C)		✓	
	Bio_5	Maximum temperature of warmest month (°C)			✓
	Bio_6	Minimum temperature of the coldest month (°C)			✓
	Bio_7	Annual temperature range (°C)	✓	✓	
	Bio_8	Mean temperature of wettest quarter (°C)			
	Bio_9	Mean temperature of driest quarter (°C)	✓		
	Bio_10	Mean temperature of warmest quarter (°C)			
	Bio_11	Mean temperature of coldest quarter (°C)			
	Bio_12	Annual precipitation (mm)	✓		
	Bio_13	Precipitation of wettest month (mm)	✓		
	Bio_14	Precipitation of driest month (mm)	✓		✓
	Bio_15	Precipitation seasonality (mm)	✓	✓	✓
	Bio_16	Precipitation of wettest quarter (mm)		✓	
	Bio_17	Precipitation of driest quarter (mm)			
	Bio_18	Precipitation of warmest quarter (mm)		✓	✓
	Bio_19	Precipitation of coldest quarter (mm)		✓	
Soil	At10	≥10 °C accumulated temperature	✓	✓	
	SOM	Soil organic matter (g/kg)			✓
	Thickness	Soil thickness (cm)	✓	✓	✓
	pH	Soil pH value			
	CEC	Cation exchange capacity (cmol (+)/kg)	✓	✓	
	Cly	Proportion of clay particles (%)	✓	✓	✓
	SLIT	Proportion of slit particles (%)			
	SAND	Proportion of sand particles (%)	✓	✓	✓
Terrain	Bd	Bulk density (g/cm ³)	✓	✓	✓
	Slope	Slope (°)	✓	✓	✓
Hydrology	DEM	Elevation/m	✓	✓	✓
	WTD	Depth to water table (m)	✓	✓	✓
Human activities	D_Water	Distances to river networks (m)	✓	✓	✓
	D_Road	Distances to main roads (m)	✓	✓	✓
	D_Residence	Distances to residence (m)	✓	✓	✓
	GDP	Gross Domestic Product (10 ⁴ yuan/km ²)	✓		✓
	POP	Population count (person/km ²)		✓	

(3) Model setting and evaluation

To calibrate the MaxEnt model, 25% of the crop distribution data were randomly selected as test data for model validation. The maximum number of background points was set to 10,000. To reduce model uncertainty, a bootstrap resampling procedure was performed with 10 replicates, and the average of the results was used. All other parameters were maintained at their default settings. The regularization multiplier was used to control model complexity and alleviate overfitting, while the number of replicates determined the iterations for random subsampling. The options “Create response curves” and “Do jackknife to measure variable importance” were enabled to assess the contribution and influence of individual environmental variables. Model performance was assessed by the receiver operating characteristic (ROC) curve, which plots the true positive rate (sensitivity) against the false positive rate (1-specificity). The area under the ROC curve (AUC) was used

as an indicator of model accuracy, with values closer to 1 indicating superior predictive performance.

2.5. Potential Conflict Identification for Crop Suitability

When two or more crops exhibit similar suitability levels for the same land parcel, potential land use conflicts may arise. In this study, four suitability levels for each of the three crops were combined, resulting in 64 possible combinations. These combinations were subsequently classified into four conflict types (Table 7). When a plot contains two types of poorly suitable or unsuitable categories, it is classified as no conflict. When a plot includes three types of poorly suitability or two types of moderate suitability, it is defined as weak conflict. When a plot contains one type of high suitability along with one type of moderate suitability, or three types of moderate suitability, it is categorized as moderate conflict. When a plot contains two or three types of high suitability, it is classified as severe conflict.

Table 7. Type of three crop suitability conflict and their expected changes.

Rice Suitability-Maize Suitability-Soybean Suitability	Primary Types of Conflict	Second Level Types of Conflict	Expected Changes
111 211, 311, 312, 321, 322, 411, 412, 421, 422 121, 131, 132, 141, 142, 231, 232, 241, 242 112, 113, 114, 123, 124, 213, 214, 223, 224 222	No potential conflict	Unsuitable for all three crops Rice suitability dominance Maize suitability dominance Soybean suitability dominance	Develop other agricultural or non-agricultural uses besides the three main crops. It is more suitable for rice cultivation. It is more suitable for maize cultivation. It is more suitable for soybean cultivation.
222 331, 332, 221	Mild potential conflict	Mild conflict among the suitability of the three crops Mild conflict between rice suitability and maize suitability	Combine farming and breeding with flexible crop rotation. Suitable for rice or maize cultivation, but with a lower suitability than that in the moderate conflict zones between rice and maize.
313, 323, 212		Mild conflict between rice suitability and soybean suitability	Suitable for rice or soybean cultivation, but with a lower suitability than that in the moderate conflict zones between rice and soybean.
122, 133, 233		Mild conflict between maize suitability and soybean suitability	Suitable for maize or soybean cultivation, but with a lower suitability than that in the moderate conflict zones between maize and soybean.
333, 343, 433, 334	Moderate potential conflict	Moderate conflict among the suitability of the three crops	Determine the optimal crop through a multi-objective decision optimization model considering economic benefits, ecological impacts, and policy orientation.
341, 342, 431, 432		Moderate conflict between rice suitability and maize suitability	Suitable for rice or maize cultivation, but with a lower suitability than that in the severe conflict zones between rice and maize.
314, 324, 413, 423		Moderate conflict between rice suitability and soybean suitability	Suitable for rice or soybean cultivation, but with a lower suitability than that in the severe conflict zones between rice and soybean.
134, 143, 234, 243		Moderate conflict between maize suitability and soybean suitability	Suitable for maize or soybean cultivation, but with a lower suitability than that in the severe conflict zones between maize and soybean.
444 441, 442, 443 414, 424, 434 144, 244, 344	Severe potential conflict	Severe conflict among the suitability of the three crops Severe conflict between rice suitability and maize suitability Severe conflict between rice suitability and soybean suitability Severe conflict between maize suitability and soybean suitability	Determine the optimal crop through a multi-objective decision optimization model considering economic benefits, ecological impacts, and policy orientation. Suitable for rice or maize cultivation. Suitable for rice or soybean cultivation. Suitable for maize or soybean cultivation.

Note: Suitability levels 1, 2, 3, and 4 represent unsuitable, poorly suitable, moderately suitable, and highly suitable, respectively.

In addition, all pairwise combinations among the four suitability levels between any two crops were evaluated, resulting in 16 possible combinations (Table 8). These combinations were also categorized into four levels of conflict. Land parcels composed of two highly suitable crop types were classified as experiencing severe conflict. Parcels containing either two moderately suitable types or one highly suitable and one moderately suitable type were categorized as having moderate conflict. Parcels with either two marginally suitable types or one marginally suitable and one moderately suitable type were considered to have mild conflict. All other combinations were classified as no conflict.

Table 8. Conflict type combinations between two crop suitability.

Crop1 Suitability—Crop2 Suitability	Primary Types of Conflict	Second Level Types of Conflict
11	No potential conflict	Unsuitable for two crops
21, 31, 41, 42		Crop 1 suitability dominance
12, 13, 14, 24		Crop 2 suitability dominance
22, 23, 32	Mild potential conflict	Mild conflict between two crops
33, 34, 43	Moderate potential conflict	Moderate conflict between two crops
44	Severe potential conflict	Severe conflict between two crops

Note: Suitability levels 1, 2, 3, and 4 represent unsuitable, poorly suitable, moderately suitable, and highly suitable, respectively.

3. Result

3.1. Accuracy Assessment of Crop Mapping

To assess the reliability of the crop mapping results, confusion matrices were constructed using the sample sets for the three crops (Table 9). An overall classification accuracy above 85% was achieved for all crop maps in the three northeastern provinces from 2000 to 2023, with the highest accuracy recorded in 2020 (95.4%) and an average accuracy of 91.05% over the study period. The coefficient for Kappa fluctuated between 0.795 and 0.931. The peak was observed in 2020, and the five-year average was 0.857. Among the three crops, rice exhibited the highest classification performance, with both producer accuracy (PA) and user accuracy (UA) exceeding 90%, averaging 95.36% and 94.53%, respectively. For maize, both PA and UA were above 77%, with average values of 89.97% and 89.81%. Soybean showed relatively lower accuracies, with both PA and UA above 75%, averaging 86.11% and 89.05%, respectively. In summary, rice demonstrated the highest classification accuracy, followed by maize, with soybean exhibiting the lowest accuracy among the three crops.

Table 9. Confusion matrix for crop mapping.

Year	Class	Rice	Maize	Soybean	PA (%)	UA (%)	OA (%)	Kappa
2000	Rice	444	12	6	96.77	92.02	86.37	0.795
	Maize	15	340	69	77.41	83.08		
	Soybean	3	61	298	83.89	82.90		
2005	Rice	424	17	13	95.55	93.46	89.79	0.847
	Maize	18	360	40	84.46	90.36		
	Soybean	10	34	361	89.29	85.59		
2010	Rice	259	12	4	94.18	91.2	91.01	0.849
	Maize	23	626	24	93.02	91.25		
	Soybean	2	48	259	83.82	90.24		
2015	Rice	261	5	1	97.75	98.86	95.09	0.912
	Maize	2	611	18	96.83	94.44		
	Soybean	1	31	251	88.69	92.96		
2020	Rice	441	10	1	97.57	96.5	95.40	0.931
	Maize	10	387	14	94.16	93.03		
	Soybean	6	19	417	94.34	96.53		
2023	Rice	253	17	0	90.36	95.11	88.61	0.805
	Maize	8	494	24	93.92	86.67		
	Soybean	5	59	210	76.64	86.07		

3.2. Comparison with Statistical Yearbook Data

To validate the classification results at the regional scale, the mapped crop areas for each prefecture-level city were compared with official statistics obtained from agricultural yearbooks (Figure 3). Due to the unavailability of crop area statistics for 2023, only data from 2001 to 2021 were used to derive statistical areas for the years 2000 to 2020; thus, crop areas for 2023 were not included in the comparison. For rice, the coefficient of determination (R^2) ranged from 0.77 to 1.00, and the RMSE ranged from 371.5 to 1469.8 km². For maize, R^2 ranged from 0.73 to 0.97 and RMSE from 768.3 to 2270.2 km². For soybean, R^2 ranged from 0.82 to 0.97, with RMSE values between 499.1 and 1594.8 km². Specifically, rice showed the lowest R^2 (0.77) in 2000 and the highest (1.00) in 2020; maize had the lowest R^2 (0.73) in 2000 and the highest (0.97) in 2015; soybean exhibited the lowest R^2 (0.82) in 2015 and the highest (0.97) in 2020. Overall, the dataset demonstrated reliable accuracy at the regional level.

