Spatialization of Actual Grain Crop Yield Coupled with Cultivation Systems and Multiple Factors: From Survey Data to Grid

Jingxin Li, Hongqi Zhang and Erqi Xu

Key Laboratory of Land Surface Pattern and Simulation, Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China; lijx.18b@igsnrr.ac.cn (J.L.); zhanghq@igsnrr.ac.cn (H.Z.)
2 University of Chinese Academy of Sciences, Beijing 100049, China
*
Correspondence: xueq@igsnrr.ac.cn

Received: 26 April 2020; Accepted: 7 May 2020; Published: 11 May 2020

Abstract: The spatialization of actual grain crop yield helps to understand the spatial heterogeneity of yield and support for the precise farming and targeted farmland management. However, inadequate consideration and quantification of anthropogenic factors affecting the estimation of actual yield distribution easily cause uncertainties in recent researches. Here, we developed a new grain crop yield spatialization (GCYS) model in order to downscale the yield from county to grid scale. The GCYS model is composed of four modules: (a) cultivated land Net Primary Productivity (NPP) calculation module, (b) comprehensive agricultural system construction module, (c) key factors establishment module, and (d) integration and validation module. Its novelty is to realize the actual grain crop yield spatialization from county scale to grid scale by quantifying and spatializing the comprehensive agricultural system when considering the diversity of cultivated structure and field management factors. Taking Guizhou and Guangxi Karst Mountains Region as a study-area, the GCYS model is developed and tested. The determination coefficients of regression analysis between agricultural survey data and spatialization results of paddy rice yield calculated by our model reach 0.72 and 0.76 in 2000 and 2015, respectively (p < 0.01). The results visualize the spatial pattern of actual grain crop yield at the grid scale, which show a gradually decreasing trend from southeast to northwest. With an increase in potential yield and improvement of field management technologies, the actual average yield of grain crops per unit increased from 4745.10 kg/ha of 2000 to 5592.89 kg/ha of 2015. Especially in high-yield zones in southeast area, mechanized cultivation became the dominated factor, rather than chemical fertilizer application. This demonstrates that our model can provide a reference for agricultural resource utilization and policy-making.

Keywords: actual yield spatialization; grain crop; cultivation system; key factors; accuracy assessment

1. Introduction

The actual grain crop yield is affected to different degrees by the physical environment and social factors [1–4], such as irregular topography [1,3], irrigation [5,6], soil nutrients [1,7], and the application of pesticides and chemical fertilizers [8,9]. These factors can generate grain crop yield differences, even at a fine scale [10–14]. Understanding the spatial distribution of the grain crop yield of specific regions plays a crucial role in policy-making and market management [15,16]. However, agricultural survey data at a county level mask the spatial heterogeneity of these factors. Spatialization downscales the grain crop yield statistics from administrative units to grid by coupling the physical and anthropogenic factors. The downscaled grid data can contribute to understanding the spatial heterogeneity of agricultural resource distribution. Identifying the major obstacles or opportunities at
A grid scale can help in maximizing the productivity of cultivated land, optimizing the allocation of various physical resources, and decreasing production losses that are caused by inappropriate farming practices or land use patterns [15,17].

Some achievements have been made for single grain crop productivity spatialization, including paddy rice [15,18], maize [12,18], wheat [6,19,20], and barley [21], and multiple grain crops at global [17,18,22,23], national [24–26], and regional [5,15,27,28] scales because of the advantages of raster data. The estimation of the grain crop yield can be divided into five productivity levels, including photosynthetic productivity, light and temperature productivity, climatic productivity, soil productivity and land productivity [16]. The estimations of previous studies have been determined by the main yield restriction factors according to the particular crop growth characteristics or local environmental conditions [16,27,28]. For example, the AEZ (Agro-Ecological Zoning) model is representative of a step-by-step revision method and it is widely used in various scenarios and at a range of research scales [18,20,29,30]. Indicators, such as NPP (net primary productivity; NPP is defined as the net carbon retained by plants that is the difference between photosynthesis and respiratory action) [15,31], GPP (gross primary productivity; the total carbon retained by plants through photosynthesis) [32,33], or NDVI (normalized difference vegetation index, \( \text{NDVI} = \frac{(NIR - R)}{(NIR + R)} \), NIR and R are reflectivity at near-infrared and infrared bands) [19,34] are also used to evaluate the crop growth states or estimate the productivity of the crops and they are gradually becoming mainstream methods. Based on comprehensive datasets, crop growth models, such as the WOFOST (World Food Studies) [35,36], can also be used to precisely predict the per unit yield in micro-scale regions [37].

However, limitations still exist in recent crop spatialization researches. First, most studies focus on potential productivity [27,38,39], because of the inadequacy of actual yield research, which is a reflection of reality. Few studies on actual productivity are mostly determined by physical indicators [16,27,40,41]. The impact of social or anthropogenic factors on productivity and its spatial distribution, resulting in differences between actual productivity and actual yield, remains unclear [1,2,9]. Furthermore, an accurate single crop yield spatialization is not sufficient for estimating the productivity of a study area because administrative divisions include diverse planting structures, but not just one grain crop [42]. A comprehensive cultivation system model needs to be developed, rather than a conventional single crop yield model, to realize spatialization [17], as that will be more valuable for government policy-making. Lastly, most spatialization studies using crop productivity estimates based on models or special indicators do not integrate the survey data into result evaluation. The similarities and differences between the spatialization results and agricultural survey data [15,17] require thorough analysis. This research not only spatialized the actual grain crop yield under comprehensive crop structures, but also integrated agricultural survey data in the spatialization process to assess and calibrate the data.

In hilly areas, the spatialization of grain crop yield is more challenging, because there is a high spatial heterogeneity of physical and anthropogenic factors, while these could have positive or negative effects on grain crop outputs. In the Guizhou and Guangxi Karst Mountain Region, complex topography, shortage of surface water, and severe rocky desertification restrict agricultural development [43]. In recent years, advancing urbanization, continuous population growth, and conversion of limited farmlands have severely impacted food supplies [44]. Therefore, the accurate identification of the distribution of actual grain crop yield is more critical [44], and it will contribute to the sustainable agricultural production and local food security. The aim at our research is to develop a novel grain crop yield spatialization (GCYS) model, which included: (1) a comprehensive agricultural cultivation system that is suitable for local farming conditions, (2) the spatialization of the core physical and anthropogenic factors correlated to actual grain crop yield, and (3) an accuracy assessment and calibration process at the grid and county scale. By integrating the comprehensive cultivation system and anthropogenic factors, the spatialization results will correspond better to actual grain crop yield, which can be tested through the correlation analysis between the agricultural survey data and spatialized results. Mapping the grain crop yield will provide more credible information for the regional policy-making and resource utilization planning.
2. Materials and Methods

2.1. Study Area

The Guizhou and Guangxi karst mountain region is a typical karst region in Southwest China, located between 22°8′54″ N–28°12′27″ N and 104°18′27″ E–110°20′40″ E. It covers 211,400 km² within the Guizhou and Guangxi provinces (Figure 1). The area lies in the subtropical monsoon humid climate zone with a mean annual temperature of 19 °C and a mean annual precipitation of 1350 mm. The terrain gradually descends from northwest to southeast with an elevation range of 0 to 2848 m. The topography in this region is mainly composed of the positive terrain that is represented by Fengcong and the negative terrain represented by depressions. The sloping land is common in karst regions, because the terrain is dominated by positive terrain with great amplitude relief, and the expansion of construction land occupied relatively flat and limited negative terrain regions [43]. The agricultural survey data show that the per capita grain holding in 2015 is about 300 kg in this region, which is only 66% of the national level and barely met basic food requirements.

![Figure 1. Location and terrain of the Guizhou and Guangxi karst mountains region.](image)

2.2. Data Description

2.2.1. Data Collection and Preprocessing

The administrative divisions at the county level are adjusted between 2000 and 2015; therefore, the names and boundaries of counties in the study area needed to be unified before analysis. We consolidated data within inconsistent counties in order to match the administrative division adjustment. Ninety counties in the Guizhou and Guangxi karst mountain region are used for analysis.

We collected the main agricultural production survey data for grain crops (i.e., rice, wheat, maize, soybean, and potato) from the Ministry of Agriculture and Rural Affairs of the People’s Republic of China, including cultivated land area, irrigation area, use of chemical fertilizer, sown area, and crop yield for the five main crops in 90 counties. We selected these data at the county level because there are no available data at the land parcel unit or a more detailed scale. These agricultural survey data, except...
the five grain crop yields, constituted the comprehensive agricultural system, which is an important module of our spatialization model. The light and temperature potential productivity and climatic potential productivity are collected from the national standard of the Regulation for Gradation on Agricultural Land Quality in each county. The potential productivity of the main grain crops and calculations from the appointed meteorological observation station data of each county are revised based on the topographic characteristics of the station to the mean elevation and slope of the cultivated land of the county, according to the research conducted by Wang et al. [45]. The revised dataset reflects the maximum actual crop productivity of each grain crop at the county level.

