Effects of Climate Change on Corn Yields: Spatiotemporal Evidence from Geographically and Temporally Weighted Regression Model

Bing Yang 1, Sensen Wu 1,2 and Zhen Yan 3,4,*

1 School of Earth Sciences, Zhejiang University, Hangzhou 310027, China; zjuyangbing@zju.edu.cn (B.Y.);
wusensengis@zju.edu.cn (S.W.)
2 Zhejiang Provincial Key Laboratory of Geographic Information Science, Hangzhou 310028, China
3 Center of Agricultural and Rural Development, School of Public Affairs, Zhejiang University, Hangzhou 310058, China
4 Laboratory of Agricultural and Rural Development Intelligent Computation, School of Public Affairs, Zhejiang University, Hangzhou 310058, China
* Correspondence: yanzhen@zju.edu.cn; Tel.: +86-571-5633-7015

Abstract: Food security has been one of the greatest global concerns facing the current complicated situation. Among these, the impact of climate change on agricultural production is dynamic over time and space, making it a major challenge to food security. Taking the U.S. Corn Belt as an example, we introduce a geographically and temporally weighted regression (GTWR) model that can handle both temporal and spatial non-stationarity in the relationship between corn yield and meteorological variables. With a high fitting performance (adjusted $R^2$ at 0.79), the GTWR model generates spatiotemporally varying coefficients to effectively capture the spatiotemporal heterogeneity without requiring completion of the unbalanced data. This model makes it possible to retain original data to the maximum possible extent and to estimate the results more reliably and realistically. Our regression results showed that climate change had a positive effect on corn yield over the past 40 years, from 1981 to 2020, with temperature having a stronger effect than precipitation. Furthermore, a fuzzy c-means algorithm was used to cluster regions based on spatiotemporally changing trends. We found that the production potential of regions at high latitudes was higher than that of regions at low latitudes, suggesting that the center of productive regions may migrate northward in the future.

Keywords: climate change; corn yield; geographically and temporally weighted regression; spatiotemporal variation of relationships; U.S. Corn Belt

1. Introduction

Global food security is a major challenge to the sustainable development goals and implementation of the 2030 Paris Agreement on Climate Change. Climate solutions aiming at effective adaption to climate change are needed, along with further research into its dual role in requiring a response from agri-producers and as a natural factor in food security. In addition to socioeconomic factors such as scientific and technological progress, farming costs and benefits, and land use patterns, these natural factors including soil quality and climatic conditions [1–3] are essential to crop production. Beginning in the 1950s, as the scientific evidence of human greenhouse gas emissions affecting important agricultural factors such as temperature and precipitation has continued to accumulate, scholars and experts have become convinced that climate change will significantly impact crop yields [4–8]. Therefore, measuring the potential impact of meteorological variability on crop yield is critical for farm management and food security [9,10].

Many studies have been conducted to understand the impact of climate change on crop yield. For example, using regression analysis of historical data to link past yields
with weather realizations, Lobell showed that at national and regional scales temperature changes were more important than precipitation changes for maize, rice, and wheat [11]. Scholars have found that climate extremes have a much stronger impact on crop yield in rain-fed areas than in irrigated areas [12]. In addition, under different conditions, meteorological factors such as precipitation have been shown to have different effects on corn yield [13]. Several studies have shown that the degree of negative susceptibility to excessive precipitation and of positive susceptibility to precipitation under drought conditions are of similar magnitude [14].

In order to research the effects of climate change on crop yield, two modeling strategies have been extensively applied, namely, crop simulation [13,15] and statistical models [2,16,17]. The crop simulation model, known as the process-based model, can accurately control for all crop growth factors (i.e., temperature, humidity, soil, fertilizer, etc.) and examine the impact of different climate change scenarios on crop growth, development, and yield [18]. The decision support system for agrotechnology transfer (DSSAT), for example, typically requires various input data such as crop variety, soil conditions, and farm management [19]. However, crop growth simulation experiments are difficult to implement at large scales, which is necessary in order to obtain professional technical information and can additionally incur high acquisition costs [11]. The second approach is statistical modeling, which extends crop production variables in order to perform regression estimations.

Initially, scholars can estimate the impact of weather by analyzing year-to-year weather changes and their impact on yield using time series models [2,20]. However, considering the long-term impact of climate change on yield changes, as production processes are updated (as exemplified by variety improvement), time-invariant analysis cannot adequately capture the impact [11]. In addition, cross-sectional models that account for spatial autocorrelation have reflected the historical effectiveness of adaptation in maintaining the yield of a crop across different climates.

Next, in light of the heterogeneity of crop production systems and significant interannual meteorological variability, the regression coefficients of meteorological variables are expected to vary both spatially and temporally [6,16,21–23]. As a result, panel models that consider both time and space have been widely used [24–26]. In the past, panel data have allowed for the use of either fixed or random effects models to control for unobserved heterogeneity. Here, we have improved the estimation efficiency by conducting panel data analysis.

However, the panel model has a fatal drawback in that it requires balanced input data and does not perform well with unbalanced data. Therefore, data completion has become a necessary process for research. The most commonly used data completion methods consider similar yield data over similar times or spaces and preset the distribution characteristics of yield in advance. This likely produces interference when determining the relationships between crop outcomes and various climatic factors.

To use original unbalanced data directly and estimate the results more reliably and realistically, we must develop a novel model. At the same time, this new model should deal with both spatial and temporal non-stationary processes. Luckily, a frequently used cross-section model, geographically weighted regression (GWR), has demonstrated a powerful ability to capture related spatial datasets [27–29]. Here, we further propose a spatiotemporal regressive model, the geographically and temporally weighted regression (GTWR) model, which provides a broader and more flexible inference basis for spatiotemporal yield variability analysis by accounting for the fusion of temporal and spatial heterogeneity [30]. Meanwhile, using local regression, GTWR can deal effectively with unbalanced data. GTWR has already been widely used to determine spatiotemporal relationships in areas such as housing prices and marine ecology [31–33]. Thus, we can use the GTWR model to handle both the unbalanced data and the spatiotemporal non-stationary characteristics of the coefficients.