3.3. Spatial Temporal Distribution of Three Major Crops from 2000 to 2023

The mapped areas of rice, maize, and soybean were validated against yearbook-reported statistics (Figure 4). Overall, except for soybean in 2020—where the mapped area was slightly lower than the statistical area—all other years showed larger mapped areas than those reported in the yearbooks. However, the trends of change for the three crops were generally consistent between the two methods. Maize and rice exhibited overall increasing trends, while soybean showed a declining trend. Specifically, maize area increased steadily before 2015 and then slightly declined afterward. Soybean area experienced a modest increase before 2010, followed by a slight decrease, and began to recover after 2015. Rice area steadily increased until 2020, after which it declined slightly.

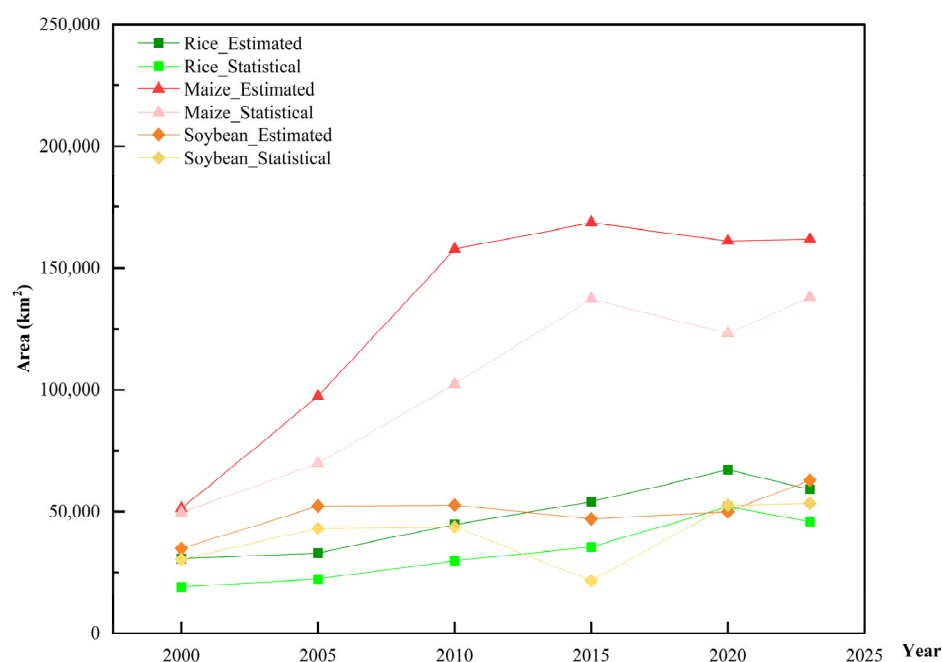


Figure 4. Verification of the accuracy of classified areas.

The three crops exhibited pronounced spatial heterogeneity in their distribution (Figure 5). Soybean was primarily concentrated in Heilongjiang Province, particularly in Heihe City. Rice was mainly cultivated along river systems, with the largest planting areas found in the Sanjiang Plain. The area under corn cultivation has continuously expanded, predominantly across the Songnen Plain and the Liaohe Plain. From 2000 to 2023, maize experienced the largest increase in cultivated area, followed by soybean, while rice showed the smallest expansion. Analysis of spatial transfer patterns among the three major

crops revealed that the dominant conversion type was the shift from soybean to maize (Figure S3, Table S9).

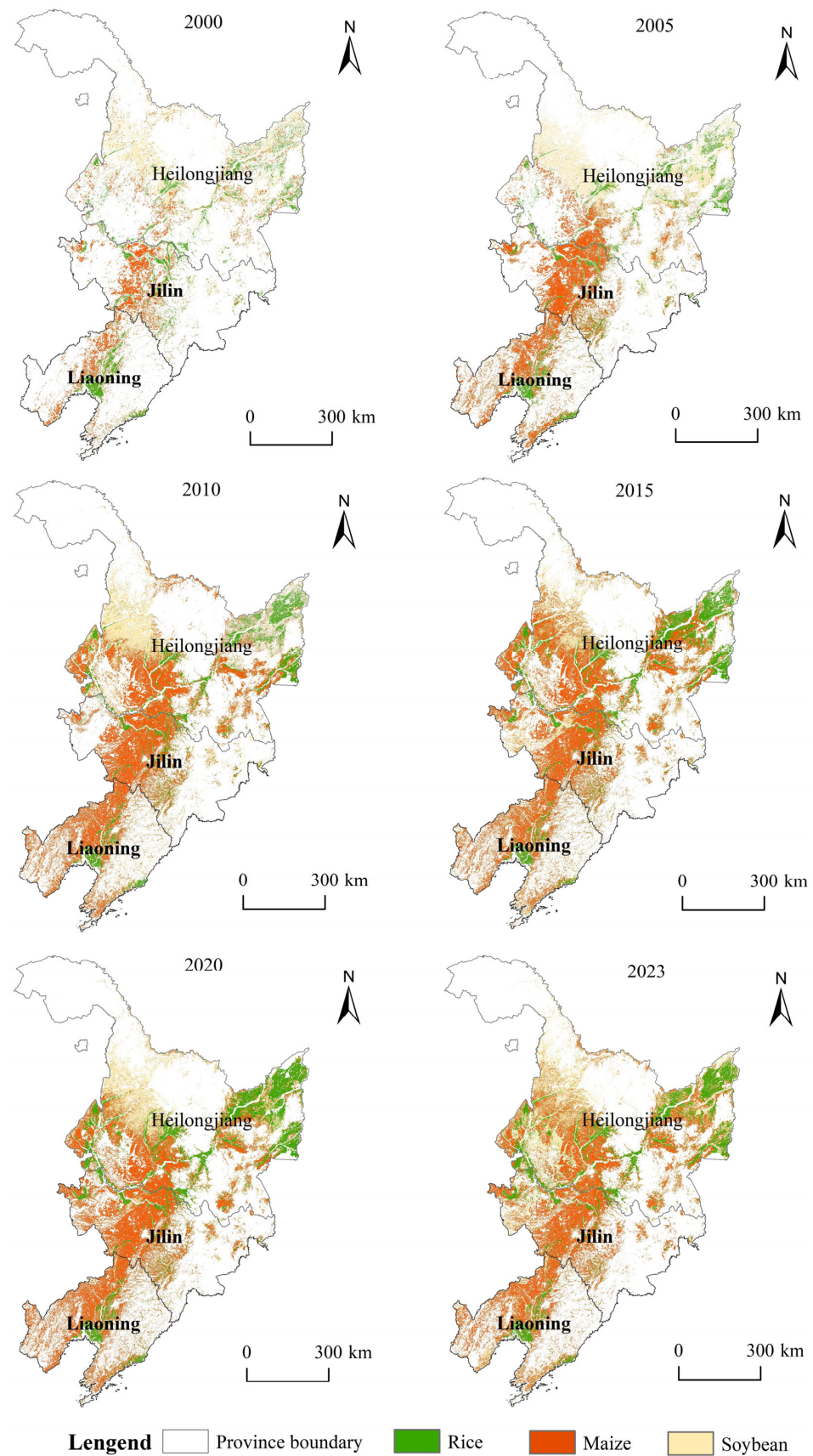


Figure 5. Crop classification map from 2000 to 2023.

3.4. Potential Distribution and Suitability Area of Crop

First, a farmland suitability map was generated using the multi-criteria suitability evaluation model, in which suitability scores were classified into three levels: highly suitable, moderately suitable, and marginally suitable (Figure S4). The highly suitable farmland areas were then extracted and overlaid with stable crop plots, with their intersection defined as high-quality representative regions. On this basis, representative sample points were created, resulting in 3624 rice points, 3082 maize points, and 1906 soybean points (Figure 6).

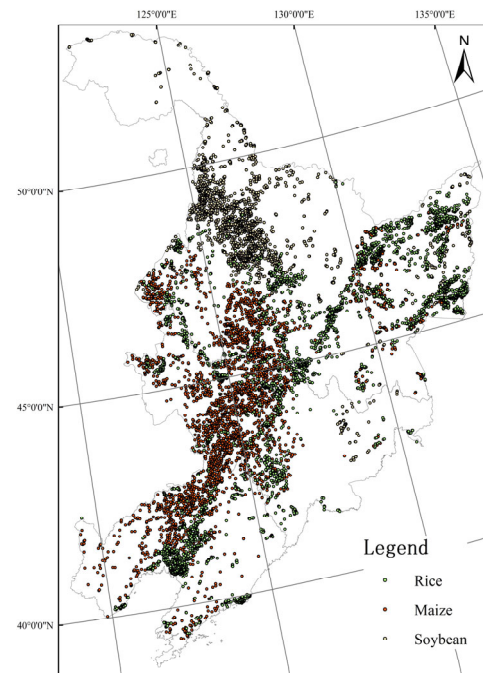


Figure 6. Spatial distribution of sample points for crop suitability assessment.

Using these high-quality sample points together with environmental variables, suitability distribution maps of rice, maize, and soybean were produced with the MaxEnt model. Each model was run 10 times, and all AUC values exceeded 0.80, indicating that the MaxEnt models constructed using environmental variables provided reliable and high-accuracy predictions (Figure S5). Natural conditions and human activities were intertwined in shaping the spatial distribution of crop suitability. Climate, topography, and soil formed the foundation for crop production, whereas human activities determined the feasibility and spatial distribution of cultivation. The importance and contributions of environmental variables were provided in the supplementary materials (Figures S6–S10).

Currently, the suitable area for rice cultivation is estimated at 311,869 km² (Figure 7). The highly suitable zones are primarily located in the Sanjiang Plain, including cities such as Hegang, Jiamusi, Shuangyashan, and Jixi, as well as the southern part of the Lower Liaohe Plain, such as Panjin, covering approximately 76,540 km², or 9.8% of the total area of the three northeastern provinces. The moderately suitable area accounts for 86,661 km² (11.1%), while marginally suitable zones comprise 19%, mainly distributed in the northern and southeastern parts of the Songnen Plain, including cities such as Qiqihar, Suihua, Siping, and Changchun. For maize, the total suitable area is approximately 329,056 km². The highly suitable area encompasses 102,349 km² (13.1%), mainly in the eastern and southern regions of the Songnen Plain and the Lower Liaohe Plain, including cities such as Suihua, Harbin, Changchun, Siping, Tieling, and Shenyang. The moderately suitable zone covers 111,117 km² (14.2%), and the marginally suitable area accounts for 14.8%, including cities such as Daqing, Baicheng, Chaoyang, and Dalian. Soybean had the widest suitable range,

with a total area of 376,692 km². The highly suitable area is concentrated at the junction of the Lesser Khingan Mountains and the Songnen Plain, particularly in southwestern Heihe, northeastern Qiqihar, and northeastern Suihua, covering 78,879 km² (10.1%). The moderately suitable zone extends across 105,525 km² (13.5%), while marginally suitable areas account for 24.6%, mainly distributed across the central and western parts of the Songnen Plain and the Sanjiang Plain, including southwestern Qiqihar, Daqing, Baicheng, Hegang, Jiamusi, Shuangyashan, and Jixi. Overall, compared to rice and maize, soybean shows greater climatic adaptability and is more suitable for cultivation in northern regions with limited heat and water resources. The highly suitable zones of the three crops overlap but cover a relatively small area, whereas the highly and moderately suitable zones of maize and soybean exhibit much larger overlaps. This indicates competition among the three crops, with the most pronounced competition occurring between maize and soybean.