2.2.2. Key Factors of Yield Spatialization

Ten parameters in five categories are applied to analyze the spatial differences of actual crop yields at different stages based on previous studies of the Guizhou and Guangxi karst mountain region (Table 1). The land use types, categorized into six first level and twenty-five second level classes, are interpreted from Landsat images. The average interpretation accuracy according to field surveys and random sample checks are 92.9% and 97.6%, respectively. Paddy field and dry farmland distributions are extracted from all land use categories and they are used for actual yield spatialization, and the area of each cultivated land plot is calculated. The spatial scale of the land use classification and the topographic data is 30 m. The spatial scale of climate data, soil data, and economic data is 1 km. However, as the spatial heterogeneity of these data are embodied at a larger scale than land use and topography in practice, a grid scale with 1 km resolution is unable to reflect the fine spatial pattern. The resolution of 250 m is set as the appropriate grid scale when considering the complex topography of the study region and the need to maintain the details on land use as far as possible. Climatic data are used to calculate the NPP of cultivated land; topographic and soil data are used to modify the available soil water capacity. The area of each cultivated land plot and the topographic data are the key parameters for mechanized cultivation, and economic data are used to spatialize chemical fertilizer application. These factors are the key to determining the spatialization results of the actual grain crop yield.

### Table 1. Description of the spatial datasets and attributes.

| Factors                  | Spatial Scale        | Temporal Scale       | Data Source                                                                 |
|--------------------------|----------------------|----------------------|-----------------------------------------------------------------------------|
| Land use data            |                      |                      |                                                                             |
| Spatial pattern of cultivated land | 30 m                | 2000 and 2015        | Data Center for Resources and Environmental Sciences, Chinese Academy of Sciences. |
| Area of each cultivated land plot | -                   | 2000 and 2015        |                                                                             |
| Climatic data            |                      |                      |                                                                             |
| Mean annual temperature  | 1 km × 1 km          | 2000 and 2015        | Data Center for Resources and Environmental Sciences, Chinese Academy of Sciences. |
| Mean annual precipitation| 1 km × 1 km          | 2000 and 2015        |                                                                             |
| Topographic data         |                      |                      |                                                                             |
| Elevation                | 30 m × 30 m          | 2009                 | International Scientific Data Service Platform, Computer Network Information Center, Chinese Academy of Sciences. Derived from DEM in Arcgis. |
| Slope                    | 30 m × 30 m          | 2009                 |                                                                             |
| Soil data                |                      |                      |                                                                             |
| Ratio of sandy soil      | 1 km × 1 km          | 1995                 | Data Center for Resources and Environmental Sciences, Chinese Academy of Sciences. |
| Ratio of silt soil       | 1 km × 1 km          | 1995                 |                                                                             |
| Ratio of clay soil       | 1 km × 1 km          | 1995                 |                                                                             |
| Soil organic matter     | 1 km × 1 km          | 1995                 |                                                                             |
| Economic data            |                      |                      |                                                                             |
| GDP                      | 1 km × 1 km          | 2000 and 2015        | Data Center for Resources and Environmental Sciences, Chinese Academy of Sciences. |

2.3. The Grain Crop Yield Spatialization Model

NPP is defined as the net carbon retained by plants [46], which can be calculated and converted into potential productivity based on growth characteristic parameters; the converted yield is the light and temperature potential productivity. However, for a rainfed scenario, the yield should be
revised by the available soil water capacity (AWC) to climate productivity. The potential productivity must be modified by anthropogenic field management practices to realize the transformation from potential productivity to actual crop yield, because factors, such as mechanization and chemical fertilizer application, have generated a gradually increasing impact on grain crop yield. Finally, the spatialization results should be calibrated at the grid and county levels, respectively, to improve the accuracy of actual grain crop yield at different scales.

Based on the framework mentioned above, the GCYS model consists of four modules (Figure 2): (a) cultivated land NPP calculation, (b) comprehensive agricultural system construction, (c) key factors establishing, and (d) integration and validation. The GCYS model can realize the actual crop yield spatial allocation, which couples a comprehensive cultivation system with the spatial heterogeneity that is generated by human management factors.

![Diagram](image-url)

**Figure 2.** The framework of the grain crop yield spatialization (GCYS) model. NPP—Net Primary Productivity,  $R^2$—determination coefficients, c—calculated results of determination coefficients, p—probability level of obtaining the statistic, given that the null hypothesis is true.

In the cultivated land NPP calculation module, NPP is estimated based on the Thornthwaite Memorial model, which is an improved version of the Miami model [47]. Once the NPP is obtained, it can be converted to potential yield by using the parameters of crop growth characteristics. However, comprehensive crop productivity rather than the specific single crop yield is the target of our research. Therefore, we improve the input parameters of the Thornthwaite Memorial model to fit the cultivation system of each county in this module.

In the agricultural system construction module, a comprehensive cultivation system conforming to reality is constructed, instead of only estimating a single crop yield distribution, in order to generate the comprehensive crop productivity parameters that are needed to be applied in the first module. Therefore, a standard cultivation system, a total cultivation structure, a multiple cropping index, and crop growth categories in paddy field and dry farmland for each county are coupled in our agricultural system.

In the key factors establishing module, physical and anthropogenic factors, which can influence the grain crop yields, are integrated to achieve a better response of actual grain crop yield spatialization. Several improvements are made in this module, which is based on the SAORES (Spatial Assessment
and Optimization Tool for Regional Ecosystem Services) model [48]. When considering the complex topographic and lithological characteristics of the karst mountain region, the AWC is revised in dry farmland, according to the slope difference. In addition, human management factors, including mechanization and chemical fertilizer application, are spatialized to further revise the potential yield close to reality.

In the model integration and validation module, a complete model integrating the former three modules is established. Subsequently, the grid value is calibrated by the data collected from the national standard and revised by terrain characteristics in the previous step to achieve higher reliability at the grid level. Finally, the reliability of the model is evaluated at the county level.

2.3.1. Cultivated Land NPP Calculation Module

First, the NPP of cultivated land is calculated by the Thornthwaite Memorial model [49], which is driven by precipitation and temperature data:

$$\text{NPP}_{ck} = 3000 \times \left(1 - e^{\frac{-0.0009695 \times \sqrt{\frac{1}{1 + 1.05 \times P_{ck}}} - 20}{3000 + 25 T_{ck} + 0.05 T_{ck}^3}}\right)$$

where $\text{NPP}_{ck}$ indicates the annual potential NPP in grid $k$ of county $c$ (g·C·m$^{-2}$·a$^{-1}$), $V_{ck}$ represents the actual annual evapotranspiration (mm) in grid $k$ of county $c$, $L_{ck}$ represents the annual average evapotranspiration (mm) in grid $k$ of county $c$, $P_{ck}$ is the annual average precipitation (mm) in grid $k$ of county $c$, and $T_{ck}$ is the annual average temperature ($^\circ$C) in grid $k$ of county $c$.

Subsequently, the mixed potential grain crop yield (i.e., rice, wheat, maize, soybean, and potato) is calculated according to the improved Equation (2) [50–52], based on the specific crop growth characteristics. However, the actual crop cultivation distribution is difficult to obtain in different years, and some variable factors influence the grain crop category in one site, like market price fluctuation, and the comprehensive crop productivity rather than specific grain crop yield in one grid is the target of spatialization. The input parameters were improved from Equation (3) to Equation (5), which combines the proportional area of various crops in paddy field or dry farmland, according to the standard cultivation system. The calculations are as follows:

$$\text{NPP}_{ck} = \frac{\text{PY}_{ck} \times (1 - M_{cc} \times m_c \times 0.45 \text{gC/g})}{H_{lc} \times f_{AG,c}} = \text{PY}_{ck} \times t$$

where $\text{NPP}_{ck}$ indicates the mixed potential crop yield in grid $k$ of county $c$ (kg/ha), which means that $\text{PY}_{ck}$ is the light and temperature potential productivity of paddy field or dry farmland; $M_{cc}$, $H_{lc}$, and $f_{AG,c}$ are the mixed water content (%) for different cultivation systems, harvest index (the ratio of the crop harvest and economic production), and the ratio of the aboveground biomass and the total biomass in county $c$, respectively; for the parameters referring to crop $n$, if grid $k$ in county $c$ is located in a paddy field, then $n = i$, namely, $i$ is the grain crop category that is planted in the paddy field of county $c$; or else, $n = j$, and $j$ represented the crops cultivated in the dry farmland of county $c$; $m_c$ is a calibration coefficient for multiple indices: if cropping in county $c$ take place once a year, then $m_c = 0.5$, otherwise, if it occurs more than once a year, its value is 1.0; 0.45 gC/g is the carbon conversion coefficient; and, $t$ means the transform coefficient of converting NPP to potential grain crop yield. The parameters in Equation (2) can be calculated, as follows:

$$M_{cc} = \sum_n \left(\text{Area}_{P_{nc}} \times M_{cn}\right)$$

$$H_{lc} = \sum_n \left(\text{Area}_{P_{nc}} \times H_{ln}\right)$$

where $\text{Area}_{P_{nc}}$ is the proportional area of various crops in paddy field or dry farmland.
\[ f_{AG,c} = \sum_n (\text{Area}_{Pnc} \times f_{AG,n}) \] (5)

where \( \text{Area}_{Pnc} \) represents the sown area proportion of the crop \( n \) planted in the paddy field or dry farmland of county \( c \); \( M_{rc} \), \( HI_{rn} \), and \( f_{AG,n} \) are the water content (\%), harvest index, and ratio of the aboveground biomass in county \( c \) (Table 2). Therefore, the NPP is the comprehensive net primary productivity weighted by sown area proportion of the mixed five grain crops, i.e., rice, wheat, maize, soybean, and potato. Once the \( \text{Area}_{Pnc} \) become known quantity, the mixed potential grain crop yield in grid \( k \) of county \( c \) can be calculated.

Table 2. The coefficients for estimation of cultivated land Net Primary Productivity (NPP) [50–52].

| Index | \( M_{rc} \) | \( HI_{rn} \) | \( f_{AG_n} \) |
|-------|-------------|-------------|-------------|
| Rice  | 0.1375      | 0.45        | 0.80        |
| Wheat | 0.1250      | 0.37        | 0.83        |
| Maize | 0.1550      | 0.49        | 0.85        |
| Soybean | 0.1250   | 0.25        | 0.87        |
| Potato | 0.1330      | 0.50        | 0.20        |

2.3.2. Agricultural System Construction Module

The purpose of the agricultural system construction module is to obtain the proportional area for each grain crop in the irrigation (paddy field) or rainfed (dry farmland) scenarios, and then light and temperature productivity could be obtained using Equation (2). The paddy field area is not included in the agricultural survey data; therefore, the irrigation area is used for estimation instead. Based on the survey data that were collected from the Regulation for the Gradation on Agriculture Land Quality, the multiple cropping index showed that crops in nine counties are harvested once a year, and three counties had triple cropping. Thus, most of the ninety counties harvest at least twice a year. For the counties where the multiple cropping index is larger than one, the cropping pattern of rice combined with another grain crop is universal.

The sown area in paddy yield is equal to the sown area of rice in the counties that did not have multiple cropping of rice with other grain crops. In contrast, in the counties practicing multiple cropping of rice with rainfed crops, no more than two crops are planted. For example, the standard double or triple cropping pattern is “a-b” or “a-a-b” (“-” indicates the planting order of multiple crops in a year; “a” and “b” stands for the different grain crops), and the sown area estimation of these grain crops in paddy field is calculated, as follows:

\[
\text{SurplusArea}_{ac} = \text{IrrigationArea}_c \times MCI_c - \text{SownArea}_{ac} \] (6)

where \( \text{IrrigationArea}_c \) is the irrigation area of county \( c \) (ha), \( MCI_c \) is the multiple cropping index in county \( c \), \( \text{SownArea}_{ac} \) is the sown area of the crop \( a \) in county \( c \) (ha), and \( \text{SurplusArea}_{ac} \) is the surplus sown area with irrigation in county \( c \) after planting crop \( a \). Subsequently, an allocation on crop area is made while using Equation (7):

\[
\text{SownArea}_{Pc} = \begin{cases} 
\text{SownArea}_{ac} + \text{SownArea}_{bc}, & \text{if } \text{SurplusArea}_{ac} \geq \text{SownArea}_{bc}; \\
\text{IrrigationArea}_c \times MCI_c, & \text{if } \text{SurplusArea}_{ac} < \text{SownArea}_{bc}.
\end{cases} \] (7)

\[
\text{SownArea}_{Pbc} = \text{SownArea}_{Pc} - \text{SownArea}_{ac} \] (8)

\[
\text{SownArea}_{Dbc} = \text{SownArea}_{bc} - \text{SownArea}_{Pbc} \] (9)

where \( \text{SownArea}_{bc} \) is the sown area of the crop \( b \) (ha) in county \( c \), \( \text{SownArea}_{Pc} \) is the total area of grain crops cultivating in the paddy field (ha) in county \( c \), and \( \text{SownArea}_{Pbc} \) and \( \text{SownArea}_{Dbc} \) are areas (ha) of crop \( b \) cultivating in the paddy field and dry farmland in county \( c \), respectively. If the \( \text{SurplusArea}_{ac} \) is not less than the sown area of crop \( b \) in county \( c \), then the total area of crop \( a \) and \( b \) is the sown area.
of grain crops in paddy fields in county $c$; if the $SurplusArea_c$ is less than the sown area of crop $b$, the $SurplusArea_c$ becomes the sown area of crop $b$ in the paddy field. After carrying out the calculation of Equation (7), the area of crop $b$ in the paddy field and dry farmland in county $c$ can be obtained based on Equations (8) and (9).

After the sown area of grain crops in the paddy field or dry farmland is determined, the total sown area in the two scenarios is calculated, as follows:

$$Area_{PNC} = \begin{cases} \frac{SownArea_{PPi}}{\sum SownArea_{PPi}} & \text{if crop } n \text{ located in paddy field}, n = i \\ \frac{SownArea_{PDj}}{\sum SownArea_{PDj}} & \text{if crop } n \text{ located in dry farmland}, n = j \end{cases}$$

(10)

where $SownArea_{PPi}$ and $SownArea_{PDj}$ are the sown areas of crop $i$ that are planted in paddy field (ha) and crop $j$ planted in dry farmland (ha) of county $c$, respectively; $Area_{PPi}$ or $Area_{PDj}$ are the proportional areas (%) of crop $i$ or $j$ in paddy field or dry farmland in county $c$, respectively. Substituting the results of $Area_{PNC}$ into Equations (3)–(5), the values of $Mc_c$, $HI_c$, and $f_{AG,c}$ could be obtained, and then calculated the mixed potential grain crop yield (i.e., rice, wheat, maize, soybean, and potato) in grid $k$ of county $c$ ($PY_{ck}$) based on Equation (2).

2.3.3. Key Factors Establishing Module

In the key factors establishing module, three physical or anthropogenic factors, including the AWC revised by topography, mechanization, and chemical fertilizer application, are calculated in order to improve the spatial heterogeneity accuracy of the grain crop yield at the grid level.

**Soil Available Water Capacity**

The AWC is the volume of soil water that should be available to plants (unit: $10^{-2} \text{ cm}^3 \text{ cm}^{-3}$) [53]. The AWC is related to the ratio of sand, silt, clay soil, and soil organic matter based on previous experiments in the study area [54]. The revision of AWC is only calculated in the dry farmland. When considering the sum proportion of sand, silt, clay soil is 1, the quantitative relations are as follows:

$$AWC_k = 53.778 + 0.592 \times (SAN\%)_k + 0.004 \times (SAN\%)_k^2 + 0.669 \times (SIL\%)_k + 0.001 \times (SIL\%)_k^2 + 0.014 \times (SAN\% \times SIL\%)_k - 2.688 \times (C\%)_k + 0.501 \times (C\%)_k^2$$

(11)

where $(SAN\%)_k$, $(SIL\%)_k$, and $(C\%)_k$ are the ratios of the sand, silt, and organic matter of grid $k$ in county $c$, respectively, and the $AWC_k$ is the available soil water capacity of grid $k$.