The purpose of the present research was to reveal persuasive spatial and temporal relationships between meteorological variables and corn yield using county-level data in
the U.S. Corn Belt from 1981 to 2020. In order to demonstrate the advantages of this model, we further compared it with ordinary least squares (OLS) (Equation S1) [6], time-linear regression (time-LR, a time series method) (Equation S2) [34], and GWR (a cross-sectional method) (Equation S3) [27] in terms of model fitting performance (Table S1).

The rest of the paper is structured as follows. We introduce the study area and data sources in Section 2. In Section 3, we describe in detail the analytical methods used and propose a set of methods for investigating variation in corn yield with meteorological variables based on the GTWR model. In Section 4, we analyze how climatic variables affect yield changes in time and space. Finally, in Section 5 we summarize our findings, provide conclusions, and discuss the strengths and limitations of the study.

2. Study Area and Data

2.1. Study Area

In this study, we focused on the main corn cropping regions in the U.S. (Figure 1). This region, called the Corn Belt, covers 13 States (Illinois, Indiana, Iowa, Kansas, Kentucky, Michigan, Minnesota, Missouri, Nebraska, Ohio, Wisconsin, North Dakota, and South Dakota), and has a humid continental climate with rich and deep soils and is located in the center of the country [35,36]. The Corn Belt is the world’s largest specialized corn-producing area, and has the highest yield [20]; corn production here accounts for three-quarters of the country’s total.

![Figure 1. Location of the Corn Belt.](https://www...)

2.2. Data Collection and Preprocessing

Corn yield data from 1981 to 2020 were collected from the National Agricultural Statistics Service (NASS) of the U.S. Department of Agriculture (USDA) (https://www...
nass.usda.gov/, accessed on 24 March 2021) at the county level (unit: tons/acre) and converted to tons/ha. Counties that had too much missing data or no longer grew corn were eliminated.

From the existing data, yield records were occasionally missing for a variety of reasons, making the original yield data unbalanced. For example, certain counties skipped the survey for many years. Fortunately, the method used in this study could handle unbalanced data well, and the preprocessing method no longer completed missing data. Our final sample included 1159 counties and 41,693 county-year records (Table 1).

Table 1. Corn yield data at the county level (1981–2020).

<table>
<thead>
<tr>
<th>State Name</th>
<th>County Count</th>
<th>Records</th>
</tr>
</thead>
<tbody>
<tr>
<td>ILLINOIS</td>
<td>102</td>
<td>3965</td>
</tr>
<tr>
<td>INDIANA</td>
<td>92</td>
<td>3517</td>
</tr>
<tr>
<td>IOWA</td>
<td>99</td>
<td>3931</td>
</tr>
<tr>
<td>KANSAS</td>
<td>105</td>
<td>3737</td>
</tr>
<tr>
<td>KENTUCKY</td>
<td>113</td>
<td>3775</td>
</tr>
<tr>
<td>MICHIGAN</td>
<td>81</td>
<td>2603</td>
</tr>
<tr>
<td>MINNESOTA</td>
<td>88</td>
<td>3077</td>
</tr>
<tr>
<td>MISSOURI</td>
<td>114</td>
<td>3597</td>
</tr>
<tr>
<td>NEBRASKA</td>
<td>93</td>
<td>3524</td>
</tr>
<tr>
<td>OHIO</td>
<td>86</td>
<td>3320</td>
</tr>
<tr>
<td>WISCONSIN</td>
<td>71</td>
<td>2565</td>
</tr>
<tr>
<td>NORTH DAKOTA</td>
<td>53</td>
<td>1747</td>
</tr>
<tr>
<td>SOUTH DAKOTA</td>
<td>65</td>
<td>2335</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1159</strong></td>
<td><strong>41,693</strong></td>
</tr>
</tbody>
</table>

Daily county-level meteorological data were obtained from Parameter Elevation Regression on Independent Slopes (PRISM) (https://prism.oregonstate.edu, accessed on 1 August 2021) using the Applied Climate Information System API (2017). The PRISM variables used for analysis were daily mean air temperature (TEMP), maximum air temperature (T_max), minimum air temperature (T_min), total precipitation (PCPN), maximum vapor pressure deficit (VPD_max), and minimum vapor pressure deficit (VPD_min) [37], all in metric units (°C, mm, hPa).

As PRISM meteorology data are available as continuous surfaces (spatial raster layers) with a spatial resolution of 4 km, zonal statistics were obtained by overlapping the data with county boundaries to estimate the average values of climate-related factors at the county level [6].

3. Methodology

We proposed a novel spatiotemporal analytical framework to measure the potential impact of meteorological variability on crop yield in the U.S. Corn Belt (as shown in Figure 2). First, we needed to remove the portion affected by the time trend from the corn yield data, for which we used the linear detrend method. This enabled us to focus only on climate data. The original climate variables were obtained with new variables that were more in line with the requirements of crop production, such as effective accumulated temperature. A powerful spatiotemporal regression model, the GTWR model, was then implemented to discover the relationships between climate change and crop production by deriving spatiotemporal coefficients. Furthermore, the spatiotemporal coefficients were clustered using a fuzzy c-means algorithm to discover regions with similar trends over the past 40 years.
3.1. Linear Detrend Preprocessing

In order to remove the influence of crop growing practices and conditions [38,39], we detrended the corn yield for county $i$ at year $t$ by removing the linear fits for the yield model using Equation (1):

$$Y_{d,i,t} = Y_{i,t} - Y^*_{i,1981}$$

where $Y_{d,i,t}$ represents the detrended values, $Y_{i,t}$ represents the observed yield, and $Y^*_{i,t}$ represents the predicted yield; $Y^*_{i,1981}$ indicates the initial value in the year 1981.