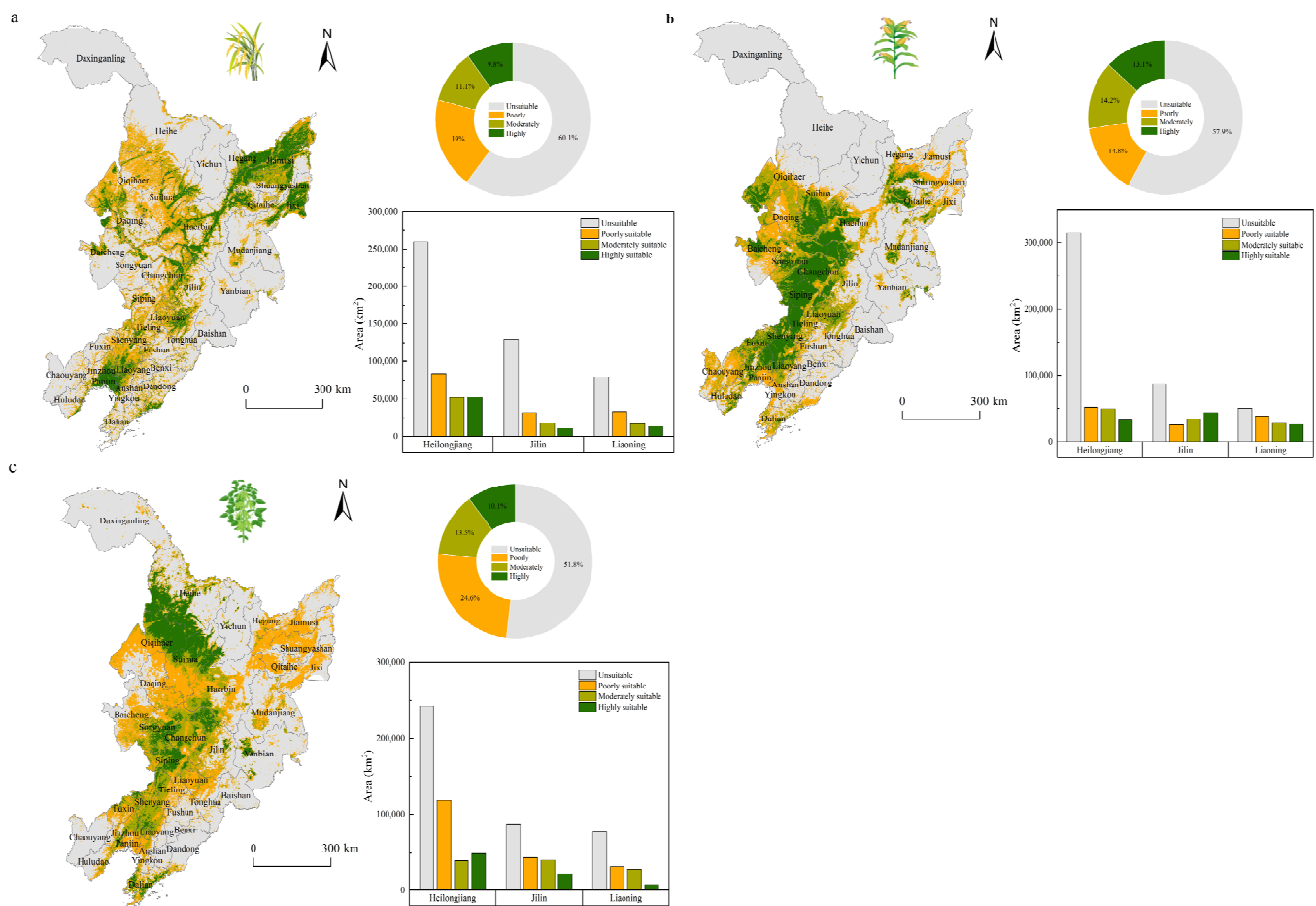


Figure 7. Potential distribution of suitability for rice (a), maize (b) and soybean (c).

3.5. Consistent Analysis of Crop Distribution and Suitability

To identify spatial opportunities for optimizing crop structure, the actual distribution of the three crops in 2023 was compared with their modeled suitability zones (Figure 8, Table 10). For rice planted in 2023, 57.39% of the area was highly suitable, 21.91% moderately suitable, 12.94% poorly suitable, and 7.77% unsuitable. With the exception of poor or unsuitable zones in central Qiqihar, Daqing, and northeastern Baicheng, most areas demonstrated good spatial consistency. For maize, 39.69% of the planted area was classified as highly suitable, 25.94% as moderately suitable, 17.57% as poorly suitable, and 16.81% as unsuitable. Marginal and unsuitable zones were mainly concentrated in northeastern Qiqihar, northern Suihua, Heihe, central and eastern Shuangyashan, north-

eastern Jiamusi, southern Baicheng, and Chaoyang. For soybean, 28.89% of the planting area was highly suitable, 16.23% moderately suitable, 36.11% poorly suitable, and 18.77% unsuitable. Marginal and unsuitable areas were concentrated in southwestern Qiqihar, Daqing, southwestern Suihua, Hegang, Jiamusi, Shuangyashan, Jixi (Heilongjiang), and Baicheng (Jilin). Overall, the current spatial pattern of actual cultivation does not align with the suitability distribution, particularly for maize and soybean, where mismatches are highly pronounced. For instance, while the northeastern regions of Heihe and Suihua are highly suitable for soybean cultivation, they are currently planted with maize. This indicates that soybean-suitable areas have been encroached upon by maize, highlighting the need to further optimize planting spatial allocation in the future.

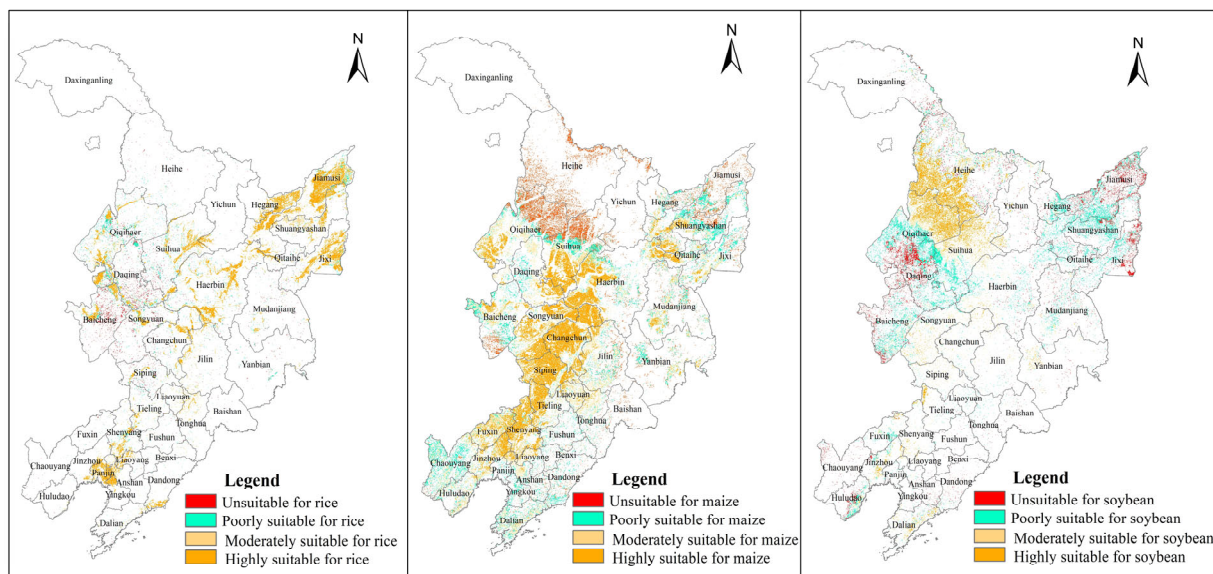


Figure 8. Alignment between actual planting areas and suitability for rice, maize and soybean.

Table 10. Consistency between crop cultivation change and suitability distribution.

Year	Crop Type	Unsuitable (km ²)	Poorly Suitable (km ²)	Moderate Suitable (km ²)	Highly Suitable (km ²)
2000→2023	Rice	4616	7685	13,014	34,094
	Maize	28,087	29,365	43,342	66,322
	Soybean	15,322	29,481	13,254	23,590
	Rice→Others	765	2478	3591	5012
	Maize→Others	5557	3829	5794	7006
	Soybean→Others	4214	7812	3179	8219
	Others→Rice	4423	6706	9617	19,783
	Others→Maize	26,406	26,714	36,526	48,471
	Others→Soybean	14,281	27,222	11,942	16,906

Additionally, a spatial overlay analysis was conducted to examine the consistency between areas that experienced crop type conversion from 2000 to 2023 and their corresponding suitability levels (Table 10). The results showed that during this period, a total of 11,846 km² of rice cultivation area was lost, of which 42.31% had been highly suitable for rice production. Meanwhile, 40,529 km² of land from other land uses or crops was converted to rice, with 48.81% of this area classified as highly suitable. For maize, 22,186 km² was lost, with 31.58% of the lost area being highly suitable. At the same time, 138,117 km² was converted to maize, of which 35.09% was highly suitable. Soybean experienced a loss of 23,424 km², with 35.09% of the lost area being highly suitable, while 70,351 km² was

gained from other land uses or crops, of which only 24.03% was highly suitable. Overall, the percentage of highly suitable areas in the lost cropland was consistently lower than that within the newly gained cropland for all three crops. This indicates that, despite crop type conversions over the study period, the total cultivation area for each crop remained adequately supported. Moreover, the proportion of highly suitable land among the newly gained areas was higher than that among the lost areas for both rice and maize, suggesting that high-quality planting areas were relatively well preserved or enhanced. In contrast, soybean showed the opposite pattern, with a lower proportion of highly suitable land in the newly converted areas compared to the lost ones, highlighting the need to prevent further loss of prime land suitable for soybean cultivation.