Equation (13) could be used to estimate the mean $AWC_k$ value of most soil categories. However, the complex terrains in the karst mountain region impacted $AWC_k$. An experimental research in Puding County, which is located in our study area, showed a linear decrease in the capacity until the slope is up to $40^\circ$ [53]. Therefore, the $AWC_k$ of farmland where the slope is larger than $6^\circ$ is revised according to the slope land grade of the technical regulation of the Third Nationwide Land and Resources Survey, as follows:

$$CAWC_k = AWC_k \times \frac{\frac{\text{Max}(slope_k) - \text{Slope}_k}{\text{Max}(slope_k)}}$$

(12)

$$NCAWC_k = \frac{\text{CAWC}_k}{\text{Max}(\text{CAWC}_k)}$$

(13)

where $Slope_k$ is the slope of grid $k$, when the slope degree is larger than $6^\circ$, $\text{Max}(\text{Slope}_k)$ is the maximum slope value of whole study area, $S_k$ is the calibration coefficient of $AWC_k$, and $CAWC_k$ is the actual soil water capacity calibrated by slope degree. If the slope degree of grid $k$ is less than $6^\circ$, the calibration coefficient $S_k$ is 1; $\text{Min}(\text{CAWC}_k)$ and $\text{Max}(\text{CAWC}_k)$ are the minimum and maximum value of $CAWC_k$ in the whole study area, respectively, and $NCAWC_k$ is the normalization result of $CAWC_k$. For the paddy field, the value of $NCAWC_k$ is 1.
Mechanized Cultivation

The farmland patch area and grain crop categories are the two main factors impacting the mechanized cultivation rate. Previous researches have shown that the length of farmland patches that are available for mechanized cultivation are usually larger than 200 m \[55\], and the slope is no more than 6° \[56\]. Restricted by the spatial resolution of the raster data in the present study, we select the area of one grid cell as the lower limit for mechanized cultivation.

The mechanized cultivation process included three steps: plowing, seeding, and harvesting. The contribution of the mechanized cultivation rate to the increase in production of grain crops is calculated, as follows:

\[
YIR_c = \sum_n (M_{nc} \times Y_{nc} \times P_{nc}) 
\]

\[
M_{nc} = m_{nc1} \times m_{nc2} \times m_{nc3} \tag{15}
\]

\[
P_{nc} = \frac{SownArea_{nc}}{\sum_n SownArea_{nc}} \tag{16}
\]

where \(YIR_c\) is the share of increasing yield due to the introduction of mechanized cultivation (kg/ha) in county \(c\); \(m_{nc1}, m_{nc2},\) and \(m_{nc3}\) are the mechanized cultivation rates of crop \(n\) in the cropping process of plowing, seeding, and harvesting (\%) in county \(c\); \(P_{nc}\) is the sown area proportion of crop \(n\) in county \(c\); and, these values are collected from the China Statistical Yearbook and China Agricultural Machinery Yearbook; \(SownArea_{nc}\) is the sown area of crop \(n\) (ha) in county \(c\); \(Area_{nc}\) is the proportional area of crop \(n\) of the total sown area (%) in county \(c\); \(M_{nc}\) is the comprehensive mechanization rate of crop \(n\) (%) in county \(c\); and, \(Y_{nc}\) is the contribution rate of mechanization to the increase in grain crop yield in county \(c\). \(Y_{nc}\) is set as 9.21% in 2000 and 8.73% in 2015, according to Zhou et al. \[57\], respectively. As for the potato and soybean, the mechanization rate of the latter two steps is extremely low, and the sown areas of both are relatively lower than the other three grain crops in most counties, so we regarded the \(M_{nc}\) values of potato and soybean as 0.

Chemical Fertilizer Application

Li. et al. \[58\] showed that the amount of chemical fertilizer application is significantly correlated with the farmers’ household income, which is determined by using questionnaires. However, this indicator is not included in the agricultural survey data of all counties each year. Therefore, the chemical fertilizer application is quantified, as follows:

\[
P_{Vc} = GDP_{ck} \times \frac{PV_{c}}{GDP_{c}} \tag{17}
\]

\[
C_{fck} = \frac{PV_{ck}}{\sum_k PV_{ck}} \times C_{fc} \tag{18}
\]

where \(PV_{c}\) and \(GDP_{c}\) are the planting industry production value (10^4 yuan) and the gross regional domestic product (10^4 yuan) of county \(c\) (\(c = 1, 2, \ldots, 90\)), respectively; \(GDP_{ck}\) is the GDP value of grid \(k\) in county \(c\) (10^4 yuan); therefore, \(PV_{ck}\) is the planting industry production value of grid \(k\) in county \(c\) (10^4 yuan), and the \(\sum_k PV_{ck}\) is the total planting industry production value of county \(c\) (10^4 yuan); \(C_{fc}\) is the net application of chemical fertilizer of county \(c\) (kg/ha), which is collected from the available agricultural survey data; and, \(C_{fck}\) is the spatialization value of the net application of chemical fertilizer in grid \(k\) of county \(c\) (kg/ha). The upper limit is set as 1000 kg/ha when considering the amount of chemical fertilizer applied to grain crops. The mean crop yield increase from the fertilizer application is calculated, as follows \[28\]:

\[
\Delta Y_{ck} = 2.2875 \times C_{fck} \times \left(6.58 \times e^{-0.047991 \times C_{fck}} + 1.08\right) \tag{19}
\]
where \( \Delta Y_{ck} \) is the mean increasing yield in grid \( k \) of county \( c \) (kg/ha). After revising the potential yield by the three factors, respectively, the maximum reduction or increase in the yield, namely, the dominated factor, can be identified among the three factors. Then the area proportion of each domination factors can be calculated, which help to understand the sphere of each factor influence.

2.3.4. Integration and Validation Module

The former three modules are integrated to determine the spatialization of crop yield in this module, and then the reliability of the model is verified; and finally, the grid values are calibrated to be lower than the revised crop productivity collected through the national standard, while the spatialized results are revised being the same as agricultural survey data at the county level.

The three factors developed in the second module are combined further to describe the spatial heterogeneity of the actual grain crop yield, while the revised and normalized available soil water capacity, and anthropogenic management factors, including mechanized cultivation and chemical fertilizer application, are considered both in dry farmland and paddy field, as follows:

\[
Y_{ck} = PY_{ck} \times \text{NCAWC}_k \times (1 + YIR_c) + \Delta Y_{ck}
\]  

(20)

where \( Y_{ck} \) is the final actual crop yield in grid \( k \) of county \( c \) after revised by the three key factors (kg/ha).

Our results indicating the comprehensive productivity of grain crop yield is difficult to be accurately assessed. However, it is feasible to assess the accuracy of the GCYS model through the available crop distribution extent. We collected the rice distribution within the pure pixels of paddy field based on our comprehensive agricultural system module, and calculated the average per unit area of paddy rice of each county in ArcGIS 10.2. The average per unit yield of paddy field can be calculated according to the yield and sown area of rice. The significant correlation test is used to validate the reliability of our GCYS model. First, the data normal distribution is examined before the correlation analysis. We set the null hypothesis that there is no correlation between the two groups. Otherwise, the alternative hypothesis, that a significant correlation existed, is accepted. The p-value is set as 0.01. The linear regression is applied to analyze the function relationship and the difference between the two groups.

From the micro-scale perspective, the crop yield in a grid should be less than the light and temperature productivity of paddy field or the climate productivity of dry farmland, which are calculated through observational data and topographic calibration. The calibration at the grid level is as follows:

\[
Y_{ck} \leq \begin{cases} 
\text{Max}(RY_{ci}), & \text{if } k \text{ located in paddy field} \\
\text{Max}(RY_{cj}), & \text{if } k \text{ located in dry farmland}
\end{cases}
\]  

(21)

where if the grid \( k \) is located in paddy field of county \( c \), \( \text{Max}(RY_{ci}) \) indicates the maximum per unit area yield of crop \( i \) of county \( c \); or else, \( \text{Max}(RY_{cj}) \) is the maximum per unit area yield of crop \( j \) of county \( c \).

2.4. The Application of the GCYS Model

Grain crops are a basic and important source of nutritional calories for humans, and grain crop production is one of the most important indicators of food security and healthy diet [59]. Because all the grain crop yields are estimated in our model, the spatialization result is used to calculate the calories of local grain crops by using the following formula [59]:

\[
Q_{ck} = 100 \times Y_{ck} \times EP_c \times C_c
\]  

(22)

\[
EP_c = \sum_n (EP_n \times \text{AreaP}_{nc})
\]  

(23)

\[
C_c = \sum_n (C_n \times \text{AreaP}_{nc})
\]  

(24)
where \( Q_{ck} \) is the total calorie in grid \( k \) of county \( c \) (kcal/ha); \( EP_n \) is the edible proportion of crop \( n \) (%); \( EP_c \) is the comprehensive edible proportion of grain crops of county \( c \) (%); \( C_n \) is the calorie of crop \( n \) (kcal/100 g); and, \( C_c \) is the comprehensive calorie of grain crops of county \( c \) (kcal/100 g).