3.2. Interval Accumulated Temperature Method

Generally, the accumulated temperature in a certain temperature range is defined as the effective accumulated temperature and is referred to as growing degree days (GDD) [35,40]. GDD serves as a necessary thermal resource for crop growth. In an extreme temperature range, the accumulated temperature is defined as harmful accumulated temperature, known as killing degree days (KDD) [6,41]; this directly indicates the heat stress. The daily accumulated temperature is summed based on the crop growth period to measure the impact of temperature on crop yield. These two variables can be calculated using Equations (2)–(4):

$$TEMP = \frac{(T_{max} + T_{min})}{2}$$

$$GDD = \sum_{period} TEMP - TB$$

$$KDD = \sum_{period} Tmax - Tu$$

where $Tu$ and $Tb$, the upper and lower bounds, respectively, are set as 29 °C and 8 °C, respectively [6,42].

3.3. GTWR Model

To fit the inherent spatial non-stationarity of geographic relationships, we harness the idea of local smoothing, specifically a GWR model, to change the regression coefficients alongside changes in spatial location [43,44]. The GWR model is an extension of the OLS model, which embeds the spatial relationship changes caused by spatial position differentiation into the calculation of regression coefficients.
Huang introduced time and space factors to balance the differences in time and space scales, embedded the changes in the time dimension into the GWR model, proposed a method for constructing time and space distance, and established a time and space GTWR model [30]. This derivation proves the theoretical feasibility of replacing time and space factors with space-time factors, thereby reducing computational complexity. Combined with the relationship between crop yield and climate factors studied in this article, the mathematical expression of the GTWR model is described by Equation (5):

$$
\begin{align*}
\text{dY}_{(u_i,v_i,t_i)} &= \hat{\beta}_{(u_i,v_i,t_i),0} + \hat{\beta}_{(u_i,v_i,t_i),1}GDD_{(u_i,v_i,t_i)} + \hat{\beta}_{(u_i,v_i,t_i),2}KDD_{(u_i,v_i,t_i)} + \\
&\hat{\beta}_{(u_i,v_i,t_i),3}VPD_{(u_i,v_i,t_i)} + \hat{\beta}_{(u_i,v_i,t_i),4}PCPN_{(u_i,v_i,t_i)} + \epsilon_{(u_i,v_i,t_i)}
\end{align*}
$$

where \((u_i,v_i,t_i)\) represents the space-time coordinates of the sample county \(i\), \(\beta_{(u_i,v_i,t_i),0}\) is the regression constant, \(\beta_{(u_i,v_i,t_i),1-4}\) is the regression coefficient of each independent variable and is a function of space-time position \((u_i,v_i,t_i)\), \(\epsilon_{(u_i,v_i,t_i)}\) is the error term of sample county \(i\) and is subject to \(\epsilon_i \sim N(0,\sigma^2)\), and \(\text{Cov}(\epsilon_i,\epsilon_j) = 0\) (\(i \neq j\)).

The Gaussian and bi-square kernel functions are currently the two most commonly used kernel functions. The mathematical expression of the Gaussian kernel function is described using Equation (6):

$$
w_{ij} = \exp\left(-\frac{d_{ij}^2}{b^2}\right)$$

The Gaussian kernel function assigns weights to all sample points when calculating regression coefficients. In reality, when a certain distance is exceeded, the spatial influence of geographic element relationships should decrease to 0. To improve the computational efficiency of the model, a bi-square kernel function is adopted, as shown in Equation (7):

$$
w_{ij} = \left\{ \begin{array}{ll}
1 - \left(\frac{d_{ij}}{b}\right)^2 & , \text{if } d_{ij} < b \\
0, & \text{otherwise}
\end{array} \right.$$

where \(d_{ij}\) represents the distance between sample points \(i\) and \(j\). In the GWR model, \(d_{ij}\) represents spatial distance, and in the GTWR model, it represents the space-time distance. Bandwidth \(b\) is a parameter describing the non-negative decreasing relationship between weight and distance.

### 3.4. Soft Clustering Using the Fuzzy C-Means Algorithm

Clustering is an important tool for agricultural zone analysis [45]. As the next step in the analysis of spatiotemporal regions, we needed to find an efficient spatiotemporal clustering method that could be used to examine the spatiotemporal relationship between yield and meteorology. In agronomy, the agglomerative hierarchical clustering method [46] and machine learning [47] have been widely implemented to solve this problem. However, the former method cannot be used for large amounts of data and the latter method has a complicated learning process, making both methods unsuitable for this study. We discovered that fuzzy clustering (sometimes referred to as soft clustering or soft k-means) can efficiently handle spatiotemporal coefficient data [48–50].

The fuzzy c-means algorithm, which has been extensively applied for microarray data analysis and clustering, is based on the iterative optimization of an objective function to minimize variations in objects within clusters. It meets the clustering needs when evaluating the impact of meteorological variables on crop yield because the influence coefficients obtained from the GTWR model are microarray data.

The fuzzy c-means algorithm is very similar to the k-means algorithm, and is implemented as follows:

1. Choose a number of clusters;
2. Assign coefficients randomly to each data point based on the cluster to which it belongs;
3. Repeat until the algorithm has converged (i.e., the change in coefficients between two iterations is no more than $\epsilon$, which is the assigned sensitivity threshold);
4. Compute the centroid for each cluster and then, for each data point, compute its coefficients based on the cluster to which it belongs.

We used an R package, Mfuzz, to implement this soft clustering method. This helped us to determine when and where the effects of meteorological variables on corn yield followed similar trends in time and space.