4. Discussion

4.1. Comparison with Published Crop Distribution Data Products

Classification performance was assessed by comparison with five existing datasets for validation. We calculated the total planting areas of the three crops for 36 prefecture-level cities based on published data products and compared them with our results. The area calculation was conducted on the GEE platform using the regional statistics function, which aggregates the pixel areas of each crop within the administrative boundaries of each prefecture-level city. To eliminate the effects of spatial resolution differences, all published crop distribution maps were resampled to a 1 km resolution. First, the mapping results of three major crops in 2015 and 2020 by Xuan [20] were examined (Table 11). According to Xuan, the R^2 values in 2015, based on comparisons with statistical yearbook data, were 0.75 for rice, 0.96 for maize, and 0.77 for soybean. In contrast, the corresponding R^2 values from this study were slightly higher at 0.79, 0.97, and 0.82, respectively. In 2020, all three crops reported by Xuan achieved R^2 values exceeding 0.90, and similarly, R^2 values in this study were also above 0.90 for all three crops. Overall, the results from this study demonstrated a stronger correlation with official statistical records, whereas Xuan's dataset exhibited better internal consistency.

Table 11. Comparison of this study (b) with Xuan's crop mapping (a).

Year	Crops	(a) R^2	RMSE (km ²)	(b) R^2	RMSE (km ²)
2015	Rice	0.75	696.3	0.79	1469.8
	Maize	0.96	786.7	0.97	1147.1
	Soybean	0.77	586.4	0.82	1090.4
2020	Rice	0.99	215.4	1	714.2
	Maize	0.92	923.5	0.93	1518
	Soybean	0.91	903.6	0.97	527.1

The rice classification maps from this study were compared with those produced by Luo [18] and Shen [33] for the years 2005, 2010, and 2015 (Figure 9). The R^2 values for Luo ranged from 0.64 to 0.86, with RMSE values between 370.5 km² and 749.3 km². Shen achieved R^2 values ranging from 0.74 to 0.86, and RMSE values between 266.5 km² and 1579 km². Among the three datasets, Shen's maps showed the strongest correlation with official statistics, followed by the CTNP dataset, while Luo's maps exhibited the weakest correlation. However, in terms of internal consistency with statistical data, Luo's classification showed the highest agreement compared to the other datasets. Similarly, the maize classification results from this study were compared with those from Luo [18] and Peng [34] for the same years. Luo's R^2 values ranged from 0.70 to 0.87, with RMSE values between 1088.2 km² and 1530.1 km². Peng reported R^2 values between 0.69 and

0.91, and RMSE values ranging from 877.6 km² to 1685.0 km². Among the three datasets, the CTNP results demonstrated the highest correlation with official statistics, while Peng's maps provided the greatest consistency.

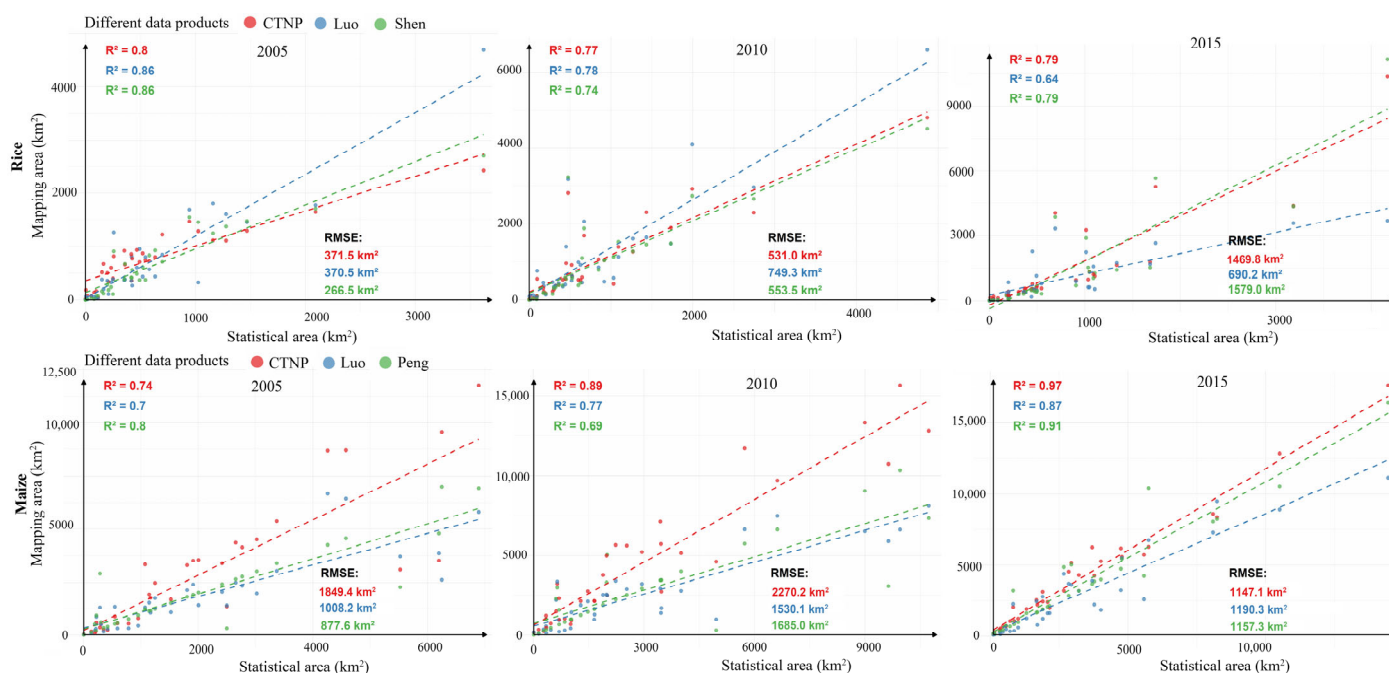


Figure 9. Comparison of different data products with data from statistical yearbooks.

Finally, the soybean classification maps generated in this study were compared with the soybean distribution maps produced by Li [23] for Heilongjiang Province (2000–2020) and by Liu [35] for the three northeastern provinces over the same period (Table 12). In Li's maps, the coefficient of determination (R^2) ranged from 0.76 to 0.96, with RMSE values between 911.8 km² and 2533.9 km². In Liu's maps, R^2 ranged from 0.84 to 0.99, and RMSE ranged from 358 km² to 1466.6 km². The differences in RMSE between our maps and those of Li and Liu ranged from 194.4 to 1016.5 km² and from 58.9 to 198 km², respectively. These results indicate a relatively high level of consistency between our soybean distribution maps and those of previous studies, thereby supporting the accuracy of our classification results.

Table 12. Comparison of this study (b) with Li's soybean mapping in Heilongjiang Province and Liu's soybean mapping in the three northeastern provinces (a).

Region	Year	(a) R^2	RMSE (km ²)	(b) R^2	RMSE (km ²)
Heilongjiang	2000	0.84	911.8	0.85	717.4
	2005	0.96	2533.9	0.84	1517.4
	2010	0.91	1295.9	0.87	2762.9
	2015	0.76	1021.4	0.9	1420.2
	2020	0.94	1548.4	0.97	821.9
Three Northeastern Provinces	2000	0.84	558.0	0.85	499.1
	2005	0.92	1001.7	0.9	870.7
	2010	0.9	1466.6	0.87	1594.8
	2015	0.91	892.4	0.82	1090.4
	2020	0.99	358	0.97	527.1

4.2. Drivers of Crop Pattern Transformation

In recent decades, significant changes have occurred in land use and crop planting structure across the three northeastern provinces [69,70]. Extensive regions of grasslands, forests, and wetlands have been converted into farmland, leading to a decline in vegetation cover and an increase in soil erosion. The expansion of cultivated land has heightened agricultural water demand, resulting in a substantial rise in groundwater extraction and further exacerbating water scarcity. Additionally, the prevalence of monoculture systems, such as continuous maize or soybean cropping, has hindered soil fertility restoration and contributed to the degradation of agroecosystems [71,72].

The transformation of cropping patterns in the three northeastern provinces has been driven by a combination of natural conditions, economic returns, policy incentives, and technological advancements [73–75]. Among these, natural factors provide the foundational conditions for agricultural restructuring. Global warming has enhanced thermal resources in the region, alleviating temperature constraints on crop growth and facilitating the northward expansion of cultivation areas [76]. Rice, being the most profitable crop, has particularly benefited from rising temperatures, which have created a more favorable environment for paddy fields in higher latitudes [77]. As rice expanded northward, it began to compete with maize, prompting the transitional zones of maize and soybean cultivation to shift further north [78].

Economic returns have been the dominant factor influencing crop structure adjustment. From the perspective of individual farmers, market prices play a decisive role in crop selection. Since China joined the World Trade Organization (WTO), low-priced imported soybeans have flooded the domestic market, severely undermining the competitiveness of local soybean production. In response, many farmers switched to maize, which offered higher profitability, leading to a rapid decline in soybean acreage and a substantial shift in the maize-soybean planting pattern. In 2016, the cancellation of the temporary maize reserve policy resulted in a sharp drop in maize prices. Consequently, farmers turned to rice cultivation in pursuit of higher returns, resulting in an expansion of rice at the expense of maize [79].

Policy interventions have also played a crucial role in shaping farmers' planting decisions [23]. The Chinese government has introduced a set of agricultural stimulus policies designed to promote food production. The abolition of agricultural taxes in 2004 facilitated a sustained increase in grain crop areas. To buffer against the dramatic drop in global agricultural commodity prices following the global financial crisis, and to avoid the adverse effects of falling grain prices on farmers' incomes, the government introduced a minimum purchase price policy for rice and temporary reserve policies for maize and soybean. During this period, rice acreage steadily increased, while soybean acreage declined after a brief recovery. Maize planting expanded significantly—even in areas unsuitable for its cultivation—revealing the distorting effects of price-support policies on the crop structure. Given the limited effectiveness of the temporary reserve mechanism, the Chinese government launched agricultural market-oriented reforms to reduce direct intervention. In 2015, the Ministry of Agriculture and Rural Affairs issued policy guidelines to adjust maize production in the so-called “Sickle Bend” region, recommending reductions in maize acreage and increased soybean cultivation. In 2016 and 2017, the temporary maize reserve and soybean target price policies were replaced by a producer subsidy combined with market-based procurement. A subsidy mechanism identical to that applied to maize was introduced for soybean producers, with the aim of improving the relative profitability of soybean through differentiated support [80]. These policy shifts have encouraged farmers to resume soybean planting, laying the foundation for a gradual recovery in its cultivation

area [81]. High-resolution, long-term crop mapping provides essential data support for evaluating the effectiveness of crop structure adjustment policies.