3. Results

3.1. Spatial Pattern Analysis of the Potential Productivity of Grain Crops

3.1.1. Spatial Pattern of Net Primary Productivity

NPP calculation is the first step for estimating the potential yield of grain crops. The mean mixed NPP value of cultivated land is 1629.92 in 2000 and 1765.00 g-C·m\(^{-2} \cdot a^{-1} \) in 2015, respectively (Figure 3a,b). In 2000, the NPP distribution values are higher in the central region, and lower in the north and south regions of the study area. Relatively higher elevation and complex karst terrain are distributed in the northern region, and temperature became the limiting factor for NPP. The lowest annual average temperature of the whole area is 9.16 °C in the north, which is significantly lower than the maximum average temperature in the whole study area of 23.13 °C. Precipitation is the limiting factor in the southern region, with the lowest precipitation of 1029 mm in this region, which is far below the 1849 mm in the central area. In 2015, precipitation is a dominant factor in NPP spatial pattern changes. The higher precipitation area moves slightly to the south, and the NPP shows a similar spatial distribution.

3.1.2. Spatialization of Potential Productivity

Generally, the spatial pattern of the mixed transform coefficient decreased slightly by 0.02 from 2000 to 2015. The mixed coefficient of paddy field decreased by 0.0052. However, the change of that in dry farmland is −0.0311 due to the diversity of cropping categories of dry farmland. The most significant change is found in Dafang County, in the north of the study area, where there is a maximum reduction of 0.31 due to remarkable changes in the cropping structure. In 2000, the sown area of potato is just 1.05%, but the ratio sharply increased to 40.35% in 2015, while the sown areas of the other four crops decreased. The mixed transform coefficient of potato from NPP to potential productivity is the lowest among the five grain crops based on Table 3; thus, the mixed transform coefficient of Dafang County shows the greatest reduction (Figure 4). A similar situation is observed across the northeast and northwest.
Table 3. The coefficients for calculation of grain crop calories.

| Index       | EP (%) | C (kcal/100 g) |
|-------------|--------|----------------|
| Rice        | 100    | 1405           |
| Wheat       | 78     | 1532           |
| Maize       | 100    | 1527           |
| Soybean     | 100    | 1412           |
| Potato      | 100    | 77             |

Data collected from agricultural survey data of Ministry of Agriculture and Rural Affairs of PRC.

Figure 4. (a) The mixed transform coefficient $t$ in 2000; (b) The mixed transform coefficient $t$ in 2015.

The mixed potential grain yield (i.e., rice, wheat, maize, soybean, and potato) is calculated (Figure 5). The lower mixed potential grain crop yields are mainly distributed in the northeast and southwest regions in 2000, and the higher mixed potential grain crop yields are mainly located in the central part of the study area. In contrast, the relatively low grain crop yield region expanded in the northern area in 2015, while the regions with the higher values showed a southeast shift. The minimum and the maximum per unit area potential yield both increased from 2000 to 2015, and the mean value of the whole study area increased from 14,600.57 kg/ha in 2000 to 15,521.49 kg/ha in 2015.

Figure 5. (a) The mixed potential grain yield derived from NPP in 2000; and, (b) The mixed potential grain yield derived from NPP in 2015.
3.2. Key Factors Analysis of Actual Yield

Figure 6 shows the areas of different dominant factors influencing the grain crop yield. In general, the AWC is the limiting factor affecting the largest area. The proportion of dry farmland increased from 2000 to 2015. Although field management technologies, including the mechanization level and chemical fertilizer application, improved rapidly, the AWC increased slightly from 59.08% to 59.29% (Table 4).

![Figure 6](image_url)

**Figure 6.** (a) Dominant key factors distribution in 2000; (b) Dominant key factors distribution in 2015 (AWC: Available soil water capacity; Mechan: Mechanization; CFA: Chemical fertilizer application).

**Table 4.** The reduction or increasing effect of grain crop yield and the area proportion of dominated factors in 2000 and 2015.

| Dominated Factors                | 2000                           | 2015                           |
|----------------------------------|--------------------------------|--------------------------------|
|                                  | Reduction/Increasing of Yield(kg/ha) | Dominated Area Proportion of Total Cultivated Land | Reduction/Increasing of Yield(kg/ha) | Dominated Area Proportion of Total Cultivated Land |
| Available soil water capacity    | 0–15334.63(−)                  | 59.08%                         | 0–17947.08(−)                 | 59.29%                             |
| Mechanization                    | 0–194.90(+)                    | 0.23%                          | 0–990.36(+)                   | 17.47%                             |
| Chemical fertilizer application  | 0–2470.50(+)                   | 40.69%                         | 48.86–2470.50(+)              | 23.24%                             |

The impact of the proportional area of mechanization is the smallest among the three key factors, and it had little effect with the area proportion of 0.23% in 2000. However, its dominated area proportion increased to 17.47% in 2015. In the southern region, where the terrain is relatively flat and multiple crops are harvested at least twice a year, the advantages of mechanization gradually become the dominant factor, thereby playing a more important role of increasing production in 2015.

The chemical fertilizer application proved to be the most efficient field management technology for improving crop yield during the study period, even though the effect of mechanization grew rapidly. From 2000 to 2015, the dominant area of chemical fertilizer application decreased by 17.45%, but the mean per unit area of production increased from 332.27 kg/ha to 347.52 kg/ha.
3.3. Accuracy Assessment and Spatial Pattern Analysis of Actual Grain Crop Yield

3.3.1. Accuracy Assessment and Calibration at Two Scales

Figure 7 shows the regression lines of the per unit area yield in county level between the spatialized results of GCYS model and agricultural survey data. The determination coefficients of paddy field in 2000 and 2015 are 0.72 and 0.76, respectively, and the model results are significantly correlated with the agricultural survey data \( (p < 0.01) \) and have a higher fitting degree when compared with previous studies [16,28,35]. The slopes of the two curves are 0.95 and 1.23, respectively, indicating that the spatialization results are close to reality, although a slight underestimation and overestimation occurred for paddy rice yield estimation per unit area at the county level in 2000 or 2015.

![Regression analysis of per unit area yield in county level between spatialized results calculated by GCYS model and agricultural survey data of (a) per unit area yield of paddy rice in 2000; (b) per unit area yield of paddy rice in 2015.](image)

The significant correlation of spatialization results and agricultural survey data verified the reliability of our model. Based on this, the spatialization results were revised with the agricultural survey data at the county level, and then calibrate any grid productivity values higher than potential productivity. However, the effect of field management technologies, like chemical fertilizer application in lower potential yield regions, will be larger than that in other regions. Although we have taken these factors into account, the uncertainty of the simulated results still exists, especially the points that are located in lower value area of the plot while deviating from the straight line.

3.3.2. Spatialization Results of Actual Grain Crop Yield

The mean mixed grain crop yield (i.e., rice, wheat, maize, soybean, and potato) of the total cultivated lands is 4745.10 kg/ha in 2000 and 5592.89 kg/ha in 2015, respectively. In 2000, the mean mixed grain crop yield of paddy field and dry farmland are 8495.48 kg/ha and 2894.82 kg/ha, respectively. Additionally, the yield changed to 8473.25 kg/ha and 3676.96 kg/ha, respectively, by 2015. In general, although the local spatial pattern showed variations at different degrees, the gradient of the actual yield in the whole study area increased from northwest to southeast, consistent with elevation gradient (Figure 8). In the north, the actual grain crop yield decreased in 2015. The AWC, the dominant factor of actual grain crop yield, is relatively steady, but changes in local temperature and precipitation led to a decrease in NPP. Meanwhile, the change in cropping structure decreased the transform coefficient from NPP to potential yield (Figure 4), thus limiting potential productivity and the actual yield. In the southern region, the productivity of low actual yield obviously improved, and the high value region is more concentrated and moved southward. The increase in potential yield and its southward shift was attributed to the rapid mechanized cultivation development in these regions.
Huanjiang County is used as an example to analyze the spatial pattern change in actual grain crop yield. Huanjiang County is a typical karst plateau-peak-cluster-depression transition region that is located in the middle of the study area, and the elevation gradually decreased from northeast to southwest. The average per unit area yield of mixed actual grain crop is 2395.84 kg/ha in 2000. Except for farmlands with a low productivity (less than 1000 kg/ha) scattered on the steep mountains in the north, the lower values mainly surrounded the mountains in the east (Figure 9). Relatively high productivities in the paddy field are found in the southern karst depression areas. In 2015, the average per unit area yield of mixed actual grain crop increased to 4505.41 kg/ha (an increase of 88.51%). The proportional area of increasing yields dominated by mechanization increased from 0% in 2000 to 13.07% in 2015 in the southern flat region. However, this substantial improvement does not alter the spatial pattern of the whole county. Especially in the northern and eastern mountains, AWC determined by terrain and soil attributes is the dominant limiting factor for actual grain crop yield, although the per unit area yield value is improved by other factors.