4. Results
4.1. How Detrended Yield and Meteorological Variables Changed over 40 Years

After applying the linear detrending treatments from 1981 to 2020, the corn yield trends were very flat (Figure 3), indicating that time-dependent factors were almost completely eliminated. As the improvements in production brought about by technological advances and human behavior were removed, we were able to focus on the impact of climate change on yield. In addition, low values fluctuated more widely than high values, indicating that corn yield was susceptible to negative volatility. One of the purposes of this study was, therefore, to reduce the negative impact of meteorological variables on corn yield by determining how climate change affects yields.

![Figure 3. Detrended corn yield (de_yield) from year 1981 to 2020, with the trend line shown in red.](image)

Table 2 shows the result of correlation analysis for 40 years of 41,693 county-years of recorded yield data along with the TEMP, GDD, KDD, VPD, and PCPN variables aggregated over growing seasons. Traditionally, the growing season for the Corn Belt runs from May 1 to October 31. However, for simplicity, researchers may define the corn growing season as June, July, and August, i.e., the summer [51,52]. We found that all of the meteorological factors in summer had better relevance with the detrended yield than in the growing season, especially for TEMP, GDD, and PCPN.
Table 2. Correlation analysis results of meteorological variables and detrended yield, with all p-values < 0.001. TEMP_all, GDD_all, KDD_all, VPD_all, and PCPN_all data are aggregated from May 1 to October 31, whereas TEMP_s, GDD_s, KDD_s, VPD_s, and PCPN_s data are aggregated over the summer months of June, July, and August.

<table>
<thead>
<tr>
<th></th>
<th>TEMP_all</th>
<th>GDD_all</th>
<th>KDD_all</th>
<th>VPD_all</th>
<th>PCPN_all</th>
<th>TEMP_s</th>
<th>GDD_s</th>
<th>KDD_s</th>
<th>VPD_s</th>
<th>PCPN_s</th>
<th>De-yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEMP_all</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDD_all</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KDD_all</td>
<td>0.70</td>
<td>0.72</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VPD_all</td>
<td>0.56</td>
<td>0.58</td>
<td>0.92</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCPN_all</td>
<td>0.20</td>
<td>0.18</td>
<td>−0.26</td>
<td>−0.43</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TEMP_s</td>
<td>0.93</td>
<td>0.95</td>
<td>0.80</td>
<td>0.66</td>
<td>0.10</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDD_s</td>
<td>0.93</td>
<td>0.95</td>
<td>0.80</td>
<td>0.66</td>
<td>0.10</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KDD_s</td>
<td>0.67</td>
<td>0.70</td>
<td>0.99</td>
<td>0.91</td>
<td>−0.26</td>
<td>0.80</td>
<td>0.80</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VPD_s</td>
<td>0.46</td>
<td>0.48</td>
<td>0.90</td>
<td>0.96</td>
<td>−0.44</td>
<td>0.61</td>
<td>0.61</td>
<td>0.91</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCPN_s</td>
<td>0.07</td>
<td>0.06</td>
<td>−0.30</td>
<td>−0.39</td>
<td>0.78</td>
<td>−0.02</td>
<td>−0.02</td>
<td>−0.32</td>
<td>−0.47</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>De-yield</td>
<td>−0.06</td>
<td>−0.07</td>
<td>−0.30</td>
<td>−0.26</td>
<td>0.19</td>
<td>−0.17</td>
<td>−0.17</td>
<td>−0.32</td>
<td>−0.32</td>
<td>0.27</td>
<td>1.00</td>
</tr>
</tbody>
</table>

For temperature-dependent factors in the summer, the correlation coefficient between KDD and the detrended yield was −0.32, which was much higher than the coefficient of −0.17 between both TEMP and GDD and the detrended yield; the correlation between TEMP and GDD was close to 1. All three temperature-related factors were negatively correlated with corn yield. For water-dependent variables, the coefficients of VPD (−0.32) and PCPN (0.27) were not significantly different in terms of absolute value, although the yield was negatively correlated with VPD and positively correlated with PCPN. Ultimately, we chose GDD, KDD, VPD, and PCPN in the summer months as the factors for our next experiment.

From Figure 4, it can be seen that GDD increased over the past 40 years, while KDD decreased. This indicates that although the temperature increased, instances of extremely hot weather decreased, in turn implying that the impact of global warming on the U.S. Corn Belt may not have been completely negative. Furthermore, the upward trend in PCPN and the downward trend in VPD were very clear. Figure 4 shows that the annual distribution of meteorological variables in the Corn Belt was uneven, showing several sudden decreases or increases in accumulated temperature or precipitation between 1981 and 2020. Cumulative temperature GDD and KDD plummeted in 1992, 2004, 2009, and 2014, whereas KDD increased abruptly in 1983, 1988, and 2012, implying that the former years were much cooler and the latter years were hotter than average summers. In 1988 and 2012, there was a sharp rise in VPD and a sudden decline in PCPN, implying extreme dry weather.

Combining the detrended corn yield and meteorological variables, we found a significant nadir in 2012 that coincided with notable heat and drought. In 1983, 1988, 1993, and 2002, when corn production fell deeply, several climate conditions were either higher or lower than normal conditions.
4.2. Spatiotemporal Model Fitting and Performance in Assessing Climate Change Impact on Yield

We applied several spatiotemporal models to the detrended yield and meteorological variables in summer that were chosen in the previous section to fit reliable yield–meteorology relationships. It can be seen from Table 3 that the GTWR outperformed alternative models, namely, OLS, time-LR, and GWR, in fitting corn yield. GTWR had the highest adjusted $R^2$ of 0.79, obtained using a bi-square kernel. The GTWR-bi-square had the lowest RMSE of all models, with a value of 0.59 tons/ha. Furthermore, GTWR, which is based on spatiotemporally non-stationary processes, outperformed GWR, which only accounts for spatial autocorrelation, in terms of almost all factors. GTWR similarly outperformed time-LR, which only considers temporal autocorrelation, and OLS, which considers neither autocorrelation. The OLS and time-LR models performed exceptionally poorly compared to other traditional studies, as we used unbalanced data without data completion in order to demonstrate the superiority of the GWR and GTWR models on such tasks.