4.3. Improvements in the Prediction of Potential Crop Planting Zones Suitability

Many studies had already applied the MaxEnt model to predict the potential distribution of various crops or animals. As a supervised machine learning model, the MaxEnt approach was based on the principle of inferring a species' probability distribution function under known constraints while maximizing entropy. Consequently, the quality of the training dataset largely determined the accuracy and reliability of model predictions. However, previous research mainly relied on records from historical plant specimen databases or biodiversity information websites, which were often unevenly distributed, inconsistent in quality, and insufficiently representative [36]. For example, Estes et al. [82] used the MaxEnt model to create two maize suitability models: one based on national crop distribution sample points and another based only on high-productivity locations. The results showed that suitability outcomes derived from national crop distribution points had weak correlations with yield, whereas those based on high-productivity points exhibited stronger correlations with yield. To improve the quality and representativeness of sample points, we employed a multi-criteria suitability evaluation model, which required substantial prior knowledge, to filter sample points. This approach significantly enhanced the spatial representativeness of the dataset.

External validation based on independent data sources is essential for assessing the accuracy of suitability models. To evaluate the reliability of the crop suitability model developed in this study, two independent validation approaches were employed. First, the fundamental assumption of validation is that an effective suitability model should be capable of predicting observed agricultural productivity [79]. We collected the average yield per unit area data of rice, maize, and soybean in each prefecture-level city of the three northeastern provinces in 2023 from the Heilongjiang Statistical Yearbook, Jilin Statistical Yearbook, and Liaoning Statistical Yearbook. These datasets are entirely independent of the model construction process. The raster-based suitability results were aggregated by prefecture-level administrative boundaries to calculate the mean suitability index for each city. The mean suitability index was then matched with the corresponding average yield to construct the validation dataset. Using OriginPro 2024, we conducted linear regression analyses to explore the correlations between the two variables. The validation results (Figure 10) revealed significant positive correlations between crop suitability and actual yield. The Pearson correlation coefficients were $r = 0.628$ ($p < 0.05$), $r = 0.674$ ($p < 0.05$), and $r = 0.641$ ($p < 0.05$) for rice, maize, and soybean, respectively, with corresponding coefficients of determination $R^2 = 0.394$, $R^2 = 0.455$, and $R^2 = 0.410$. These findings indicate that the suitability indices explain approximately 39.4%, 45.5%, and 41.0% of the spatial variation in actual yield for rice, maize, and soybean, respectively, with statistically significant relationships. This result provides robust external evidence supporting the model's capability to capture the key biophysical factors determining land productivity, thereby confirming the validity of the proposed crop suitability model.

Second, we further validated the model in two representative verification areas. Hailun City and Youyi County in Heilongjiang Province were selected as typical regions representing the undulating hilly area and the flat plain area of Northeast China, respectively. Based on field survey data from these two regions, the evaluation results were compared with on-site sample points. Field observations focused on slope, surface gravel abundance, tillage layer thickness, and irrigation–drainage conditions to assess the model's accuracy. A total of 27 valid sample points were collected, covering most parts of Hailun City and Youyi County. The verification results (Tables S12 and S13) showed that the spatial differ-

ences in suitability levels were consistent with field observations, confirming that the crop suitability evaluation results in this study generally correspond well with actual conditions. Future work will further expand the spatial scope of validation to enhance the robustness of the model.

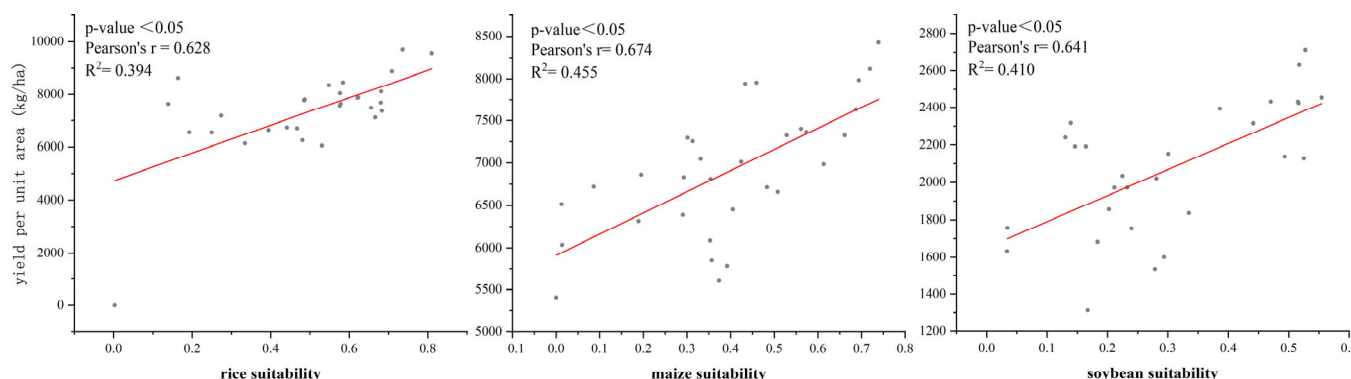


Figure 10. Correlation analysis between the average yield data and the average suitability data of the three crops at the prefecture level. The gray dots represent the average suitability and average grain yield level of each prefecture-level city.

4.4. Potential Suitability Conflicts Among Different Crops

A spatial overlay analysis of suitability maps for rice, maize, and soybean revealed that most areas exhibited no conflict in crop suitability. High-conflict zones were primarily located in the southeastern Songnen Plain and the Lower Liaohe Plain (Figure 11). In the potential conflict zones between rice and maize, the highest proportion of conflict was observed in Liaoyuan City, where conflicts affected over 80% of the administrative area. Other cities with conflict proportions exceeding 40% included Qitaihe, Shenyang, Jinzhou, Liaoyang, and Tieling, while the remaining cities had conflict proportions below 40%. For rice and soybean, the highest potential conflict ratio occurred in Shenyang, exceeding 75%. Other cities where potential conflict exceeded 40%, in descending order, were Tieling, Changchun, Jinzhou, Suihua, Jilin, Liaoyang, and Harbin. All other cities showed less than 40% conflict. In terms of maize and soybean, Siping exhibited the highest conflict proportion, reaching 91.58%. Other cities with conflict ratios exceeding 70% included Changchun, Songyuan, Liaoyuan, Shenyang, Panjin, and Tieling. Cities such as Harbin, Daqing, Jilin, Dalian, Jinzhou, Fuxin, and Liaoyang had conflict ratios above 40%. Remaining cities were below the 40% threshold. In the areas where all three crops were suitable and overlapped, Shenyang had the highest conflict proportion, affecting 84.49% of its administrative area. Other cities with over 40% conflict in descending order were Changchun, Siping, Songyuan, Liaoyuan, Jinzhou, Tieling, and Liaoyang. All other cities showed lower levels of potential conflict.

4.5. Policy Implications

In light of the aforementioned findings, several policy recommendations are proposed. First, regional land use planning should incorporate high-resolution cropland and crop suitability maps to ensure that highly suitable areas are prioritized in spatial development strategies. Second, the natural conditions that support crop suitability must be respected and managed—particularly key environmental factors such as soil thickness, sand content, and elevation. For instance, soil erosion should be controlled to preserve soil thickness [83], and appropriate soil amendments should be applied to improve the fertility of sandy soils [84]. Moreover, social and economic demands strongly influence farmers' planting decisions. In this regard, government intervention through incentive and compensation

policies could encourage farmers to grow crops in areas that are most suitable for their cultivation, thereby promoting a more rational spatial allocation of crop production.

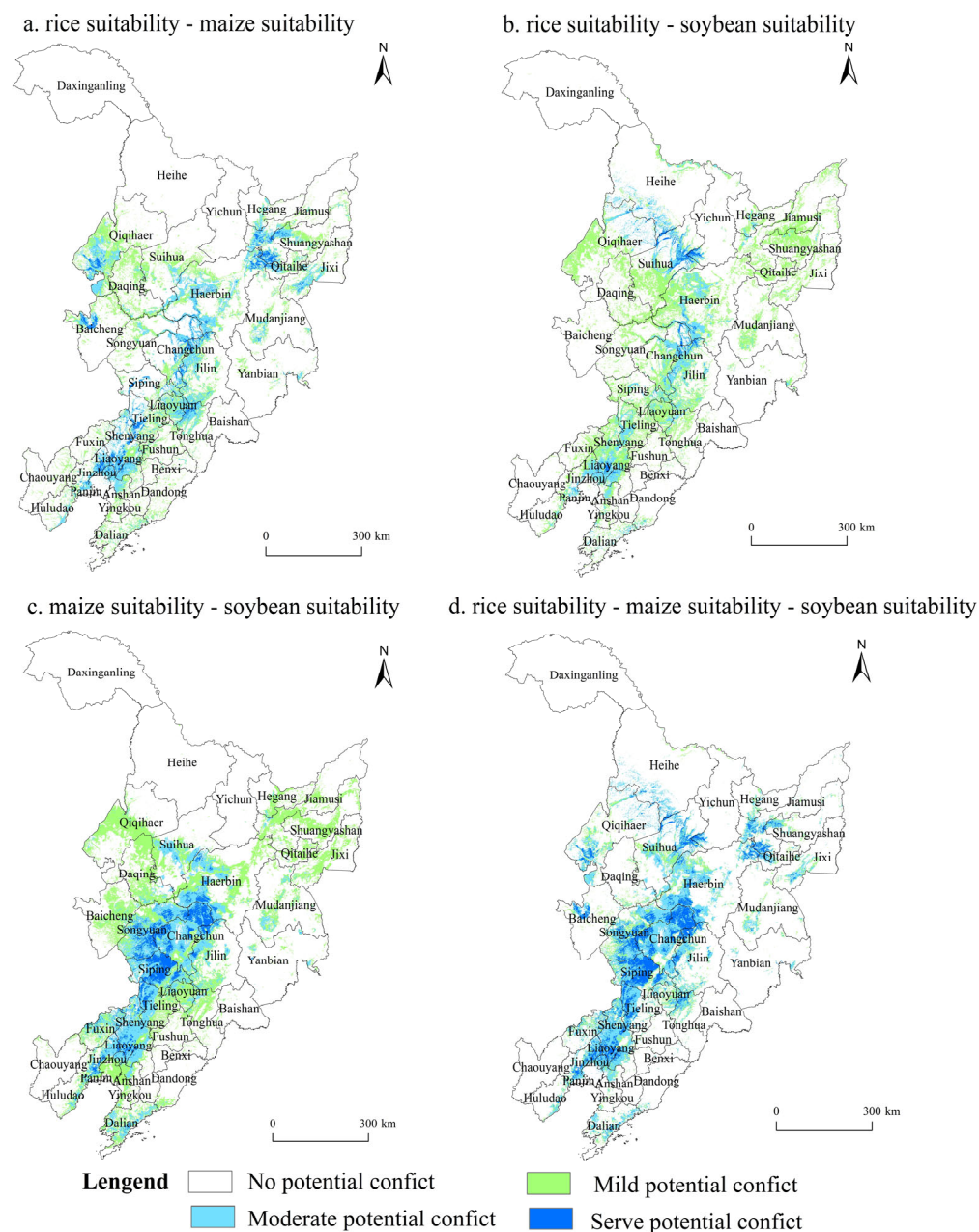


Figure 11. Potential conflicts in the suitability of three crops. (a): Potential conflicts in the suitability of rice and maize. (b): Potential conflicts in the suitability of rice and soybean; (c): Potential conflicts in the suitability of maize and soybean. (d): Potential conflicts in the suitability of rice, maize and soybean.