3.4. The Spatio-Temporal Change in Actual Grain Crop Yield

The mixed average grain crop yield (i.e., rice, wheat, maize, soybean, and potato) increased by 847.79 kg/ha from 2000 to 2015. We set (0,4000), (4000,8000), and (8000,15,000) kg/ha as the ranges for low, medium, and high productivity fields. It shows that the per unit area yield improved from 2000 to
2015 (Figure 10), mainly due to the decreasing percentage of low productivity fields and the increasing percentage of medium productivity fields. The per unit area yield were reclassified as seven groups with an interval of 2000 kg/ha. The percentage of the lowest productivity fields registers the largest decline, decreasing from 11.84% in 2000 to 0.71% in 2015, and the percentage of medium productivity fields (4000–6000 kg/ha) increases from 14.81% to 25.20%, showing the largest increase.

![Figure 10. Interval distribution of the mixed actual grain yield, i.e., rice, wheat, maize, soybean, and potato, in 2000 and 2015.](image)

From 2000 to 2015, the cultivated land area decreased, and a total of 119,112.50 ha of farmland is lost. Of the total land lost, 50.70% is paddy field and 49.30% is dry farmland, which is mainly located in the northern karst mountains. Even though 56,781.25 ha of land was newly cultivated by 2015, its average per unit yield is 4255.75 kg/ha, which is less than the lost farmland per unit yield of 4485.20 kg/ha (Figure 11). Therefore, from the perspective of general grain crop productivity, the productivity of supplementary cultivated land declined. However, the average per unit yield of total cultivated land increased from 4745.10 kg/ha to 5592.89 kg/ha, which arose from the beneficial effects from the physical environment and improved field management on existing farmland.

![Figure 11. The statistical characteristics of cultivated land change from 2000 to 2015.](image)
3.5. The Grain Crop Calorie Spatialization Based on the GCYS Model

Figure 12 shows the mixed grain crop (i.e., rice, wheat, maize, soybean, and potato) calorie map. The spatial distribution of calorie outputs is higher in the south and lower in the north. The average calories for 2000 and 2015 are $6.37 \times 10^6$ kcal/ha and $6.63 \times 10^6$ kcal/ha, respectively, and the per capita calories obtained from grain crops each day are about 654.76 kcal/person and 1025.77 kcal/person, respectively.

![Figure 12](image-url)

**Figure 12.** (a) The mixed grain crop calorie distribution in 2000; (b) The mixed grain crop calorie distribution in 2015.

It should be noted that the grain crop calorie distributions are not the same as the actual yield spatialization results, because they are related to the planting structure and the multiple indices. For example, the grain crop calories of relatively flat paddy field in the southeast, where rice is harvested twice a year, would be far greater than the calories of dry farmland in the northern region, where the standard cultivation system only allows harvest once a year and potatoes occupied a large percentage area. The unit mass calories of the potato are just one-fifth of the other four grain crops. In the counties with a low proportion of potato cultivation, the spatial pattern of calorie spatialization results is similar to that of the grain crop yield. Therefore, the variation of calorie values is larger than that in yield in the whole study area.

4. Discussion

This research aims to realize the spatialization of actual grain crop yield. Knowledge of cultivation systems and planting structures are essential preconditions for analyzing the grain crop production [1,28,60]. Therefore, we built a comprehensive agricultural system, which integrates the standard cultivation system, the planting structure, including all five grain crops (i.e., rice, wheat, maize, soybean, and potato), and the cropping difference between the paddy field and dry farmland at the county level in the Guizhou and Guangxi karst mountain region in China. This agricultural system reflects the agricultural production rather than only estimating one crop at the county level [6,12,18,21]. The characteristics of comprehensiveness and localization of the cultivation system underpin the basic framework of the GCYS model.

Previous studies focused on the physical factors limiting crop productivity [27,28,51]. However, improvements in field management technologies have resulted that anthropogenic factors gradually become the key elements of increasing crop yields [3]. Such yield differences at a micro-scale cannot
been described in accurate detail just by using physical limiting factors in previous studies. Therefore, based on former studies, questionnaires, and experiments in the Guizhou and Guangxi karst mountain region, we can spatialize the effects of mechanization and chemical fertilizer applications on improved crop production to establish the key factors module of the GCYS model. These factors highlighted the spatial differences in the grain crop yield at the grid level. The reliability of our GCYS model is validated by the significant correlations between spatialization results and agricultural survey data of paddy rice yields, and the determination coefficients are 0.72 and 0.76 in 2000 and 2015, respectively. Furthermore, validation at the county and grid scale helped to improve the accuracy of the GCYS model.

The cultivation area of grain crops in the paddy field or dry farmland is quite different in our study. The proportions of farmland used for grain crop cultivation in paddy fields are 74.22% and 71.84% in 2000 and 2015, respectively, according to the estimation of our agricultural system module. In contrast, the average proportions of dry farmland are only 49.97% in 2000, and decreased to 43.13% in 2015. The lower proportion of grain crops in dry farmland might be less accurately estimated than in paddy fields, because greater uncertainty existed in practice, especially whether the grain crops are cultivated on low or high productivity farmland [35]. This problem is hard to solve, unless the precise grain crop distribution can be obtained [15,35,42]. However, precise mapping of grain crop distribution is also difficult to implement, particularly in large-scale studies [11,23]. The validation and calibration process at both the county and grid scale of GCYS model helped to improve the accuracy of the GCYS model.

The per unit area yield in a grid is the mixed grain crop yield following the cropping structure within its county. Although the mixed process decreased the spatial heterogeneity of the grain crop yield, the grid yield is a comprehensive reflection of farmland productivity, field management level, and cultivated systems, where these factors are relatively stable or controllable. Therefore, drastic yield fluctuations because of changes in the type of cultivated grain crop or sown area resulting from unstable factors, like policy [16] or market forces [17], are avoided. Additional refinement of crop yield spatialization might come from the quantification of high-resolution nitrogen or other flows in cultivated land [8,34,61,62]. The relationship between chemical fertilizer application and crop yield would provide us a new perspective to improve actual yield spatialization. Furthermore, algorithms, such as cross-entropy and the general algebraic modeling system (GAMS) [17,26], can be coupled with our model to improve spatial allocation and optimization procedure, realizing a higher accuracy in the GCYS model.

The GCYS model can spatialize the agricultural survey data of grain crop yields coupled with physical and anthropogenic factors, while indicating the main limiting or improvement factors for local grain crop yields. With multiple demands for food security, economic development, and ecological protection [63,64], more informative and precise results, accurately spatialized through the GCYS model, will help to improve field management measures [17], generate more rational farmland use planning, and improve land use efficiency [63,65]. It is important to explore the production potential of cultivated land, especially where the scale of farmland is limited by physical or anthropogenic factors [3].

With the advances in remote sensing technology, efforts have been made to search for efficient ways of extracting the spatial distributions of major grain crops [15,66]. Some new indicators, such as the evaporative stress index (ESI), have been found to be significantly correlated with crop yields at the field scale [67]. Combined with the GCYS model, the accuracy of actual grain crop yield spatialization would be potentially improved. In order to avoid fluctuations caused by unstable factors, the GCYS model could also be applied to predict grain crop yield and spatialization patterns in combination with agricultural development planning in the future and provide a reference for decision-making.

5. Conclusions

The spatialized grain crop yield helped to explore the major obstacles or key factors for crop yields, contributed to an understanding of the spatial heterogeneity of agricultural resource distributions,
and might serve for agricultural development planning in the future. The novelty of the GCYS model is the construction of a comprehensive agricultural system, then coupled with key physical and anthropogenic factors to realize accurate and actual grain crop yield spatialization from agricultural survey data at a finer-scale grid, and finally refined the details of spatial heterogeneities.

The GCYS model is applied and tested while taking Guizhou and Guangxi Karst Mountains Region as the case study area. The results show that the mean per unit yield of mixed grain crops (i.e., rice, wheat, maize, soybean, and potato) weighted by sown area proportion of each grain crop of the total cultivated land is 4745.10 kg/ha and 5592.89 kg/ha in 2000 and 2015, respectively. Although the average per unit area yield of mixed grain crops increased from 2000 to 2015, its value of newly reclaimed cultivated lands trended downward in general. The main limiting and improvement factors are AWC and chemical fertilizer application, respectively. Meanwhile, the dominated proportional area of mechanization increased rapidly from 0.23% in 2000 to 17.47% in 2015. The significant correlation between agricultural survey data and spatialization results for rice, with a determination coefficient of 0.72 and 0.76 in 2000 and 2015, respectively. This validated the reliability of the GCYS model. The slopes of regression lines are close to 1, indicating the high accuracy of our model. Finally, a mixed grain crop calorie map is generated based on the actual yield spatialization results. This helps to identify the calories, where people could obtain from local grain crop cultivation. Emerging grain crop recognition and distribution mapping methods that are based on remote sensing technology can be combined with our GCYS model to further optimize the accuracy of spatialization, and our results are expected to provide a reference for agricultural development or land use planning in the future.