Figure 4. Changes in meteorological variables over the summer months from 1981 to 2020.

- **GDD (°C)**: $y = 1.0712x + 1295.4$
- **KDD (°C)**: $y = -0.7406x + 149.15$
- **VPD (hPa)**: $y = -1.8138x + 1067.2$
- **PCPN (mm)**: $y = 1.7183x + 262.66$
Table 3. Summary of parameters and performances of ordinary least square (OLS), time linear regression (LR), geographically weighted regression (GWR) with Gaussian kernel, GWR with the bi-square kernel, geographically and temporally weighted regression (GTWR) with Gaussian kernel, and GTWR with bi-square kernel models; values in round brackets are the first and third quantiles of the coefficients.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>OLS</th>
<th>Time-LR</th>
<th>GWR-Gaussian</th>
<th>GWR-Bi-Square</th>
<th>GTWR-Gaussian</th>
<th>GTWR-Bi-Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient of GDD</td>
<td>0.0012</td>
<td>0.0021</td>
<td>0.0019</td>
<td>0.0020</td>
<td>0.0012</td>
<td>0.0010</td>
</tr>
<tr>
<td>(tons/ha/°C)</td>
<td>(−0.0033, 0.0044)</td>
<td>(−0.0000, 0.0036)</td>
<td>(−0.0002, 0.0042)</td>
<td>(−0.0016, 0.0039)</td>
<td>(−0.0020, 0.0041)</td>
<td></td>
</tr>
<tr>
<td>Coefficient of KDD</td>
<td>−0.0063</td>
<td>−0.0121</td>
<td>−0.0114</td>
<td>−0.0119</td>
<td>−0.0020</td>
<td>−0.0017</td>
</tr>
<tr>
<td>(tons/ha/°C)</td>
<td>(−0.0171, −0.0071)</td>
<td>(−0.0153, −0.0075)</td>
<td>(−0.0168, −0.0072)</td>
<td>(−0.0016, −0.0060)</td>
<td>(−0.0106, 0.0069)</td>
<td></td>
</tr>
<tr>
<td>Coefficient of VPD</td>
<td>0.0006</td>
<td>−0.0005</td>
<td>−0.0005</td>
<td>−0.0006</td>
<td>−0.0007</td>
<td>−0.0006</td>
</tr>
<tr>
<td>(tons/ha/Pa)</td>
<td>(−0.0025, 0.0014)</td>
<td>(−0.0021, 0.0010)</td>
<td>(−0.0024, 0.0013)</td>
<td>(−0.0030, 0.0016)</td>
<td>(−0.0031, 0.0019)</td>
<td></td>
</tr>
<tr>
<td>Coefficient of PCPN</td>
<td>0.0019</td>
<td>0.0004</td>
<td>0.0005</td>
<td>0.0004</td>
<td>0.0022</td>
<td>0.0022</td>
</tr>
<tr>
<td>(tons/ha/mm)</td>
<td>(−0.0018, 0.0021)</td>
<td>(−0.0012, 0.0021)</td>
<td>(−0.0016, 0.0020)</td>
<td>(−0.0001, 0.0041)</td>
<td>(−0.0003, 0.0043)</td>
<td></td>
</tr>
<tr>
<td>Moran’s I of residuals</td>
<td>0.53</td>
<td>0.44</td>
<td>0.33</td>
<td>0.32</td>
<td>0.22</td>
<td>0.19</td>
</tr>
<tr>
<td>Adj.R²</td>
<td>0.14</td>
<td>0.44</td>
<td>0.59</td>
<td>0.63</td>
<td>0.73</td>
<td>0.79</td>
</tr>
<tr>
<td>MAE</td>
<td>0.89</td>
<td>0.74</td>
<td>0.66</td>
<td>0.64</td>
<td>0.50</td>
<td>0.44</td>
</tr>
<tr>
<td>RMSE (tons/ha)</td>
<td>1.18</td>
<td>0.96</td>
<td>0.88</td>
<td>0.86</td>
<td>0.67</td>
<td>0.59</td>
</tr>
</tbody>
</table>

Based on the Moran’s I of the residuals in Table 3, all of the experimental models were significantly positive, indicating that neighboring counties tended to have similar degrees of estimation bias. As the GTWR model considered spatial autocorrelation, it had fewer spatial clustered residuals. Unexpectedly, incorporating the temporal non-stationarity in the GWR model significantly reduced the spatial autocorrelation of residuals in GTWR to 0.22 and 0.19, compared to 0.33 and 0.32 for the GWR ($p < 0.01$ for all years and models). These findings prove that temporal non-stationarity plays an important role in resolving spatial autocorrelation.

The improved performance of the GTWR model indicates that non-stationary spatiotemporal processes are more suitable for the description of the relationship between meteorology and yield. GTWR with a bi-square kernel showed a slightly better performance than GTWR with a Gaussian kernel, making the former the best model for our study.

We then used the 40-year average coefficient weights of the meteorological variables obtained from the GTWR-bi-square model, shown in Table 3, to illustrate the general relationship between meteorology and yield. GDD and PCPN had positive effects on yield. For every degree increase in GDD, the yield increased by 1 kg/ha, while for each millimeter increase in PCPN the yield increased by 2.2 kg/ha. However, the impact of KDD and VPD on yield was negative. Yield decreased by 1.7 or 0.6 kg/ha for every degree increase in KDD or every hectopascal increase in VPD. Coincidently, combined with the 40-year climate trends shown in Figure 4, we predicted that all four climate variables would have a positive impact on yield in the summer. This implies that nearly 40 years of climate change has had a positive impact on yields in the U.S. Corn Belt.