4.6. Limitations

Although this study offers valuable theoretical and practical insights for optimizing crop production structure, several limitations and uncertainties remain. First, the crop distribution maps relied on the staple crop cropland mask, and potential errors in this mask might have affected the mapping accuracy. To minimize this impact, multiple cropland and crop products were introduced for cross-validation and overlay during the mapping process, thereby improving the reliability of the staple crop cropland mask. Future research should explore binary classification methods for cropland and non-cropland to indepen-

dently generate mask datasets. Meanwhile, sample quality is a key factor influencing the performance of model-based extraction. Reliable, sufficient, and representative sample points are essential for high-accuracy crop mapping. In this study, we adopted the strategy of generating sample points for 2000–2020 from other data products instead of relying solely on field surveys. However, potential errors inherent in these data products may be further propagated into the Random Forest prediction model. Consequently, the quality of the samples may affect the classification accuracy across the two-decade period. Future studies should explore more diverse approaches for sample generation. Second, only the distributions of rice, maize, and soybean were identified and evaluated for suitability. This choice was directly related to the construction of the staple crop mask; when generating the maximum mask area, only pixels corresponding to these three crops were considered, while other crops and non-cropland pixels were excluded. As a result, vegetables, oil crops, and other agricultural types cultivated in Northeast China were not covered. This limitation indicated that the results mainly reflected the spatial distribution and suitability of staple crops, and could not be directly generalized to all crop types. Future studies could expand the mask coverage and incorporate additional crop samples to enable more comprehensive crop mapping and suitability assessments. Third, many parameters in the random forest classification model were set to default values, which might have influenced model performance and accuracy. Future work should explore hyperparameter optimization through approaches such as grid search or Bayesian optimization. Fourth, most of the environmental variables used in the MaxEnt model in this study are natural factors, while socioeconomic factors are relatively few. In future research, more socioeconomic data should be incorporated to better capture the impacts of human activities. Lastly, future climate change may continue to alter the suitability of crop cultivation. Consequently, it is necessary to further investigate crop suitability under future climate scenarios to alleviate the potential negative impacts of global warming on agricultural productivity.

5. Conclusions

Understanding the planting structure and agroecological suitability of primary grain crops in the three northeastern provinces is essential for evaluating agricultural stability, guiding production restructuring, and ensuring national food security. This study mapped the spatial distribution of rice, maize, and soybean at a 30 m resolution from 2000 to 2023, and assessed the suitability and key influencing factors for each crop. Overlay analysis of the suitability maps revealed potential competition zones, and a comparison between actual cropping areas and suitability maps was conducted to identify spatial mismatches. The main findings are as follows. (1) The overall mapping accuracy from 2000 to 2023 exceeded 85%, with an average Kappa coefficient of 0.857. The crop maps were consistent with official statistics and similar data products, demonstrating their reliability in capturing the spatiotemporal dynamics of rice, maize, and soybean cultivation in the region. (2) Over the past two decades, the planting area of the three major grain crops has expanded, with a notable trend of soybean being replaced by maize. (3) Soybean had the largest area of suitable land, followed by maize and then rice. Significant suitability conflicts were observed between maize and soybean. The Songnen Plain, Liaohe Plain, and Sanjiang Plain were identified as regions with high suitability for crop production. (4) Within the actual cropping areas, the proportion of moderate-to-high suitability land was 79.3% for rice, 65.63% for maize, and 45.12% for soybean. This indicates that there remains room for optimization of the cropping structure within existing farmland. Overall, this study provides valuable references for regulating farmland use, adjusting agricultural zoning, and formulating agricultural policies in the three northeastern provinces.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/agronomy15112587/s1>, Figure S1: Spatial distribution of sample points for crop mapping; Figure S2: Spearman correlation analysis of environmental variables; Figure S3: Spatial shift in crop types from 2000 to 2023; Figure S4: Spatial distribution of cultivated land use suitability (CLUS); Figure S5: ROC curves for rice, maize and soybean; Figure S6: Jackknife for regularised training gain in rice, maize and soybean; Figure S7: Percentage contribution of environmental variables for rice, maize and soybean; Figure S8: Response curves for environmental variables affecting rice; Figure S9: Response curves for environmental variables affecting maize; Figure S10: Response curves for environmental variables affecting soybean; Table S1: Numbers of images for each period; Table S2: Detailed settings of parameters; Table S3: Pairwise comparison matrix of first level indicators; Table S4: Pairwise comparison matrix of climate; Table S5: Pairwise comparison matrix of topography; Table S6: Pairwise comparison matrix of soil; Table S7: Pairwise comparison matrix of hydrology; Table S8: Pairwise comparison matrix of location conditions; Table S9: Weights of CLUS indicators; Table S10: Data source and main parameters; Table S11: Crop type transfer matrix from 2000 to 2023; Table S12: Field survey table of crop suitability in Youyi County; Table S13: Field survey table of crop suitability in Hailun City.

Author Contributions: X.W.: conceptualization; methodology; software; validation; formal analysis; investigation; data curation; visualization; writing—original draft; writing—review and editing. H.Z.: conceptualization; validation; resources; writing—review and editing; supervision; project administration; funding acquisition. G.Z.: methodology; data curation; writing—review and editing. X.Q.: methodology; data curation; writing—review and editing. C.C.: data curation; writing—review and editing. J.Q.: writing—review and editing. S.F.: writing—review and editing. T.W.: writing—review and editing. H.H.: data curation. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Key Technology R&D Program of Ministry of Science and Technology of China (Grant No. 2023YFD1500103) and National Natural Science Foundation of China (Grant No. 42171261).

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the corresponding authors on request.

Acknowledgments: We are grateful to the editors and reviewers for their insightful and useful comments and suggestions, which have significantly enhanced the quality and presentation of this research.

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

References

1. Yue, Y.; Su, P.; Gao, Y.; Zhang, P.; Wang, R.; Zhang, A.; Jiang, Q.; Ma, W.; Zhou, Y.; Wang, J.A. Mapping Global Crop Distribution. In *Atlas of Global Change Risk of Population and Economic Systems*; Springer Nature Singapore: Singapore, 2022; pp. 131–148.
2. Dong, J.; Xiao, X.; Menarguez, M.A.; Zhang, G.; Qin, Y.; Thau, D.; Biradar, C.; Moore, B. Mapping paddy rice planting area in northeastern Asia with Landsat 8 images, phenology-based algorithm and Google Earth Engine. *Remote Sens. Environ.* **2016**, *185*, 142–154. [\[CrossRef\]](#)
3. Qiu, B.; Yu, L.; Yang, P.; Wu, W.; Chen, J.; Zhu, X.; Duan, M. Mapping upland crop–rice cropping systems for targeted sustainable intensification in South China. *Crop J.* **2024**, *12*, 614–629. [\[CrossRef\]](#)
4. Li, L.; Friedl, M.A.; Xin, Q.; Gray, J.; Pan, Y.; Froking, S. Mapping Crop Cycles in China Using MODIS-EVI Time Series. *Remote Sens.* **2014**, *6*, 2473–2493.
5. Xie, W.; Zhu, A.; Ali, T.; Zhang, Z.; Chen, X.; Wu, F.; Huang, J.; Davis, K.F. Publisher Correction: Crop switching can enhance environmental sustainability and farmer incomes in China. *Nature* **2023**, *618*, E26. [\[CrossRef\]](#)
6. Li, X.; Wu, K.; Hao, S.; Yue, Z.; Ran, Z.; Ma, J. Mapping cropland suitability in China using optimized MaxEnt model. *Field Crops Res.* **2023**, *302*, 109064. [\[CrossRef\]](#)
7. Peter, P.; Svetlana, T.; Hansen, M.C.; Alexandra, T.; Viviana, Z.; Ahmad, K.; Xiao-Peng, S.; Amy, P.; Quan, S.; Jocelyn, C. Global maps of cropland extent and change show accelerated cropland expansion in the twenty-first century. *Nat. Food* **2022**, *3*, 19–28.