Author Contributions: J.L. and E.X. conceived and designed the experiments; methodology, J.L. and E.X. performed the model; J.L. and H.Z. analyzed the data; J.L., H.Z., and E.X. writing and revised the paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research is funded by the Strategic Priority Research Program of Chinese Academy of Sciences (XDA20040201), and the National Natural Science Foundations of China (Grant No. 41671097, 41601095).

Acknowledgments: We are thankful to the Strategic Priority Research Program of Chinese Academy of Sciences and National Natural Science Foundations of China for their financial support of the present study, and the comments and suggestions of the editor and reviewers.

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

References
1. Woittiez, L.S.; van Wijk, M.T.; Slingerland, M.; van Noordwijk, M.; Giller, K.E. Yield gaps in oil palm: A quantitative review of contributing factors. *Eur. J. Agron.* 2017, 83, 57–77. [CrossRef]

2. Mueller, N.D.; Gerber, J.S.; Johnston, M.; Ray, D.K.; Ramankutty, N.; Foley, J.A. Closing yield gaps through nutrient and water management. *Nature* 2012, 490, 254–257. [CrossRef] [PubMed]

3. Van Ittersum, M.; Cassman, K.G.; Grassini, P.; Wolf, J.; Tittonell, P.; Hochman, Z. Yield gap analysis with local to global relevance—A review. *Field Crops Res.* 2013, 143, 4–17. [CrossRef]

4. Zhang, W.; Cao, G.; Li, X.; Zhang, H.; Wang, C.; Liu, Q.; Chen, X.; Cui, Z.; Shen, J.; Jiang, R.; et al. Closing yield gaps in China by empowering smallholder farmers. *Nature* 2016, 537, 671–674. [CrossRef] [PubMed]

5. Iqbal, M.A.; Shen, Y.-J.; Stricevic, R.; Pei, H.; Sun, H.; Amiri, E.; Penas, A.; del Rio, S. Evaluation of the FAO AquaCrop model for winter wheat on the North China Plain under deficit irrigation from field experiment to regional yield simulation. *Agric. Water Manag.* 2014, 135, 61–72. [CrossRef]

6. Liu, J.; Wiberg, D.; Zehnder, A.J.B.; Yang, H. Modeling the role of irrigation in winter wheat yield, crop water productivity, and production in China. *Irrig. Sci.* 2007, 26, 21–33. [CrossRef]

7. Lüder, R.-M.H.; Qin, R.; Richner, W.; Stamp, P.; Noulas, C. Spatial variability of selected soil properties and its impact on the grain yield of oats (*Avena sativa* L.) in small fields. *J. Plant Nutr.* 2018, 41, 2446–2469. [CrossRef]

8. Liu, J.; You, L.; Amini, M.; Obersteiner, M.; Herrero, M.; Zehnder, A.J.B.; Yang, H. A high-resolution assessment on global nitrogen flows in cropland. *Proc. Natl. Acad. Sci. USA* 2010, 107, 8035–8040. [CrossRef]

9. Meisner, M.H.; Zaviezo, T.; Rosenheim, J.A. Landscape crop composition effects on cotton yield, Lygus hesperus densities and pesticide use. *Pest. Manag. Sci.* 2016, 73, 232–239. [CrossRef]
10. Huff, F.A.; Neill, J.C. Effects of natural climatic fluctuations on the temporal and spatial variation in crop yields. *J. Appl. Meteorol.* 1982, 21, 540–550. [CrossRef]

11. Lobell, D.B.; Asner, G.P.; Ortiz-Monasterio, J.; Benning, T.L. Remote sensing of regional crop production in the Yaqui Valley, Mexico: Estimates and uncertainties. *Agric. Ecosyst. Environ.* 2003, 94, 205–220. [CrossRef]

12. Chivas, W.; Mutanga, O.; Biradar, C. Application of remote sensing in estimating maize grain yield in heterogeneous African agricultural landscapes: A review. *Int. J. Remote Sens.* 2017, 38, 6816–6845. [CrossRef]

13. Leng, G.; Huang, M. Crop yield response to climate change varies with crop spatial distribution pattern. *Sci. Rep.* 2017, 7, 1463. [CrossRef] [PubMed]

14. Dittmar, J.; Voegelin, A.; Roberts, L.C.; Hug, S.J.; Saha, G.C.; Ali, M.A.; Badruzaman, A.B.M.; Kretzschmar, R. Spatial distribution and temporal variability of arsenic in irrigated rice fields in Bangladesh. 2. Paddy soil. *Environ. Sci. Technol.* 2007, 41, 5967–5972. [CrossRef]

15. Peng, D.; Huang, J.; Li, C.; Liu, L.; Huang, W.; Wang, F.; Yang, X. Modelling paddy rice yield using MODIS data. *Agric. For. Meteorol.* 2014, 184, 107–116. [CrossRef]

16. Deng, X.; Gibson, J.; Wang, P. Management of trade-offs between cultivated land conversions and land productivity in Shandong Province. *J. Clean. Prod.* 2017, 142, 767–774. [CrossRef]

17. You, L.; Wood, S.; Wood-Sichra, U.; Wu, W. Generating global crop distribution maps: From census to grid. *Agric. Syst.* 2014, 127, 53–60. [CrossRef]

18. Stehfest, E.; Heistermann, M.; Priess, J.; Ojima, D.S.; Alcamo, J. Simulation of global crop production with the ecosystem model DayCent. *Ecol. Model.* 2007, 209, 203–219. [CrossRef]

19. Marti, J.; Bort, J.; Slafer, G.A.; Araus, J. Can wheat yield be assessed by early measurements of normalized difference vegetation index? *Ann. Appl. Biol.* 2007, 150, 253–257. [CrossRef]

20. Wang, X.; Li, X.; Fischer, G.; Sun, L.; Tan, M.; Xin, L.; Liang, Z. Impact of the changing area sown to winter wheat on crop water footprint in the North China Plain. *Ecol. Indic.* 2015, 57, 100–109. [CrossRef]

21. Qader, S.H.; Dash, J.; Atkinson, P.M. Forecasting wheat and barley crop production in arid and semi-arid regions using remotely sensed primary productivity and crop phenology: A case study in Iraq. *Sci. Total Environ.* 2018, 613, 250–262. [CrossRef] [PubMed]

22. Deryng, D.; Sacks, W.J.; Barford, C.C.; Ramankutty, N. Simulating the effects of climate and agricultural management practices on global crop yield. *Glob. Biogeochem. Cycles* 2011, 25. [CrossRef]

23. Liu, J.; Williams, J.R.; Zehnder, A.J.; Yang, H. GEPI-Modelling wheat yield and crop water productivity with high resolution on a global scale. *Agric. Syst.* 2007, 94, 478–493. [CrossRef]

24. Zhong, L.; Baoguo, L. Spatial distribution of China grain output based on land use and population density. *Trans. Chin. Soc. Agric. Eng.* 2012, 28, 1–8. [CrossRef]

25. You, L.; Wood, S. An entropy approach to spatial disaggregation of agricultural production. *Agric. Syst.* 2006, 90, 329–347. [CrossRef]

26. You, L.; Wood, S.; Wood-Sichra, U. Generating plausible crop distribution maps for Sub-Saharan Africa using a spatially disaggregated data fusion and optimization approach. *Agric. Syst.* 2009, 99, 126–140. [CrossRef]

27. Qin, Y.; Liu, J.; Shi, W.; Tao, F.; Yan, H. Spatial-temporal changes of cropland and climate potential productivity in northern China during 1990–2010. *Food Secur.* 2013, 5, 499–512. [CrossRef]

28. Li, H.; Shi, W.; An, T.; Wang, B.; Li, S.; Wang, J.; Li, S. Comparison of the modeled potential yield versus the actual yield of maize in Northeast China and the implications for national food security. *Food Secur.* 2016, 9, 99–114. [CrossRef]

29. Gathala, M.K.; Timsina, J.; Islam, S.; Krupnik, T.J.; Bose, T.R.; Islam, N.; Rahman, M.; Hassain, L.; Harun-Ar-Rashid, M.; Ghosh, A.K.; et al. Productivity, profitability, and energetics: A multi-criteria assessment of farmers’ tillage and crop establishment options for maize in intensively cultivated environments of South Asia. *Field Crops Res.* 2016, 186, 32–46. [CrossRef]

30. Seo, S.N. Evaluation of the Agro-Ecological Zone methods for the study of climate change with micro farming decisions in sub-Saharan Africa. *Eur. J. Agron.* 2014, 52, 157–165. [CrossRef]

31. Ju, W.; Gao, P.; Zhou, Y.; Chen, J.M.; Chen, S.; Li, X. Prediction of summer grain crop yield with a process-based ecosystem model and remote sensing data for the northern area of the Jiangsu Province, China. *Int. J. Remote Sens.* 2010, 31, 1573–1587. [CrossRef]