4.3. Explaining Yield Anomalies during Extreme Drought and Hot Climate Events

Historically, droughts have been major contributors to agricultural losses [53,54], and are often accompanied by high temperatures. Corn is especially vulnerable to climate change due to its high irrigation requirements. High temperatures are expected to increase water loss, making corn production susceptible to drought and high temperatures. Thus far, this study has confirmed that latitude and air temperature are closely related, with KDD having the strongest influence. We found that KDD and VPD were highly correlated, with a correlation coefficient of 0.91 in Table 2, indicating that high temperature is often accompanied by high VPD, which inhibits crop growth due to a lack of moisture.
The county-level results for GDD, KDD, VPD, PCPN and detrended yield in 1988 and 2012 are presented in Table 4. We used these results to determine how extreme drought and hot weather led to a sharp decrease in corn yield. The changes in meteorological variables in the summer for these two years were very similar. Compared with the normal climate, from the attribute of deviation ratio, GDD increased by around 10%, KDD increased by around 100%, VPD increased by around 40%, and PCPN decreased by around 35%, while the difference in detrended yield reduction was almost double. The spatial distributions of deviations in meteorological variables at the county level were quite different between 1988 and 2012, as shown in Figure 5.

Table 4. Summary of meteorological variables and detrended yield in 1988 and 2012 at the county level.

<table>
<thead>
<tr>
<th>Year</th>
<th>County</th>
<th>GDD</th>
<th>KDD</th>
<th>VPD</th>
<th>PCPN</th>
<th>De-Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Average over 40 years</td>
<td>Average for 1988</td>
<td>Average deviation</td>
<td>Deviation ratio</td>
<td>Average of coefficients</td>
</tr>
<tr>
<td>1988</td>
<td>1134</td>
<td>1311.5 °C</td>
<td>135.4 °C</td>
<td>1034.1 Pa</td>
<td>294.2 mm</td>
<td>6.33 tons/ha</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1311.5 °C</td>
<td>135.4 °C</td>
<td>1034.1 Pa</td>
<td>294.2 mm</td>
<td>6.33 tons/ha</td>
</tr>
<tr>
<td>2012</td>
<td>996</td>
<td>1316.5 °C</td>
<td>134.3 °C</td>
<td>1031.9 Pa</td>
<td>295.3 mm</td>
<td>6.33 tons/ha</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1316.5 °C</td>
<td>134.3 °C</td>
<td>1031.9 Pa</td>
<td>295.3 mm</td>
<td>6.33 tons/ha</td>
</tr>
</tbody>
</table>

Figure 5 shows that the deviation of GDD and KDD to the value of the average over 40 years was lower in most areas in 2012 than in 1988, while the marginal effect was stronger, from 0.3 kg/ha in 1988 to 2 kg/ha in 2012 for a 1 °C increase in GDD and 1.5 kg/ha in 1988 to −9 kg/ha in 2012 for a 1 °C increase in KDD, as shown in Table 4. In terms of GDD, the increase in temperature slightly boosted the yield, while in terms of KDD the extremely high temperature had a significant negative effect in 2012. As can be seen in Figure 6, although the northern regions experienced only slight increases in KDD in 2012, the negative impact was sizeable. This illustrates that higher latitudes experienced a greater impact due to extremely high temperatures, and further indicates that the southern regions had better heat resistance.

Figure 6 shows that the boundary effect of the PCPN coefficient was close to positive on a large scale, with the red part having more area than the blue part. For each additional milliliter of precipitation in Table 4, the yield in 1988 and 2012 increased by 6.1 kg/ha and 5.5 kg/ha, respectively. This indicates that precipitation and yield had a stable positive correlation. The impact of VPD on yield was relatively mild in terms of space in both 1988 and 2012, as shown in Figure 6. Then, as shown in Table 4, the average yield in 1988 and 2012 increased by −2.1 kg/ha and −1.4 kg/ha, respectively. The portion in the red circle in Figure 6 centered on Illinois, Wisconsin, and Iowa experienced a steady negative impact and was very sensitive to VPD. Nebraska and Kansas (blue circle), which are typically irrigated areas [13,55], experienced a very large increase in VPD in Figure 5, and only a small actual impact on yield in Figure 6, especially in 2012. Despite the increase in VPD, the marginal effect remained positive, indicating the significant positive effect of irrigation. Thus, the impact of VPD on yield can be alleviated manually; the effect of irrigation was evident.
Figure 5. Spatial distributions of deviations in meteorological variables in the summer months at the county level in 1988 and 2012.
In conclusion, the impact of temperature on yield increased between 1988 and 2020 and that of water declined, as it is more amenable to human intervention. Furthermore, we found that the impact of climate factors on yield has a threshold; even though the impact may not necessarily be negative, production is expected to decrease because of declining yields.

4.4. Temporal and Spatial Clustering Based on Climate Change Impact Trends

We performed fuzzy c-means clustering of the results of the GTWR-bi-square model to obtain the membership values of counties with different meteorological variables for each cluster. As a first step, we excluded 122 counties with more than 25% of the years missing. This left 1037 counties that met the requirements of the Mfuzz clustering method. In order to better understand the experimental results, we divided the effect of GDD on yield from 1981 to 2020 into four clusters, the effect of KDD into four clusters, the effect of VPD into five clusters, and the effect of PCPN into four clusters.