8. Zanaga, D.; Van De Kerchove, R.; Daems, D.; De Keersmaecker, W.; Brockmann, C.; Kirches, G.; Wevers, J.; Cartus, O.; Santoro, M.; Fritz, S. ESA WorldCover 10 m 2021 v200 (Version v200) [Data set]. Zenodo. 2022. Available online: <https://zenodo.org/records/7254221> (accessed on 1 September 2025).
9. You, L.; Sun, Z. Mapping global cropping system: Challenges, opportunities, and future perspectives. *Crop Environ.* **2022**, *1*, 68–73. [\[CrossRef\]](#)
10. Ditommaso, S.; Wang, S.; Strey, R.; Lobell, D. Mapping sugarcane globally at 10 m resolution using Global Ecosystem Dynamics Investigation (GEDI) and Sentinel-2. *Earth Syst. Sci. Data* **2024**, *16*, 4931–4947. [\[CrossRef\]](#)
11. Tian, H.; Huang, N.; Niu, Z.; Qin, Y.; Pei, J.; Wang, J. Mapping Winter Crops in China with Multi-Source Satellite Imagery and Phenology-Based Algorithm. *Remote Sens.* **2019**, *11*, 820. [\[CrossRef\]](#)
12. Zhang, H.; Li, Q.; Liu, J.; Shang, J.; Du, X.; Zhao, L.; Wang, N.; Dong, T. Crop classification and acreage estimation in North Korea using phenology features. *GIScience Remote Sens.* **2017**, *54*, 381–406. [\[CrossRef\]](#)
13. Zhong, L.; Hu, L.; Zhou, H. Deep learning based multi-temporal crop classification. *Remote Sens. Environ.* **2019**, *221*, 430–443. [\[CrossRef\]](#)
14. Belgii, M.; Csillik, O. Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis. *Remote Sens. Environ.* **2018**, *204*, 509–523. [\[CrossRef\]](#)
15. Vuolo, F.; Neuwirth, M.; Immitzer, M.; Atzberger, C.; Ng, W.-T. How much does multi-temporal Sentinel-2 data improve crop type classification? *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *72*, 122–130. [\[CrossRef\]](#)
16. Hu, X.; Zhang, S.; Li, L.; Huang, J.; Zhao, Z.; Liu, K.; Zhang, Z.; Yao, X. Major grain crop mapping in Northeast China using sample generation method and ensemble learning. *Eur. J. Agron.* **2025**, *169*, 127678. [\[CrossRef\]](#)
17. Wang, S.; Azzari, G.; Lobell, D.B. Crop type mapping without field-level labels: Random forest transfer and unsupervised clustering techniques. *Remote Sens. Environ.* **2019**, *222*, 303–317. [\[CrossRef\]](#)
18. Luo, Y.; Zhang, Z.; Li, Z.; Chen, Y.; Zhang, L.; Cao, J.; Tao, F. Identifying the spatiotemporal changes of annual harvesting areas for three staple crops in China by integrating multi-data sources. *Environ. Res. Lett.* **2020**, *15*, 074003. [\[CrossRef\]](#)
19. Zhai, Y.; Wang, N.; Zhang, L.; Hao, L.; Hao, C. Automatic Crop Classification in Northeastern China by Improved Nonlinear Dimensionality Reduction for Satellite Image Time Series. *Remote Sens.* **2020**, *12*, 2726. [\[CrossRef\]](#)
20. Xuan, F.; Dong, Y.; Li, J.; Li, X.; Su, W.; Huang, X.; Huang, J.; Xie, Z.; Li, Z.; Liu, H. Mapping crop type in Northeast China during 2013–2021 using automatic sampling and tile-based image classification. *Int. J. Appl. Earth Obs. Geoinf.* **2023**, *117*, 15. [\[CrossRef\]](#)
21. Liu, Y.; Wang, J. Revealing Annual Crop Type Distribution and Spatiotemporal Changes in Northeast China Based on Google Earth Engine. *Remote Sens.* **2022**, *14*, 4056. [\[CrossRef\]](#)
22. Li, X.; Yu, L.; Peng, D.; Gong, P. A large-scale, long time-series (1984–2020) of soybean mapping with phenological features: Heilongjiang Province as a test case. *Int. J. Remote Sens.* **2021**, *42*, 7332–7356. [\[CrossRef\]](#)
23. Yang, L.; Wang, L.; Huang, J.; Mansaray, L.R.; Mijiti, R. Monitoring policy-driven crop area adjustments in northeast China using Landsat-8 imagery. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *82*, 101892. [\[CrossRef\]](#)
24. Talukdar, S.; Naikoo, M.W.; Mallick, J.; Praveen, B.; Shahfahad; Sharma, P.; Islam, A.R.M.T.; Pal, S.; Rahman, A. Coupling geographic information system integrated fuzzy logic-analytical hierarchy process with global and machine learning based sensitivity analysis for agricultural suitability mapping. *Agric. Syst.* **2022**, *196*, 103343. [\[CrossRef\]](#)
25. Wang, Y.; Li, Y. Mapping the ratoon rice suitability region in China using random forest and recursive feature elimination modeling. *Field Crops Res.* **2023**, *301*, 109016. [\[CrossRef\]](#)
26. He, Q.J.; Zhou, G.S. Climate-associated distribution of summer maize in China from 1961 to 2010. *Agric. Ecosyst. Environ.* **2016**, *232*, 326–335. [\[CrossRef\]](#)
27. Zhao, J.; Wang, C.; Shi, X.; Bo, X.; Chu, Q. Modeling climatically suitable areas for soybean and their shifts across China. *Agric. Syst.* **2021**, *192*, 103205. [\[CrossRef\]](#)
28. Yang, J.; Hu, Q.; You, L.; Cai, Z.; Chen, Y.; Wei, H.; Xu, Z.; He, Z.; Yin, G.; Xu, B. Mapping the potential northern limits and promotion extent of ratoon rice in China. *Appl. Geogr.* **2023**, *150*, 102822. [\[CrossRef\]](#)
29. Xian, Y.; Liu, G.; Zhong, L. Will citrus geographical indications face different climate change challenges in China? *J. Clean. Prod.* **2022**, *356*, 131885. [\[CrossRef\]](#)
30. Liu, H.; Wan, W.; Zheng, M.; Li, J.; Liu, S.; Lv, W.; Zhou, Y.; Liu, Z. Study on climate suitability for maize and technical implementation strategies under conservation tillage in Northeast China. *Soil Tillage Res.* **2025**, *249*, 106473. [\[CrossRef\]](#)
31. Hong, Y.; Chen, Y.; Su, Y.; Dong, G.; Song, X.; Zhang, J.; Wang, H.; Sun, Z.; Guo, L.; Yang, J.; et al. A new framework for evaluating land suitability for Goji (*Lycium barbarum* L.) cultivation across China. *J. Environ. Manag.* **2025**, *392*, 126908. [\[CrossRef\]](#)
32. Pilevar, A.R.; Matinfar, H.R.; Sohrabi, A.; Sarmadian, F. Integrated fuzzy, AHP and GIS techniques for land suitability assessment in semi-arid regions for wheat and maize farming. *Ecol. Indic.* **2020**, *110*, 105887. [\[CrossRef\]](#)
33. Al-Taani, A.; Al-husban, Y.; Farhan, I. Land suitability evaluation for agricultural use using GIS and remote sensing techniques: The case study of Ma'an Governorate, Jordan. *Egypt. J. Remote Sens. Space Sci.* **2021**, *24*, 109–117. [\[CrossRef\]](#)

34. Khaki, B.D.; Chatrenour, M.; Navidi, M.N.; Soleimani, M.; Mirzaei, S.; Pignatti, S. Land Suitability Assessment Based on Feature-Level Fusion of Sentinel-1 and Sentinel-2 Imagery: A Case Study of the Honam Region of Iran. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2024**, *17*, 14777–14789. [\[CrossRef\]](#)
35. Franco, L.; Magalhães, M.R. Assessing the ecological suitability of land-use change. Lessons learned from a rural marginal area in southeast Portugal. *Land Use Policy* **2022**, *122*, 106381. [\[CrossRef\]](#)
36. Luo, K.; Lu, L.; Xie, Y.; Chen, F.; Yin, F.; Li, Q. Crop type mapping in the central part of the North China Plain using Sentinel-2 time series and machine learning. *Comput. Electron. Agric.* **2023**, *205*, 107577. [\[CrossRef\]](#)
37. Shen, R.; Peng, Q.; Li, X.; Chen, X.; Yuan, W. CCD-Rice: A long-term paddy rice distribution dataset in China at 30 m resolution. *Earth Syst. Sci. Data* **2025**, *17*, 2193–2216. [\[CrossRef\]](#)
38. Peng, Q.; Shen, R.; Li, X.; Ye, T.; Dong, J.; Fu, Y.; Yuan, W. A twenty-year dataset of high-resolution maize distribution in China. *Sci. Data* **2023**, *10*, 658. [\[CrossRef\]](#)
39. Liu, W.; Shu, L.; Fulu, T.; Jun, X.; Jinwei, D.; Jichong, H.; Qinghang, M.; Lichang, Y.; Hongyan, Z.; Juan, G. *30 m Annual Soybean Area Map in China from 2000 to 2022*; Zenodo: Geneva, Switzerland, 2024. [\[CrossRef\]](#)
40. Wardlow, B.D.; Egbert, S.L. Large-area crop mapping using time-series MODIS 250 m NDVI data: An assessment for the U.S. Central Great Plains. *Remote Sens. Environ.* **2008**, *112*, 1096–1116. [\[CrossRef\]](#)
41. Yang, N.; Liu, D.; Feng, Q.; Xiong, Q.; Zhang, L.; Ren, T.; Zhao, Y.; Zhu, D.; Huang, J. Large-Scale Crop Mapping Based on Machine Learning and Parallel Computation with Grids. *Remote Sens.* **2019**, *11*, 1500. [\[CrossRef\]](#)
42. Beckschäfer, P. Obtaining rubber plantation age information from very dense Landsat TM & ETM+ time series data and pixel-based image compositing. *Remote Sens. Environ.* **2017**, *196*, 89–100. [\[CrossRef\]](#)
43. You, N.; Dong, J.; Huang, J.; Du, G.; Zhang, G.; He, Y.; Yang, T.; Di, Y.; Xiao, X. The 10-m crop type maps in Northeast China during 2017–2019. *Sci. Data* **2021**, *8*, 41. [\[CrossRef\]](#)
44. Li, H.; Song, X.-P.; Hansen, M.C.; Becker-Reshef, I.; Adusei, B.; Pickering, J.; Wang, L.; Wang, L.; Lin, Z.; Zalles, V.; et al. Development of a 10-m resolution maize and soybean map over China: Matching satellite-based crop classification with sample-based area estimation. *Remote Sens. Environ.* **2023**, *294*, 113623. [\[CrossRef\]](#)
45. Kyere, I.; Astor, T.; Graß, R.; Wachendorf, M. Agricultural crop discrimination in a heterogeneous low-mountain range region based on multi-temporal and multi-sensor satellite data. *Comput. Electron. Agric.* **2020**, *179*, 105864. [\[CrossRef\]](#)
46. You, N.; Dong, J.; Li, J.; Huang, J.; Jin, Z. Rapid early-season maize mapping without crop labels. *Remote Sens. Environ.* **2023**, *290*, 113496. [\[CrossRef\]](#)
47. Yang, J.; Huang, X. The 30 m annual land cover dataset and its dynamics in China from 1990 to 2019. *Earth Syst. Sci. Data* **2021**, *13*, 3907–3925. [\[CrossRef\]](#)
48. Wang, X.; Gao, X.; Zhang, Y.; Fei, X.; Chen, Z.; Wang, J.; Zhang, Y.; Lu, X.; Zhao, H. Land-Cover Classification of Coastal Wetlands Using the RF Algorithm for Worldview-2 and Landsat 8 Images. *Remote Sens.* **2019**, *11*, 1927. [\[CrossRef\]](#)
49. Adugna, T.; Xu, W.; Fan, J. Comparison of Random Forest and Support Vector Machine Classifiers for Regional Land Cover Mapping Using Coarse Resolution FY-3C Images. *Remote Sens.* **2022**, *14*, 574. [\[CrossRef\]](#)
50. Wang, M.; Liu, Z.; Baig, M.H.A.; Wang, Y.; Li, Y.; Chen, Y. Mapping sugarcane in complex landscapes by integrating multi-temporal Sentinel-2 images and machine learning algorithms. *Land Use Policy* **2019**, *88*, 104190. [\[CrossRef\]](#)
51. Loosvelt, L.; Peters, J.; Skriver, H.; Lievens, H.; Van Coillie, F.M.; De Baets, B.; Verhoest, N.E. Random Forests as a tool for estimating uncertainty at pixel-level in SAR image classification. *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *19*, 173–184. [\[CrossRef\]](#)
52. Brinkhoff, J.; Vardanega, J.; Robson, A.J. Land Cover Classification of Nine Perennial Crops Using Sentinel-1 and -2 Data. *Remote Sens.* **2020**, *12*, 96. [\[CrossRef\]](#)
53. Ji, Y.; Gao, D.; Liu, Q.; Su, J.; Liu, Y.; Zhao, J.; Yang, Y.; Fu, Y.; Huang, G. An integrated framework for agricultural non-point source pollution control technology evaluation: Application of an improved multiple attribute decision making method. *Expert Syst. Appl.* **2023**, *229*, 120319. [\[CrossRef\]](#)
54. Jiang, S.; Meng, J.; Zhu, L. Spatial and temporal analyses of potential land use conflict under the constraints of water resources in the middle reaches of the Heihe River. *Land Use Policy* **2020**, *97*, 104773. [\[CrossRef\]](#)
55. Feng, Y.; Zhang, W.; Yu, J.; Zhuo, R. Optimization of land-use pattern based on suitability and trade-offs between land development and protection: A case study of the Hohhot-Baotou-Ordos (HBO) area in Inner Mongolia, China. *J. Clean. Prod.* **2024**, *466*, 142796. [\[CrossRef\]](#)
56. Qian, F.; Lal, R.; Wang, Q. Land evaluation and site assessment for the basic farmland protection in Lingyuan County, Northeast China. *J. Clean. Prod.* **2021**, *314*, 128097. [\[CrossRef\]](#)
57. Qian, F.K.; Wang, W.W.; Wang, Q.B.; Lal, R. Implementing land evaluation and site assessment (LESA system) in farmland protection: A case-study in northeastern China. *Land Degrad. Dev.* **2021**, *32*, 2437–2452. [\[CrossRef\]](#)
58. Shuqin, S.H.I.; Youqi, C.; Yanmin, Y.A.O.; Zhibin, L.I.; Yingbin, H.E. Assessing Natural Quality and Use Quality of Cultivated Land in Northeast China. *Resour. Sci.* **2008**, *30*, 378–384.