32. Marshall, M.; Tu, K.; Brown, J. Optimizing a remote sensing production efficiency model for macro-scale GPP and yield estimation in agroecosystems. *Remote Sens. Environ.* 2018, 217, 258–271. [CrossRef]
33. Reeves, M.C.; Zhao, M.; Running, S.W. Usefulness and limits of MODIS GPP for estimating wheat yield. *Int. J. Remote Sens.* 2005, 26, 1403–1421. [CrossRef]

34. Zhou, G.; Yin, X. Assessing nitrogen nutritional status, biomass and yield of cotton with ndvi, spad and petiole sap nitrate concentration. *Exp. Agric.* 2017, 54, 531–548. [CrossRef]

35. Cheng, Z.; Meng, J.; Wang, Y. Improving spring maize yield estimation at field scale by assimilating time-series HJ-1 CCD data into the WOFOST model using a new method with fast algorithms. *Remote Sens.* 2016, 8, 303. [CrossRef]

36. Diepen, C.; Wolf, J.; Keulen, H.; Rappoldt, C. WOFOST: A simulation model of crop production. *Soil Use Manag.* 1989, 5, 16–24. [CrossRef]

37. Wu, L.; Bai, J.; Xiao, Q.; Du, Y.; Liu, Q.; Xu, L. Research progress and prospect on combining crop growth models with parameters derived from quantitative remote sensing. *Trans. Chin. Soc. Agric. Eng.* 2017, 33, 155–166.

38. Sakuma, S.; Schnurbusch, T. Of floral fortune: Tinkering with the grain yield potential of cereal crops. *New Phytol.* 2019, 225, 1873–1882. [CrossRef]

39. Deng, X.; Huang, J.; Rozelle, S.; Uchida, E. Cultivated land conversion and potential agricultural productivity in China. *Land Use Policy* 2006, 23, 372–384. [CrossRef]

40. Myers, S.S.; Smith, M.R.; Guth, S.; Golden, C.D.; Vaitla, B.; Mueller, N.D.; Dangour, A.D.; Huybers, P. Climate change and global food systems: Impacts on food security and undernutrition. *Annu. Rev. Public Health* 2017, 38, 259–277. [CrossRef]

41. Fu, Q.; Li, B.; Hou, Y.; Bi, X.; Zhang, X. E

42. Ouyang, L.; Mao, D.; Wang, Z.; Li, H.; Man, W.; Jia, M.; Liu, M.; Zhang, M.; Liu, H. Analysis crops planting structure and yield based on GF-1 and Landsat8 OLI images. *Trans. Chin. Soc. Agric. Eng.* 2017, 33, 147.

43. Li, J.; Zhang, H.; Xu, E. A two-level nested model for extracting positive and negative terrains combining morphology and visualization indicators. *Ecol. Indic.* 2020, 109, 105842. [CrossRef]

44. Godfray, H.C.J.; Beddington, J.R.; Crute, I.R.; Haddad, L.; Lawrence, D.; Muir, J.F.; Pretty, J.; Robinson, S.; Thomas, S.M.; Toulmin, C. Food security: The challenge of feeding 9 billion people. *Science* 2010, 327, 812–818. [CrossRef] [PubMed]

45. Wang, X.; Tian, H. Approaching on the corrected methods of topographic factors for calculating the climatic potential productivity of crops. *Chin. J. Agrometeorol.* 1996, 17, 24–28.

46. Field, C.B.; Behrenfeld, M.J.; Randerson, J.T.; Falkowski, P. Primary production of the biosphere: Integrating terrestrial and oceanic components. *Science* 1998, 281, 237–240. [CrossRef]

47. Lieth, H. Modeling the primary productivity of the world. In *Primary Productivity of the Biosphere*; Springer: Berlin, Germany, 1975.

48. Zheng, Z.; Fu, B.; Hu, H.; Sun, G. A method to identify the variable ecosystem services relationship across time: A case study on Yanhe Basin, China. *Landsc. Ecol.* 2014, 29, 1689–1696. [CrossRef]

49. Lieth, H.; Box, E. Evapotranspiration and primary productivity: C.W. Thornthwaite memorial model. *Pub. Climatol.* 1972, 25, 37–46.

50. Shang, E.; Xu, E.; Zhang, H.; Liu, F. Analysis of spatiotemporal dynamics of the Chinese vegetation net primary productivity from the 1960s to the 2000s. *Remote Sens.* 2018, 10, 860. [CrossRef]

51. Yan, H.; Liu, J.; Cao, M. Spatial pattern and topographic control of China’s agricultural productivity variability. *Acta Geogr. Sin.* 2007, 62, 171–180.

52. Lobell, D.B.; Hicke, J.A.; Asner, G.P.; Field, C.; Tucker, C.J.; Los, S. Satellite estimates of productivity and light use efficiency in United States agriculture, 1982–1998. *Glob. Chang. Biol.* 2002, 8, 722–735. [CrossRef]

53. Zhou, W.; Liu, G.; Pan, J.; Feng, X. Distribution of available soil water capacity in China. *J. Geogr. Sci.* 2005, 1, 5–14.

54. Wang, S.; Su, W.; Fan, X.; Li, C.; Shi, X. Influence factors of soil moisture in karst rocky desertification region—A case study of Puding County, Guizhou Province. *Res. Soil Water Conserv.* 2010, 17, 171–175.

55. Lv, Z. Study on the Evaluation, Effect, and Consolidation of Cultivated Land Fragmentation—An Empirical Analysis Based on Quzhou County; China Agricultural University: Beijing, China, 2015.

56. Chen, C. The Mechanization Mode of Rice Cultivation in The Hills and Mountains of Southern China; Chinese Academy of Agricultural Sciences: Nanjing, China, 2012.
57. Zhou, Z.; Mu, N. Study on the influence of agricultural mechanization on grain yield increase in China. *China Price* 2015, 11, 27–29.
58. Li, Z. Farmers Use Pesticides and Fertilizers Behavior and Its Influencing Factors—Take Jiangxi Province as an Example; Jiangxi Agricultural University: Nanchang, China, 2015.
59. Ouyang, Z.; Zheng, H.; Xiao, Y.; Polasky, S.; Liu, J.; Xu, W.; Wang, Q.; Zhang, L.; Rao, E.; Jiang, L.; et al. Improvements in ecosystem services from investments in natural capital. *Science* 2016, 352, 1455–1459. [CrossRef]
60. Yan, X.; Gong, W. The role of chemical and organic fertilizers on yield, yield variability and carbon sequestration—Results of a 19-year experiment. *Plant Soil* 2010, 331, 471–480. [CrossRef]
61. Hayashi, K.; Matsumoto, N.; Hayashi, E.T.; Abdoulaye, T.; Shinjo, H.; Tabo, R.; Matsunaga, R.; Tobita, S. Estimation of nitrogen flow within a village-farm model in Fakara region in Niger, Sahelian zone of West Africa. *Nutr. Cycl. Agroecosyst.* 2012, 92, 289–304. [CrossRef]
62. Sadras, V.O.; Hayman, P.T.; Rodriguez, D.; Monjardino, M.; Bielich, M.; Unkovich, M.; Mudge, B.; Wang, E. Interactions between water and nitrogen in Australian cropping systems: Physiological, agronomic, economic, breeding and modelling perspectives. *Crop Pasture Sci.* 2016, 67, 1019–1053. [CrossRef]
63. Wang, J.; Chen, Y.; Shao, X.; Zhang, Y.; Cao, Y. Land-use changes and policy dimension driving forces in China: Present, trend and future. *Land Use Policy* 2012, 29, 737–749. [CrossRef]
64. Xu, E.; Zhang, H.; Yang, Y.; Zhang, Y. Integrating a spatially explicit tradeoff analysis for sustainable land use optimal allocation. *Sustainability* 2014, 6, 8909–8930. [CrossRef]
65. Xu, E.; Zhang, H.; Yao, L. An elevation-based stratification model for simulating land use change. *Remote Sens.* 2018, 10, 1730. [CrossRef]
66. Jiang, M.; Xin, L.; Li, X.; Tan, M.; Wang, R. Decreasing rice cropping intensity in southern China from 1990 to 2015. *Remote Sens.* 2018, 11, 35. [CrossRef]
67. Yang, Y.; Anderson, M.; Gao, F.; Wardlow, B.; Hain, C.R.; Otkin, J.A.; Alfieri, J.; Yang, Y.; Sun, L.; Dulaney, W. Field-scale mapping of evaporative stress indicators of crop yield: An application over Mead, NE, USA. *Remote Sens. Environ.* 2018, 210, 387–402. [CrossRef]