Figure 7 shows that the clustering result for GDD was distributed based on latitude, which was in line with our expectations. Cluster 3, which was located at the lowest latitude, showed relatively stable fluctuations of the boundary effect of yield; its absolute values were low, with most of the values in the range of ±0.01, as shown in Figure 8. Cluster 4, at the highest latitude, showed a clear positive trend between 2001 and 2009 in Figure 8, indicating that GDD played a significant role in promoting the yield in this region in this decade. Overall, Figure 8 shows that all four clusters had more positive than negative portions, although the impact trends across different historical periods were not similar. We found that the differences between these four regions further diverged over the past 20 years, from 2000 to 2020, in Figure 8, making the impact of GDD on yield more unsteady.
Figure 7. Four spatial clusters of county-level coefficients of GDD.

Figure 8. Impact trends of GDD from 1981 to 2020 for four clusters.

The impact of temperature-related factors was very latitude-dependent. The clustering results for KDD were completely different from those for GDD, although they showed latitude correlations, as shown in Figure 9. Cluster 2, with the highest overall latitude,
was more prone to frequent positive impact fluctuations when excluding a sudden large negative effect in 1992, as shown in Figure 10. Cluster 3, at the lowest latitude, had most coefficients in the range of $-0.02$ to $0.02$ in Figure 10, reflecting corn production more susceptible to negative effects on $KDD$. The effect of $KDD$ on yield in clusters 2 and 4, both of which were at high latitudes, varied widely and had values close to 0.2. The variation interval decreased as the latitude decreased, confirming that the lower latitude regions were better able to tolerate $KDD$ changes.

Figure 9. Four spatial clusters of county-level coefficients of $KDD$.

Figure 10. Impact trends of $KDD$ from 1981 to 2020 for four clusters.

The spatial clustering results of county-level $VPD$ coefficients, shown in Figure 11, were highly consistent with the distribution of irrigation districts in the U.S. Most irrigated
areas were spatially covered by clusters 2 and 4, which had \( VPD \) coefficients falling within the smallest variation range, as shown in Figure 12. This confirmed that the two regions affected by \( VPD \) fluctuated the least. Rainfall in the other three cluster areas was abundant, making irrigation unnecessary. These three clustered regions were strongly affected by \( VPD \) and were more susceptible to negative effects. Based on the changing trends over the past ten years shown in Figure 12, the regions in clusters 3 and 4 had reduced resistance to \( VPD \). For the regions in cluster 4, this loss in resistance may have been caused by a drop in the groundwater level, which could have caused irrigation problems. As the regions in cluster 3 are relatively far from both the Great Lakes and the Atlantic Ocean, their declining ability to withstand climate change is understandable. However, production conditions permitting, irrigation technology can be upgraded to supplement these areas.

It is clear that the spatial clustering results for \( PCPN \) clusters 2 and 4 in Figure 13 have corresponding regions in clusters 2 and 1 for \( VPD \) in Figure 11. This indicates that the moisture-related \( VPD \) and \( PCPN \) were highly correlated in these two regions. In Figure 14, cluster 2, which required irrigation, had a significant positive feedback effect on corn yield, especially from 1990 to 2010. Cluster 4, which has historically been largely negatively impacted by \( PCPN \), was positively impacted by \( PCPN \) over the past five years, even exceeding the positive impact of cluster 2. In addition, the variation interval in Clusters 2 and 4 were smaller than in the other two, illustrating that the high-latitude regions in the Corn Belt had a better tolerance for \( PCPN \) changes. In general, the magnitude of the impact faced by the four clusters has reduced in recent years. In contrast, the coefficient variation of \( PCPN \) was not as high as that of \( VPD \), which better represented the water-related impact on corn yield.

![Figure 11. Five spatial clusters of county-level coefficients of \( VPD \).](image-url)
Figure 12. Impact trends of VPD from 1981 to 2020 for five clusters.

Figure 13. Four spatial clusters of county-level coefficients of PCPN.
5. Discussion

This study aimed to discover the temporal and spatial relationships between meteorological variables and corn yield using county-level data from the U.S. Corn Belt. By implementing a spatiotemporal analytical framework based on the GTWR model and fuzzy c-means algorithm, we obtained immediate information. Climate factors had different effects on crop yield at different points in time and space, especially in extremely dry and hot climates in 1988 and 2012. There was a high degree of diversity in the trends with respect to climate change’s impact on yield in different regions. In the future, we can synthesize this information to innovatively discover and discuss these trends and their implications.

5.1. Optimistic Predictions for the Impact of Climate Change on the Corn Belt

First, from the perspective of the whole corn belt region, the impact of climate variables on corn yield showed a positive trend. The coefficients of the GTWR model were positive for GDD and PCPN, and the values increased over time. Although KDD and VPD had negative coefficients, their values decreased over time. Based on the trends in the climate variables in the summer from 1981 to 2020 in Section 4.1 in combination with the coefficients in Section 4.2, GDD promoted corn yield by 0.07 kg/ha, KDD by 2.96 kg/ha, VPD by 1.69 kg/ha, and PCPN by 1.58 kg/ha on average per year.

Fortunately, since 2012, climate change has been relatively mild and conditions have been amenable to the healthy growth of crops. This has resulted in an upward trend in future yield, with production set to increase by 6.30 kg/ha due to climate change.

Researchers have found that precipitation increases yield in nearly all crops and countries [38,46]. Furthermore, although yields tend to be reduced when temperatures rise, this is not necessarily the case in high-latitude countries, where crops particularly benefit from warming [11,35]. Our findings are consistent with previous studies in which crop production at high latitudes such as the Corn Belt in the northern U.S. appeared to benefit from a warming climate [11].
5.2. Temperature Has a Stronger Impact than Precipitation

The ways in which the effects of the climate variables changed over the past four decades year by year have been discussed in Section 5.1. We found that KDD provided the greatest yield boost to productivity, proving that temperature changes were more important than precipitation changes at both national and regional scales [11] as well as at the county scale.