59. Song, G.; Zhang, H. Spatial reconstruction of farmland rotation and fallow in the typical black soil region of Northeast China. *J. Nat. Resour.* **2022**, *37*, 2231–2246. [CrossRef]
60. Jian-Ling, A.I.; You-Qi, C.; Yan-Min, Y. THE ASSESSMENT OF NORTHEAST RESERVE FARMING LAND RESOURCE BASED ON GIS. *Econ. Geogr.* **2007**, 637–639. Available online: https://kns.cnki.net/kcms2/article/abstract?v=uSrlZFhNZxLS1FeUsOIXy_GXTyWOePuhEB2R30JsHUBDwuUNRB0IA_FnU8KuB3roUSzU0ADB4PFHb7c2L3h_ulvuApIqEAJZ6-XxB-sApfTUgjl1n2mRR2EWCMMnFLDmRw_2ooVLKeq0Opzf0_xfYdSHu-f7Y4ubBqrcLzMdFS0SVZtMkwIWSO4ftFE95v-V&uniplatform=NZKPT&language=CHS (accessed on 1 September 2025).
61. Zolekar, R.B.; Bhagat, V.S. Multi-criteria land suitability analysis for agriculture in hilly zone: Remote sensing and GIS approach. *Comput. Electron. Agric.* **2015**, *118*, 300–321. [CrossRef]
62. Li, S.; Zhao, X.; Pu, J.; Miao, P.; Wang, Q.; Tan, K. Optimize and control territorial spatial functional areas to improve the ecological stability and total environment in karst areas of Southwest China. *Land Use Policy* **2021**, *100*, 104940. [CrossRef]
63. Han, J.; Cui, J.; Yang, W.; Xu, Y.; Qin, D.; Gao, F. Analysis of Soil Erosion Change and Driving Factors in Low Hilly Areas Based on InVEST Model. *Res. Soil Water Conserv.* **2022**, *29*, 32–39.
64. Williams, J.R. The erosion-productivity impact calculator (EPIC) model: A case history. *Philosophical Transactions of the Royal Society of London. Ser. B Biol. Sci.* **1990**, *329*, 421–428.
65. Fan, Y.; Li, H.; Miguez-Macho, G. Global patterns of groundwater table depth. *Science* **2013**, *339*, 940–943. [CrossRef]
66. Yang, S.L.; Wang, H.M.; Tong, J.P.; Bai, Y.; Alatalo, J.M.; Liu, G.; Fang, Z.; Zhang, F. Impacts of environment and human activity on grid-scale land cropping suitability and optimization of planting structure, measured based on the MaxEnt model. *Sci. Total Environ.* **2022**, *836*, 155356. [CrossRef]
67. Liu, F.; Wu, H.; Zhao, Y.; Li, D.; Yang, J.-L.; Song, X.; Shi, Z.; Zhu, A.X.; Zhang, G.-L. Mapping high resolution National Soil Information Grids of China. *Sci. Bull.* **2022**, *67*, 328–340. [CrossRef]
68. Li, Y.; Hong, H. Modelling flood susceptibility based on deep learning coupling with ensemble learning models. *J. Environ. Manag.* **2023**, *325*, 116450. [CrossRef]
69. Cao, D.; Feng, J.-Z.; Bai, L.-Y.; Xun, L.; Jing, H.-T.; Sun, J.-K.; Zhang, J.-H. Delineating the rice crop activities in Northeast China through regional parametric synthesis using satellite remote sensing time-series data from 2000 to 2015. *J. Integr. Agric.* **2021**, *20*, 424–437. [CrossRef]
70. Wang, S.; Xu, X.; Huang, L. Spatial and Temporal Variability of Soil Erosion in Northeast China from 2000 to 2020. *Remote Sens.* **2023**, *15*, 225. [CrossRef]
71. Yang, D.; Liu, Y.; Wang, J. Trade-offs and synergies of food-water-land benefits for crop rotation optimization in Northeast China. *Agric. Ecosyst. Environ.* **2025**, *379*, 109377. [CrossRef]
72. Zhang, Y.; Wang, S.; Wang, H.; Ning, F.; Zhang, Y.; Dong, Z.; Wen, P.; Wang, R.; Wang, X.; Li, J. The effects of rotating conservation tillage with conventional tillage on soil properties and grain yields in winter wheat-spring maize rotations. *Agric. For. Meteorol.* **2018**, *263*, 107–117. [CrossRef]
73. Du, G.; Han, L.; Yao, L.; Faye, B. Spatiotemporal Dynamics and Evolution of Grain Cropping Patterns in Northeast China: Insights from Remote Sensing and Spatial Overlay Analysis. *Agriculture* **2024**, *14*, 1443. [CrossRef]
74. Liu, X.Q.; Liu, Y.S.; Liu, Z.J.; Chen, Z.F. Impacts of climatic warming on cropping system borders of China and potential adaptation strategies for regional agriculture development. *Sci. Total Environ.* **2021**, *755*, 13. [CrossRef]
75. Yang, Z.; Chu, L.; Wang, C.; Pan, Y.; Su, W.; Qin, Y.; Cai, C. What drives the spatial heterogeneity of cropping patterns in the Northeast China: The natural environment, the agricultural economy, or policy? *Sci. Total Environ.* **2023**, *905*, 167810. [CrossRef] [PubMed]
76. Liu, Z.; Yang, P.; Tang, H.; Wu, W.; Zhang, L.; Yu, Q.; Li, Z. Shifts in the extent and location of rice cropping areas match the climate change pattern in China during 1980–2010. *Reg. Environ. Change* **2015**, *15*, 919–929. [CrossRef]
77. Liu, Y.; Liu, X.; Liu, Z. Effects of climate change on paddy expansion and potential adaption strategies for sustainable agriculture development across Northeast China. *Appl. Geogr.* **2022**, *141*, 102667. [CrossRef]
78. Du, G.; Ma, M.; Zhang, R.; Liu, Z. Change of maize-soybean cropping patterns and its link with climate warming in Northeast China between 2000 and 2020. *Resour. Sci.* **2024**, *46*, 2251–2262.
79. Ming, G.; Yin-Sheng, X.I. Research on the Coordination Mechanism of Grain Policy in Northeast China. *J. South China Agric. Univ.* **2018**, *17*, 20–28.
80. Wei, W.X.; Si, W. Does the Reform of Soybean Subsidy Policy Realize the Expansion of Soybean Production? An Empirical Analysis Based on 124 Prefecture-level Cities in Soybean Main Producing Areas. *Chin. Rural. Econ.* **2021**, 44–65. Available online: <https://link.cnki.net/urlu/11.1262.F20211228.1117.004> (accessed on 1 September 2025).
81. Di, Y.; You, N.; Dong, J.; Liao, X.; Song, K.; Fu, P. Recent soybean subsidy policy did not revitalize but stabilize the soybean planting areas in Northeast China. *Eur. J. Agron.* **2023**, *147*, 126841. [CrossRef]

82. Estes, L.; Bradley, B.; Beukes, H.; Hole, D.; Lau, M.; Oppenheimer, M.; Schulze, R.; Tadross, M.; Turner, W. Comparing mechanistic and empirical model projections of crop suitability and productivity: Implications for ecological forecasting. *Glob. Ecol. Biogeogr.* **2013**, *22*, 1007–1018. [[CrossRef](#)]
83. Gu, Z.; Xie, Y.; Gao, Y.; Ren, X.; Cheng, C.; Wang, S. Quantitative assessment of soil productivity and predicted impacts of water erosion in the black soil region of northeastern China. *Sci. Total Environ.* **2018**, *637*, 706–716. [[CrossRef](#)]
84. Elbagory, M.; Shaker, E.M.; El-Nahrawy, S.; Omara, A.E.-D.; Khalifa, T.H. The Concurrent Application of Phosphogypsum and Modified Biochar as Soil Amendments Influence Sandy Soil Quality and Wheat Productivity. *Plants* **2024**, *13*, 1492. [[CrossRef](#)] [[PubMed](#)]

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