We found evidence to show that the influence of temperature has been strengthening and that of precipitation has been decreasing, especially in extreme climates. For example, in Table 4 of Section 4.3, we showed that KDD caused a crop yield loss of 1165.5 kg/ha in 2012, the effect being far greater than in 1988. Meanwhile, the corn yield loss caused by PCPN slightly decreased, from 633.79 kg/ha in 1988 to 611.6 kg/ha in 2012.

In previous studies, corn has been found to be more vulnerable to climate change due to its high irrigation water requirements [56,57]. However, crop yields could increase by expanding irrigated areas or intensifying irrigation [13,58], reducing the influence of natural rainfall. In this study, we selected a mature and specialized corn-producing area where the irrigation levels were already very professional. Thus, the impact of precipitation on corn yield was lower than that of temperature in our study area.

Despite recent increases in irrigation, crop yield has been reduced due to a lack of rainfall [59]. In addition, the groundwater table, on which irrigation is partly dependent, has been shown to be falling [55,60]. This has made future impact predictions not very optimistic.

5.3. Better Production Potential in High Latitudes

The effects of climate change on crop yield are expected to vary across regions, with crop yield increasing in some regions and decreasing in others, depending on the latitude and irrigation requirements of the region [1].

In Section 4.4, Figure 8 shows that GDD had more positive effects than negative effects on yield in nearly all Corn Belt regions. However, as shown in Figures 9 and 10, the ability of several high-latitude counties, such as those in Wisconsin, Minnesota, North Dakota, and South Dakota, to withstand high temperatures was very low. The loss due to extreme heat in these regions is expected to be high. In contrast, regions at lower latitudes, such as those in Kentucky, Kansas, Missouri, and Southern Illinois, were more tolerant of extreme temperatures.

In terms of water, as shown in Figures 13 and 14, most high-latitude regions were less sensitive to precipitation. Overall, selecting more heat-tolerant corn varieties or implementing greenhouse cultivation, thereby reducing the sensitivity to KDD, may contribute to greater increases in corn yield in high-latitude regions. This would mean that regions at higher latitudes would have a stronger potential to increase production despite climate change compared to those at lower latitudes.

6. Conclusions and Future Works

We applied a spatiotemporal analytical framework largely based on the GTWR model to a sample of 1159 counties in the U.S. Corn Belt, covering 13 States from 1981 to 2020. The results showed how four main meteorological variables quantitatively and qualitatively impacted corn production.

The GTWR model, which combines all data into a single regression, requires little data preprocessing, such as data completion for counties with missing years; this helps to remove human interference from the analysis of spatiotemporal relationships. Compared to previous studies, the completeness and accuracy of the original unbalanced data allowed us to obtain a more accurate prediction of climate responses. By accounting for both spatial and temporal non-stationary relationships, we were able to efficiently and robustly handle the spatiotemporal heterogeneity of the climate data.

Using the powerful spatiotemporal regression tool of the GTWR model, we innovatively uncovered several relationships between climate change and corn yield that varied
over time and space. Temporally, the U.S. Corn Belt has benefited from the increase in temperature and precipitation over the past four decades, with the temperature effect being larger. Moreover, the impact of precipitation on yield decreased, whereas that of temperature increased during extreme climate events, mainly due to the use of artificial irrigation as a water supplement. Spatially, the impact of climate change on yield was closely related to latitude, with the production potential being higher in high-latitude areas.

Although we obtained large amounts of information from the spatiotemporal analytical model, the degree of information mining was not deep enough and the direction was not comprehensive. Meanwhile, this paper lacks simulations for projected climate change scenarios, which could have widened the information presented in the regression analysis with the historical meteorological data. In the future, we will try to find more patterns between climate change and crop yield. In addition, this model can be applied to other regions in order to further test its robustness and effectiveness.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijgi11080433/s1, Table S1: Summary of parameters of ordinary least square (OLS), time linear regression (LR), geographically weighted regression (GWR) with Gaussian kernel, GWR with the bi-square kernel, geographically and temporally weighted regression (GTWR) with Gaussian kernel, and GTWR with bi-square kernel models.

Author Contributions: Methodology, Bing Yang and Sensen Wu; validation, Bing Yang and Sensen Wu; formal analysis, Bing Yang; investigation, Bing Yang; data curation, Bing Yang; writing—original draft preparation, Bing Yang; writing—review and editing, Bing Yang, Sensen Wu and Zhen Yan; visualization, Bing Yang; supervision, Sensen Wu and Zhen Yan; project administration, Bing Yang; funding acquisition, Sensen Wu and Zhen Yan. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (42001323), National Key Research and Development Program of China (2018YFB0505000), Provincial Key R&D Program of Zhejiang (2021C01031), General Project of National Natural Science Foundation of China (72073118), and International Cooperation Project of National Natural Science Foundation of China (72161147001).

Data Availability Statement: Publicly available datasets were analyzed in this study. The corn yield data presented in this study are openly available in https://www.nass.usda.gov/ (accessed on 24 March 2021). The meteorological data can be found here: https://prism.oregonstate.edu (accessed on 1 August 2021).

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

References
5. Alahacoon, N.; Edirisinghe, M. Spatial Variability of Rainfall Trends in Sri Lanka from 1989 to 2019 as an Indication of Climate Change. ISPRS Int. J. Geo-Inf. 2021, 10, 84. [CrossRef]


49. Chen, Z.; Ma, X.; Wu, L.; Xie, Z. An Intuitionistic Fuzzy Similarity Approach for Clustering Analysis of Polygons. ISPRS Int. J. Geo-Inf. 2019, 8, 98. [CrossRef]


60. Onduña Alegria, M.E.; Schütze, N.; Niyoji, D. Evaluation of Hydroclimatic Variability and Prospective Irrigation Strategies in the U.S. Corn Belt. Water 2019, 11, 2447. [CrossRef]